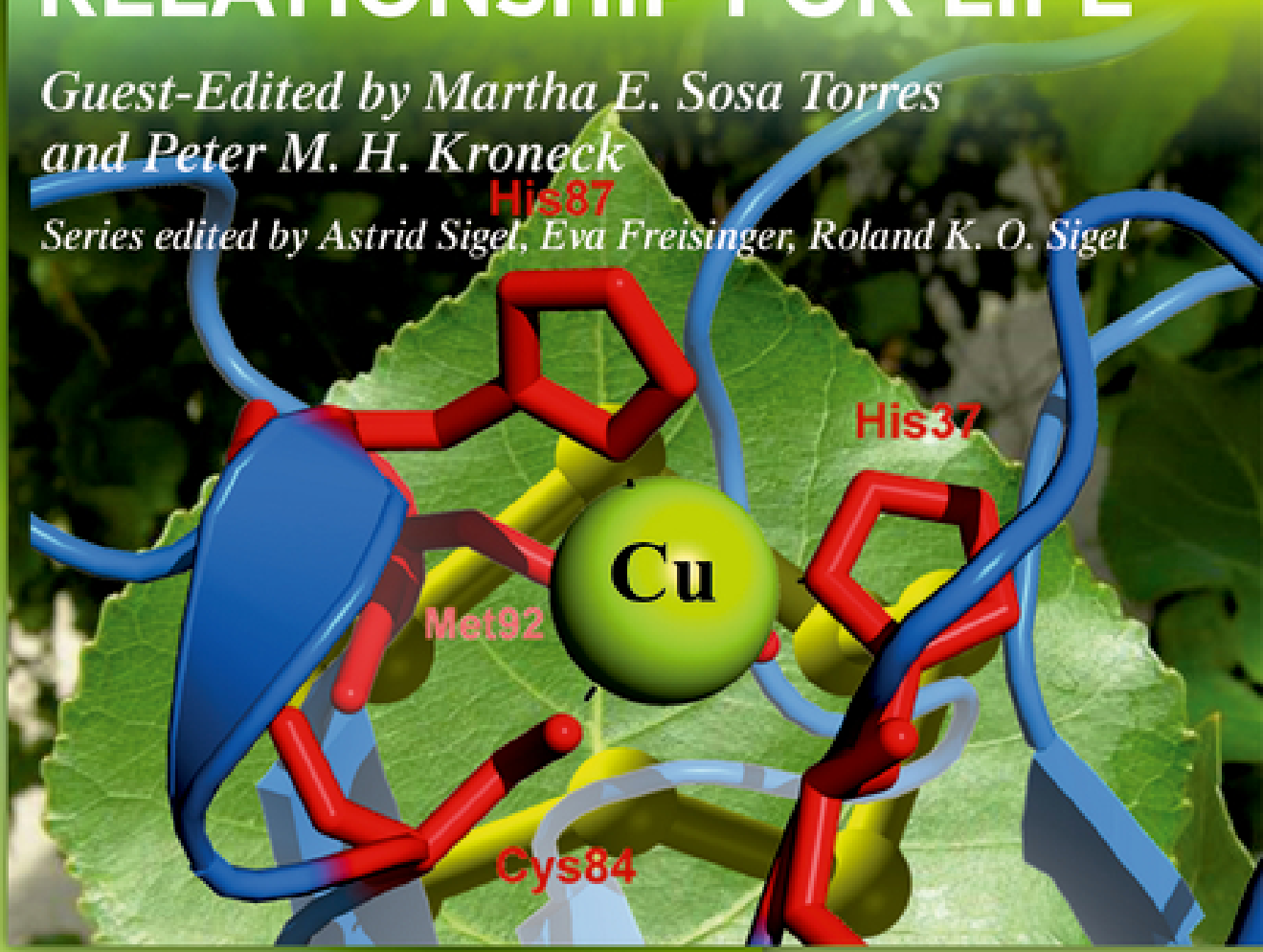
DE GRUYTER

TRANSITION METALS AND SULFUR: A STRONG RELATIONSHIP FOR LIFE

Guest-Edited by Martha E. Sosa Torres and Peter M. H. Kroneck

Series edited by Astrid Sigel, Eva Freisinger, Roland K. O. Sigel



METAL IONS IN LIFE SCIENCES 20

Astrid Sigel, Eva Freisinger, Roland K. O. Sigel
Metal Ions in Life Sciences 20

Metal Ions in Life Sciences

Edited by
Astrid Sigel, Eva Freisinger and Roland K. O. Sigel

Volume 20

Guest Editors:
Martha E. Sosa Torres and Peter M. H. Kroneck

Transition Metals and Sulfur: A Strong Relationship for Life

DE GRUYTER

Series Editors

Astrid Sigel
Department of Chemistry
Inorganic Chemistry
University of Basel
Spitalstrasse 51
CH-4056 Basel
Switzerland
<astrid.sigel@unibas.ch>

Eva Freisinger and Roland K. O. Sigel
Department of Chemistry
University of Zürich
Winterthurerstrasse 190
CH-8057 Zürich
Switzerland
<freisinger@chem.uzh.ch>
<roland.sigel@chem.uzh.ch>

Guest Editors

Martha E. Sosa Torres
Departamento de Química
Inorganica y Nuclear
Facultad de Química
Universidad Autónoma de México
Ciudad Universitaria
México, D.F. 04510
México
<mest@unam.mx>

Peter M. H. Kroneck
Fachbereich Biologie
Universität Konstanz
Universitätsstrasse 10
D-78457 Konstanz
Germany
<peter.kroneck@uni-konstanz.de>

ISBN 978-3-11-058889-7
e-ISBN (PDF) 978-3-11-058975-7
e-ISBN (EPUB) 978-3-11-058894-1
ISSN 1559-0836
e-ISSN 1868-0402
DOI: 10.1515/9783110588897

Library of Congress Control Number: 2019951922

Bibliographic information published by the Deutsche Nationalbibliothek
The Deutsche Nationalbibliothek lists this publication in the Deutsche Nationalbibliografie; detailed bibliographic data are available on the Internet at http://dnb.dnb.de.

Cover illustration: The figure was prepared by P. M. H. Kroneck; it is a derivative of “Cu binding site of plastocyanin from spinach (PDB 1AG6)” by Yikrazuul (licensed under CC BY-SA 3.0).

Typesetting: Meta Systems Publishing & Printservices GmbH, Wustermark, Germany
Printing and binding: CPI books GmbH, Leck

For further volumes: www.mils-WdG.com
www.degruyter.com

About the Editors

Martha E. Sosa Torres is Full Professor of Chemistry at the Universidad Nacional Autónoma de México (UNAM, Mexico). She received her Bachelor and Master Degree in Chemistry summa cum laude from UNAM, and her PhD from the University College London (U.K.) working with Martin L. Tobe. She did postdoctoral work with Martin Hughes, King's College London (U.K.), and she was a visiting professor at the University of Konstanz and the University of Freiburg in Germany. She received several prestigious fellowships from international institutions. In 2008 she organized the Summer School "Metal Ions in Biology Key-Elements of Life" in Cuernavaca (Mexico) sponsored by the German Academic Exchange Service (DAAD). Her research is directed towards the understanding of inorganic reaction mechanisms, with focus on dioxygen activation. Her group is interested in the spectroscopic, magnetic, and structural analysis of metal complexes, in the properties of metalloenzymes, and in Multi-Frequency Electron Paramagnetic Resonance. Furthermore, she explores multi-functional materials based on transition metal compounds including lanthanide complexes.

Peter M. H. Kroneck received his Diploma in Chemistry from the University of Basel (Switzerland) and his PhD from the University of Konstanz (Germany). He worked as a postdoctoral student with Jack T. Spence, Utah State University (Logan, USA), and he was a visiting professor with Helmut Beinert, University of Wisconsin (Madison, USA), with Israel Pecht, Weizmann Institute of Science (Rehovot, Israel), with William E. Antholine (National EPR Center, Medical College of Wisconsin Milwaukee, USA), and in 2012 he was the Robert Noyce Professor at Grinnell College (Grinnell, USA). He was the head of the Bioinorganic Research Group at the University of Konstanz, where he became a Professor of Biochemistry in 1989. In 2002 he received the Medal of the European Society of Biological Inorganic Chemistry; in 2009 he was the first Non-US scientist who chaired the Gordon Research Conference "Metals in Biology". His group's research activities have focused on structural and spectroscopic properties of copper, iron, molybdenum and tungsten enzymes and their functional roles in the biogeochemical cycles of nitrogen and sulfur.

Astrid Sigel has studied languages; she was an Editor of the *Metal Ions in Biological Systems* series (until Volume 44) and also of the "Handbook on Toxicity of Inorganic Compounds" (1988), the "Handbook on Metals in Clinical and

Analytical Chemistry" (1994; both with H. G. Seiler), and on the "Handbook on Metalloproteins" (2001; with Ivano Bertini). She is also an Editor of the *MILS* series from Volume 1 on and she coauthored about 50 papers on topics in Bioinorganic Chemistry.

Eva Freisinger is Associate Professor for Bioinorganic Chemistry and Chemical Biology (2018) at the Department of Chemistry at the University of Zürich, Switzerland. She obtained her doctoral degree (2000) from the University of Dortmund, Germany, working with Bernhard Lippert and spent three years as a postdoc at SUNY Stony Brook, USA, with Caroline Kisker. Since 2003 she performs independent research at the University of Zürich where she held a Förderungsprofessur of the Swiss National Science Foundation from 2008 to 2014. In 2014 she received her *Habilitation* in Bioinorganic Chemistry. Her research is focused on the study of plant metallothioneins with an additional interest in the sequence-specific modification of nucleic acids. Together with Roland Sigel she chaired the 12th European Bioinorganic Chemistry Conference (2014 in Zürich, Switzerland) as well as the 19th International Conference on Biological Inorganic Chemistry (2019 in Interlaken, Switzerland). She also serves on a number of Advisory Boards for international conference series; since 2014 she is the Secretary of the European Bioinorganic Chemistry Conferences (EuroBICs), and an Editorial Board Member of the *Journal of Inorganic Biochemistry.* She joined the group of editors of the *MILS* series from Volume 18 on.

Roland K. O. Sigel is Full Professor (2016) of Chemistry at the University of Zürich, Switzerland. In the same year he became Vice Dean of Studies (BSc/MSc) and in 2017 he was elected Dean of the Faculty of Science. From 2003 to 2008 he was endowed with a Förderungsprofessur of the Swiss National Science Foundation and he is the recipient of an ERC Starting Grant 2010. He received his doctoral degree summa cum laude (1999) from the University of Dortmund, Germany, working with Bernhard Lippert. Thereafter he spent nearly three years at Columbia University, New York, USA, with Anna Marie Pyle (now Yale University). During the six years abroad he received several prestigious fellowships from various sources, and he was awarded the EuroBIC Medal in 2008 and the Alfred Werner Prize (SCS) in 2009. 2015–2019 he was the Secretary of the Society of Biological Inorganic Chemistry (SBIC) and since 2018 he is the Secretary of the International Conferences of Biological Inorganic Chemistry (ICBICs). His research focuses on the structural and functional role of metal ions in ribozymes, especially group II introns, regulatory RNAs, and on related topics. He is also an Editor of Volumes 43 and 44 of the *MIBS* series and of the *MILS* series from Volume 1 on.

Historical Development and Perspectives of the Series *Metal Ions in Life Sciences*[*]

It is an old wisdom that metals are indispensable for life. Indeed, several of them, like sodium, potassium, and calcium, are easily discovered in living matter. However, the role of metals and their impact on life remained largely hidden until inorganic chemistry and coordination chemistry experienced a pronounced revival in the 1950s. The experimental and theoretical tools created in this period and their application to biochemical problems led to the development of the field or discipline now known as *Bioinorganic Chemistry*, *Inorganic Biochemistry*, or more recently also often addressed as *Biological Inorganic Chemistry.*

By 1970 *Bioinorganic Chemistry* was established and further promoted by the book series *Metal Ions in Biological Systems* founded in 1973 (edited by H. S., who was soon joined by A. S.) and published by Marcel Dekker, Inc., New York, for more than 30 years. After this company ceased to be a family endeavor and its acquisition by another company, we decided, after having edited 44 volumes of the *MIBS* series (the last two together with R. K. O. S.) to launch a new and broader minded series to cover today's needs in the *Life Sciences.* Therefore, the Sigels' new series is entitled

Metal Ions in Life Sciences.

After publication of 16 volumes (since 2006) with various publishers during the past 10 years, we are happy to join forces (from Volume 17 on) in this still growing endeavor with Walter de Gruyter GmbH, Berlin, Germany, a most experienced Publisher in the *Sciences.*

The development of *Biological Inorganic Chemistry* during the past 60 years was and still is driven by several factors; among these are (i) attempts to reveal the interplay between metal ions and hormones or vitamins, etc., (ii) efforts regarding the understanding of accumulation, transport, metabolism, and toxicity of metal ions, (iii) the development and application of metal-based drugs,

---

[*] Reproduced with some alterations by permission of John Wiley & Sons, Ltd., Chichester, UK (copyright 2006) from pages v and vi of Volume 1 of the series *Metal Ions in Life Sciences* (MILS-1).

(iv) biomimetic syntheses with the aim to understand biological processes as well as to create efficient catalysts, (v) the determination of high-resolution structures of proteins, nucleic acids, and other biomolecules, (vi) the utilization of powerful spectroscopic tools allowing studies of structures and dynamics, and (vii), more recently, the widespread use of macromolecular engineering to create new biologically relevant structures at will. All this and more is reflected in the volumes of the series *Metal Ions in Life Sciences.*

The importance of metal ions to the vital functions of living organisms, hence, to their health and well-being, is nowadays well accepted. However, in spite of all the progress made, we are still only at the brink of understanding these processes. Therefore, the series *Metal Ions in Life Sciences* links coordination chemistry and biochemistry in their widest sense. Despite the evident expectation that a great deal of future outstanding discoveries will be made in the interdisciplinary areas of science, there are still "language" barriers between the historically separate spheres of chemistry, biology, medicine, and physics. Thus, it is one of the aims of this series to catalyze mutual "understanding".

It is our hope that *Metal Ions in Life Sciences* continues to prove a stimulus for new activities in the fascinating "field" of *Biological Inorganic Chemistry.* If so, it will well serve its purpose and be a rewarding result for the efforts spent by the authors.

Astrid Sigel and Helmut Sigel
Department of Chemistry, Inorganic Chemistry
University of Basel, CH-4056 Basel, Switzerland

Roland K. O. Sigel
Department of Chemistry
University of Zürich, CH-8057 Zürich, Switzerland

October 2005
and September 2016

Preface to Volume 20

Transition Metals and Sulfur: A Strong Relationship for Life

Transition metal-sulfur sites play key roles in biology. In general, many transition metal ions – also often called trace elements – are essential to all living organisms. Notably, the number of these essential elements increased over the years, and the list of their biological functions has also grown steadily. Remarkable progress has been made in understanding the chemistry operating at the biological sites, furthermore early on it became likely that many transition metal ions harbor sulfur ligands in the metal coordination sphere. The structure and the reactivity of these metal-sulfur centers was quite different from anything found in regular coordination complexes. Yet, biomimetic inorganic chemistry, modern spectroscopic techniques, and high resolution protein crystallography made important contributions to unravel the properties of the metal-sulfur sites, with iron-sulfur and blue copper proteins as prime examples. In this volume of *Metal Ions in Life Sciences* the focus will be on some of the most intriguing, in our view, biological transition metal-sulfur sites. These include the blue type 1 Cu and the purple mixed-valent CuA electron transfer sites, the tetranuclear CuZ catalytic center of nitrous oxide reductase, the heme-thiolate complex in cytochrome P_{450}, the iron-sulfur proteins with bound inorganic and organic sulfur, the pterin dithiolene cofactor coordinated to Mo or W, the polynuclear metal clusters of nitrogenase, the coupled siroheme-[4Fe-4S] center of sulfite reductase, the NiFeS sites of hydrogenases and CO dehydrogenase, and the Zn-finger domains. We are fully aware of the excellent books and authoritative reviews on various aspects of the subject, however, it is our motivation to cover in one single volume source this exciting domain of bioinorganic chemistry.

In *Chapter 1* the reader is introduced to the general subject "What is so exciting about transition metal sulfur sites in biology", including a backward glance

Metal Ions in Life Sciences, Volume 20 Guest Editors: Martha E. Sosa Torres and Peter M. H. Kroneck
Series Editors: Astrid Sigel, Eva Freisinger, and Roland K. O. Sigel

https://doi.org/10.1515/9783110589757-202

to the early days of bioinorganic chemistry and the exploration of transition metal sulfur sites. Classically, metal ions are inorganic in nature, they have had a long and distinguished history at the interface of chemistry and biology, nowadays referred to as bioinorganic chemistry. The focus of *Chapter 2* is on the nonmetallic element sulfur. Important inorganic forms in the biosphere are elemental sulfur, sulfate, and sulfide. Because of its range of stable oxidation states (+6 to –2), sulfur plays important roles in central biochemistry and is intimately related to life on Earth. It is present in all major classes of biomolecules, such as enzymes, proteins, sugars, nucleic acids, vitamin cofactors, and metabolites. Transition metal sulfur complexes display intriguing catalytic activities. They provide surfaces and complex cavities in metalloenzymes that activate H_2, CO, N_2 or N_2O, in multi-electron, multi-proton transfer reactions. Given sulfur's diverse properties, evolution made an excellent choice in selecting sulfur as one of the basic elements of life.

Chapter 3 is devoted to one of the most studied metal sites in biology, the blue type 1 Cu center. The first coordination sphere of these Cu sites hosts two histidine (His) and one cysteine (Cys) residues, with a short Cu–S bond, as ligands in a trigonal geometry, that renders this unique Cu center an efficient and tunable electron transfer site employed in various important redox processes. Next, the binuclear mixed-valent CuA site is discussed in *Chapter 4*. CuA acts as an electron transfer hub in the terminal redox enzymes cytochrome *c* oxidase and nitrous oxide reductase. Its complex electronic structure is coupled to its capability as highly efficient electron transfer agent used in energy conservation. The insights provided by applying advanced spectroscopic techniques to native and bioengineered CuA centers are reviewed in an authoritative manner. In *Chapter 5* the tetranuclear copper sulfide CuZ center of nitrous oxide reductase is discussed. This key enzyme of the biogeochemical nitrogen cycle catalyzes the reduction of the greenhouse gas N_2O to N_2 and H_2O. There exist two forms, CuZ(4Cu1S) and CuZ(4Cu2S), leading to considerable differences in their spectroscopic and catalytic properties which are reviewed in detail. The catalytic cycle of nitrous oxide reductase has been explored, and one of the intermediates, CuZ^o, being in the $[1Cu^{2+}\text{-}3Cu^{1+}]$ oxidation state, becomes rapidly reduced by mixed-valent CuA to $[4Cu^{1+}]$. The complex nature of the tetranuclear CuZ center has posed several questions concerning its assembly and its active form *in vivo*.

Cytochromes P_{450} (CYPs) are in the focus of *Chapter 6*. Among nature's most versatile catalysts, these heme-thiolate enzymes participate in countless essential life processes. They exist in all domains of life, Bacteria, Archaea, and Eukarya, and even in viruses. CYPs fight drugs, poisonous compounds in plants, carcinogens formed during cooking, and environmental pollutants, the first line of defense to detoxify and solubilize poisonous substances by reacting them with O_2. Iron is proximally coordinated by a thiolate sulfur, and this ligation state represents the active form of the enzyme. CYPs share a sophisticated catalytic cycle that involves high-valent Fe(IV) species as key intermediates, with characteristic properties, crucial for the cleavage of strong C–H bonds.

Chapter 7 centers on iron-sulfur clusters, ubiquitous protein cofactors composed of Fe, inorganic, and organic sulfur ligands. They belong to the most an-

cient metal sites and may have been crucial for the evolution of life on Earth. They participate in many central metabolic processes like nitrogen fixation, respiration and DNA processing and repair. Organisms evolved various machineries to synthesize and then transfer iron-sulfur clusters into target proteins which is discussed in depth. Genome analysis allowed the prediction of the proteins containing iron-sulfur clusters across a broad variety of living organisms, establishing links between the size and composition of iron-sulfur proteomes and the types of organisms that encode them. Given their key functions in metabolism, dysfunctions of mutations in iron-sulfur proteins, or in proteins participating in iron-sulfur cluster biogenesis, are associated with serious human diseases. In *Chapter 8* recent exciting developments in the field of biological nitrogen fixation and nitrogenase are summarized. This enzyme mediates the reductive cleavage of the stable triple bond of gaseous N_2 at ambient conditions of temperature and pressure, driven by the hydrolysis of ATP, to yield bioavailable ammonium. At the core of nitrogenase resides a complex, iron-sulfur based cofactor that in most variants of the enzyme contains a heterometal, Mo or V, homocitrate coordinated to this heterometal, and the interstitial carbide. Spectroscopic studies helped in trapping and identifying reaction intermediates, and inhibitor- or intermediate-bound structures of the cofactors were characterized by X-ray crystallography.

The last 20 years have seen a dramatic increase in the mechanistic understanding of the reactions catalyzed by pyranopterin Mo and W enzymes which are in the focus of *Chapter 9*. These enzymes possess pyranopterin ene-1,2-dithiolate (Moco), a novel ligand in bioinorganic chemistry. A synopsis of its biosynthesis and structure is presented, along with the current understanding of the role Moco plays in enzymatic catalysis. Oxygen atom transfer (OAT) reactivity is discussed in terms of breaking strong metal-oxo bonds and the mechanism of OAT catalyzed by enzymes of the sulfite oxidase (SO) family that possess dioxo active sites. The hallmark of sulfite reductase, discussed in *Chapter 10*, is its catalytic center made of an iron-containing porphyrinoid called siroheme that is covalently coupled to a [4Fe-4S] cluster through a cysteine bridge. Siroheme is an isobacteriochlorin that is more readily oxidized than protoporphyin IX-derived hemes. The chapter summarizes (i) how microbes use sulfite reductase to survive in a range of ecosystems, (ii) how atomic-resolution structures of dissimilatory and assimilatory sulfite reductases reveal their ancient homology, (iii) how the coupled siroheme-[4Fe-4S] cluster catalyzes the six-electron reduction of sulfite to sulfide, and (iv) how siroheme is synthesized across diverse microorganisms. In *Chapter 11*, three enzymes relying on the interplay of nickel, iron, and sulfur in their active sites, [NiFe]-hydrogenases, Ni,Fe-containing carbon monoxide dehydrogenases, and acetyl-CoA synthases, are reviewed. Prokaryotes use these enzymes to catalyze reactions driving the global carbon and hydrogen cycles. Although their active sites have different compositions and assemble Ni, Fe, and S in different ways and for different purposes, they share a central role of Ni in substrate binding and activation, with sulfur linking the Ni ion to one or more Fe ions. A short overview on the properties of the three individual enzymes highlights their parallels and differences.

Finally, in *Chapter 12* the zinc-finger (ZF) motif is discussed. It represents the majority of the DNA-binding domains, including 3 % of the human genes and, besides its DNA recognition function, it is involved in many diverse processes, such as RNA packaging, transcriptional activation, regulation of apoptosis, protein folding and assembly, and lipid binding. While the amino acid composition varies from one domain to the other, a shared feature is the coordination of a structural Zn(II) ion, by a different combination of cysteine and histidine ligands, with Zn-Cys_2His_2 representing the classical finger. The current knowledge on the main classes of eukaryotic and prokaryotic ZFs is summarized, focusing on the role of the Zn(II) ion, the folding mechanism, and the DNA binding. The hypothesis of a horizontal gene transfer from prokaryotes to eukaryotes is also discussed.

In conclusion, volume 20 of the *Metal Ions in Life Sciences* series offers a wealth of in-depth information about the structural and functional properties of transition metal-sulfur sites in living organisms. Significant advances in our understanding of these important metal centers have been achieved by the application of powerful spectroscopic and biochemical techniques including the use of molecular engineering to create new biologically relevant structures, to manipulate very large biomolecules, and to design artificial metalloenzymes.

Martha E. Sosa Torres
Peter M. H. Kroneck

Contents

Contributors to Volume 20

Numbers in parentheses indicate the pages on which the authors' contributions begin.

Susana L. A. Andrade Institute for Biochemistry, Albert-Ludwigs-University Freiburg, D-79104 Freiburg im Breisgau, Germany (257)

Claudia Andreini Center for Magnetic Resonance, University of Florence, I-50019 Sesto Fiorentino, Italy <andreini@cerm.unifi.it> (199)

Trinidad Arcos-López Department of Chemistry, Center for Research and Advanced Studies (Cinvestav), México City, D.F. 07360, México (51)

Isabel Askenasy Department of Biological Science, Florida State University, Tallahassee, FL 32306-4370, USA. Current Address: Department of Biomolecular Chemistry, University of Wisconsin School of Medicine and Public Health, 440 Henry Mall, Biochemical Sciences Building, Room 4206C, Madison, WI 53706, USA (343)

Ilaria Baglivo Dipartimento di Scienze e Tecnologie Ambientali, Biologiche e Farmaceutiche, Università della Campania "L. Vanvitelli", I-81100 Caserta, Italy (415)

Marta S. P. Carepo Biological Chemistry Lab, LAQV, REQUIMTE, Departamento de Química, Faculdade de Ciências e Tecnologia, Universidade Nova de Lisboa, PT-2829-516 Caparica, Portugal <marta.carepo@fct.unl.pt> (139)

F. Miguel Castro Martinez Departamento de Física y Química Teórica, Facultad de Química, Universidad Nacional Autónoma de México, Ciudad Universitaria, México, D.F. 04510, México <miguel.castro.m@gmail.com> (165)

Simone Ciofi-Baffoni Center for Magnetic Resonance, University of Florence, I-50019 Sesto Fiorentino, Italy <ciofi@cerm.unifi.it> (199)

Gianluca D'Abrosca Dipartimento di Scienze e Tecnologie Ambientali, Biologiche e Farmaceutiche, Università della Campania "L. Vanvitelli", I-81100 Caserta, Italy (415)

Ivana Djurdjevic Institute for Biochemistry, Albert-Ludwigs-University Freiburg, D-79104 Freiburg im Breisgau, Germany (257)

Holger Dobbek Institut für Biologie, Humboldt-Universität zu Berlin, D-10115 Berlin, Germany <holger.dobbek@biologie.hu-berlin.de> (381)

Oliver Einsle Institute for Biochemistry, Albert-Ludwigs-University Freiburg, D-79104 Freiburg im Breisgau, Germany
<einsle@biochemie.uni-freiburg.de> (257)

Roberto Fattorusso Dipartimento di Scienze e Tecnologie Ambientali, Biologiche e Farmaceutiche, Università della Campania "L. Vanvitelli", I-81100 Caserta, Italy (415)

Jakob Gies Institute for Biochemistry, Albert-Ludwigs-University Freiburg, D-79104 Freiburg im Breisgau, Germany (257)

Katharina Grunau Institute for Biochemistry, Albert-Ludwigs-University Freiburg, D-79104 Freiburg im Breisgau, Germany (257)

Yulia Ilina Institut für Biologie, Humboldt-Universität zu Berlin, D-10115 Berlin, Germany (381)

Carla Isernia Dipartimento di Scienze e Tecnologie Ambientali, Biologiche e Farmaceutiche, Università della Campania "L. Vanvitelli", I-81100 Caserta, Italy <carla.isernia@unicampania.it> (415)

Jae-Hun Jeoung Institut für Biologie, Humboldt-Universität zu Berlin, D-10115 Berlin, Germany (381)

Khadanand KC Department of Chemistry and Chemical Biology, The University of New Mexico, Albuquerque, NM 87131-0001, USA <knkc@unm.edu> (313)

Martin L. Kirk Department of Chemistry and Chemical Biology, The University of New Mexico, Albuquerque, NM 87131-0001, USA <mkirk@unm.edu> (313)

Peter M. H. Kroneck Department of Biology, University of Konstanz, D-78457 Konstanz, Germany <peter.kroneck@uni-konstanz.de> (1, 19, 165)

Gaetano Malgieri Dipartimento di Scienze e Tecnologie Ambientali, Biologiche e Farmaceutiche, Università della Campania "L. Vanvitelli", I-81100 Caserta, Italy (415)

Berta M. Martins Institut für Biologie, Humboldt-Universität zu Berlin, D-10115 Berlin, Germany (381)

Marcos N. Morgada Instituto de Biología Molecular y Celular de Rosario (IBR), Departamento de Química Biológica, Facultad de Ciencias Bioquímicas y Farmacéuticas, Universidad Nacional de Rosario-CONICET, Rosario, Argentina <morgada@ibr-conicet.gov.ar> (91)

Isabel Moura Biological Chemistry Lab, LAQV, REQUIMTE, Departamento de Química, Faculdade de Ciências e Tecnologia, Universidade Nova de Lisboa, PT-2829-516 Caparica, Portugal <isabelmoura@fct.unl.pt> (139)

Daniel H. Murgida Instituto de Química Física de los Materiales, Medio Ambiente y Energía (INQUIMAE), Departamento de Química Inorgánica, Analítica y Química Física, Facultad de Ciencias Exactas y Naturales, Universidad de Buenos Aires CONICET, Buenos Aires, Argentina (91)

R. Daniel Páez López Departamento de Química Inorgánica y Nuclear, Facultad de Química, Universidad Nacional Autónoma de México, Ciudad Universitaria, México, D.F. 04510, México, <richie_daniels@comunidad.unam.mx> (165)

Sofia R. Pauleta UCIBIO, REQUIMTE, Departamento de Química, Faculdade de Ciências e Tecnologia, Universidade Nova de Lisboa, Campus da Caparica, PT-2829-516 Caparica, Portugal <srp@fct.unl.pt> (139)

Paolo V. Pedone Dipartimento di Scienze e Tecnologie Ambientali, Biologiche e Farmaceutiche, Università della Campania "L. Vanvitelli", I-81100 Caserta, Italy (415)

Liliana Quintanar Department of Chemistry, Center for Research and Advanced Studies (Cinvestav), México City, D.F. 07360, México <lilianaq@cinvestav.mx> (51)

Alfonso Rito Morales Departamento de Química Inorgánica y Nuclear, Facultad de Química, Universidad Nacional Autónoma de México, Ciudad Universitaria, México, D.F. 04510, México <ponchori@comunidad.unam.mx> (19)

Michael Rohde Institute for Biochemistry, Albert-Ludwigs-University Freiburg, D-79104 Freiburg im Breisgau, Germany (257)

Luigi Russo Dipartimento di Scienze e Tecnologie Ambientali, Biologiche e Farmaceutiche, Università della Campania "L. Vanvitelli", I-81100 Caserta, Italy (415)

Pedro D. Sarmiento Pavía Departamento de Química Inorgánica y Nuclear, Facultad de Química, Universidad Nacional Autónoma de México, Ciudad Universitaria, México, D.F. 04510, México <pedrodavids@comunidad.unam.mx> (165)

Florian Schneider Institute for Biochemistry, Albert-Ludwigs-University Freiburg, D-79104 Freiburg im Breisgau, Germany (257)

Nils Schuth Department of Chemistry, Center for Research and Advanced Studies (Cinvestav), México City, D.F. 07360, México (51)

Alejandro Solano Peralta Departamento de Química Inorgánica y Nuclear, Facultad de Química, Universidad Nacional Autónoma de México, Ciudad Universitaria, México, D.F. 04510, México <alexsolano@unam.mx> (19)

Martha E. Sosa Torres Departamento de Química Inorgánica y Nuclear, Facultad de Química, Universidad Nacional Autónoma de México, Ciudad Universitaria, México, D.F. 04510, México <mest@unam.mx> (1, 19, 165)

M. Elizabeth Stroupe Department of Biological Science and Institute of Molecular Biophysics, 91 Chieftain Way, Florida State University, Tallahassee, FL 32306-4370, USA <mestroupe@bio.fsu.edu> (343)

Christian Trncik Institute for Biochemistry, Albert-Ludwigs-University Freiburg, D-79104 Freiburg im Breisgau, Germany (257)

Alejandro J. Vila Instituto de Biología Molecular y Celular de Rosario (IBR), Departamento de Química Biológica, Facultad de Ciencias Bioquímicas y Farmacéuticas, Universidad Nacional de Rosario-CONICET, Rosario, Argentina <vila@ibr-conicet.gov.ar> (91)

Titles of Volumes 1–44 in the *Metal Ions in Biological Systems* Series

edited by the SIGELs
and published by Dekker/Taylor & Francis (1973–2005)

Volume 1: **Simple Complexes**
Volume 2: **Mixed-Ligand Complexes**
Volume 3: **High Molecular Complexes**
Volume 4: **Metal Ions as Probes**
Volume 5: **Reactivity of Coordination Compounds**
Volume 6: **Biological Action of Metal Ions**
Volume 7: **Iron in Model and Natural Compounds**
Volume 8: **Nucleotides and Derivatives: Their Ligating Ambivalency**
Volume 9: **Amino Acids and Derivatives as Ambivalent Ligands**
Volume 10: **Carcinogenicity and Metal Ions**
Volume 11: **Metal Complexes as Anticancer Agents**
Volume 12: **Properties of Copper**
Volume 13: **Copper Proteins**
Volume 14: **Inorganic Drugs in Deficiency and Disease**
Volume 15: **Zinc and Its Role in Biology and Nutrition**
Volume 16: **Methods Involving Metal Ions and Complexes in Clinical Chemistry**
Volume 17: **Calcium and Its Role in Biology**
Volume 18: **Circulation of Metals in the Environment**
Volume 19: **Antibiotics and Their Complexes**
Volume 20: **Concepts on Metal Ion Toxicity**
Volume 21: **Applications of Nuclear Magnetic Resonance to Paramagnetic Species**
Volume 22: **ENDOR, EPR, and Electron Spin Echo for Probing Coordination Spheres**
Volume 23: **Nickel and Its Role in Biology**
Volume 24: **Aluminum and Its Role in Biology**
Volume 25: **Interrelations Among Metal Ions, Enzymes, and Gene Expression**
Volume 26: **Compendium on Magnesium and Its Role in Biology, Nutrition, and Physiology**
Volume 27: **Electron Transfer Reactions in Metalloproteins**

Volume 28: **Degradation of Environmental Pollutants by Microorganisms and Their Metalloenzymes**
Volume 29: **Biological Properties of Metal Alkyl Derivatives**
Volume 30: **Metalloenzymes Involving Amino Acid-Residue and Related Radicals**
Volume 31: **Vanadium and Its Role for Life**
Volume 32: **Interactions of Metal Ions with Nucleotides, Nucleic Acids, and Their Constituents**
Volume 33: **Probing Nucleic Acids by Metal Ion Complexes of Small Molecules**
Volume 34: **Mercury and Its Effects on Environment and Biology**
Volume 35: **Iron Transport and Storage in Microorganisms, Plants, and Animals**
Volume 36: **Interrelations Between Free Radicals and Metal Ions in Life Processes**
Volume 37: **Manganese and Its Role in Biological Processes**
Volume 38: **Probing of Proteins by Metal Ions and Their Low-Molecular-Weight Complexes**
Volume 39: **Molybdenum and Tungsten. Their Roles in Biological Processes**
Volume 40: **The Lanthanides and Their Interrelations with Biosystems**
Volume 41: **Metal Ions and Their Complexes in Medication**
Volume 42: **Metal Complexes in Tumor Diagnosis and as Anticancer Agents**
Volume 43: **Biogeochemical Cycles of Elements**
Volume 44: **Biogeochemistry, Availability, and Transport of Metals in the Environment**

Contents of Volumes in the *Metal Ions in Life Sciences* Series

edited by the SIGELs

Volumes 1–4
published by John Wiley & Sons, Ltd., Chichester, UK (2006–2008)
<http://www.Wiley.com/go/mils>
<http://www.wiley.com/WileyCDA/Section/id-300350.html>

Volumes 5–9
published by the Royal Society of Chemistry, Cambridge, UK (2009–2011)
since 2015 by Walter de Gruyter GmbH, Berlin, Germany
<http://www.bioinorganic-chemistry.org/mils> <http://www.mils-WdG.com>

Volumes 10–16
published by Springer Science & Business Media BV, Dordrecht, The Netherlands (2012–2014; MILS-10 to MILS-14)
and by Springer International Publishing AG, Cham, Switzerland (2015–2016; MILS-15 and MILS-16)
<http://www.bioinorganic-chemistry.org/mils>

and from Volume 17 on
published by Walter de Gruyter GmbH, Berlin, Germany
<http://www.mils-WdG.com>

Volume 1: Neurodegenerative Diseases and Metal Ions

Volume 2: Nickel and Its Surprising Impact in Nature

Volume 3: The Ubiquitous Roles of Cytochrome P450 Proteins

Volume 4: Biomineralization. From Nature to Application

Volume 5: Metallothioneins and Related Chelators

Volume 6: Metal-Carbon Bonds in Enzymes and Cofactors

Volume 7: Organometallics in Environment and Toxicology

Volume 9: Structural and Catalytic Roles of Metal Ions in RNA

Volume 10: Interplay between Metal Ions and Nucleic Acids

Volume 11: Cadmium: From Toxicity to Essentiality

Volume 12: Metallomics and the Cell
Guest Editor: Lucia Banci

Volume 13: Interrelations between Essential Metal Ions and Human Diseases

Volume 16: The Alkali Metal Ions: Their Roles for Life

Volume 17: Lead: Its Effects on Environment and Health

Volume 22: Metal Ions in Bio-Imaging Techniques
(tentative)

Comments and suggestions with regard to contents, topics, and the like for future volumes of the series are welcome.

1

Introduction: Transition Metals and Sulfur

Martha E. Sosa Torres [1] ***and Peter M. H. Kroneck*** [2]

[1]Departamento de Química Inorgánica y Nuclear, Facultad de Química,
Universidad Nacional Autónoma de México, Ciudad Universitaria, México, D.F. 04510, México
<mest@unam.mx>

[2]Department of Biology, University of Konstanz, D-78457 Konstanz, Germany
<peter.kroneck@uni-konstanz.de>

Abstract: The number of transition metal ions which are essential to life – also often called trace elements – increased steadily over the years. In parallel, the list of biological functions in which transition metals are involved, has grown, and is still growing tremendously. Significant progress has been made in understanding the chemistry operating at the biological sites where metal ions have been discovered. Early on, based on the application of physical, chemical, and biological techniques, it became likely that numerous of these metal centers carry sulfur ligands in their coordination sphere, such as sulfide (S^{2-}), cysteine (RS^{-}), or methionine ($RSCH_3$). Notably, the structure and the reactivity of the metal active sites turned out to be quite different from anything previously observed in simple coordination complexes. Consequently, the prediction of active-site structures, based on known properties of transition metal ion complexes, turned out to be difficult and incorrect in many cases. Yet, biomimetic inorganic chemistry, via synthesis and detailed structural and electronic characterization of synthetic analogues, became an impor-

Metal Ions in Life Sciences, Volume 20 Guest Editors: Martha E. Sosa Torres and Peter M. H. Kroneck
Series Editors: Astrid Sigel, Eva Freisinger, and Roland K. O. Sigel

https://doi.org/10.1515/9783110589757-001

tant factor and helped to understand the properties of the metal active sites. Striking advances came from molecular biology techniques and protein crystallography, as documented by the publication of the first high-resolution structures of iron-sulfur proteins and the blue copper protein plastocyanin approximately five decades ago. In this volume of *METAL IONS IN LIFE SCIENCES* the focus will be on some of the most intriguing, in our view, transition metal-sulfur sites discovered in living organisms. These include the type 1 Cu mononuclear center, the purple mixed-valent $[Cu^{1.5+}\text{-}(Cys)_2\text{-}Cu^{1.5+}]$ CuA, the tetranuclear copper-sulfide catalytic center of nitrous oxide reductase, the heme-thiolate site in cytochrome P_{450}, the iron-sulfur proteins with bound inorganic (S^{2-}) and organic (Cys^{-}) sulfur, the pterin dithiolene cofactor (Moco) coordinated to either molybdenum or tungsten, the [8Fe-7S] P-cluster and the [Mo-7Fe-9S-C]-homocitrate catalytic site of nitrogenase, the siroheme-[4Fe-4S] center involved in the reduction of sulfite (SO_3^{2-}) to hydrogen sulfide (H_2S), the NiFeS sites of hydrogenases and CO dehydrogenase, and the zinc finger domains. We apologize to all researchers and their associates who have made tremendous contributions to our current knowledge of the steadily increasing transition metal sulfur sites in proteins and enzymes but are not mentioned here. These omissions are by no means intentional but merely the consequence of time and space. We are fully aware of the excellent books and authoritative reviews on various aspects of the subject, however, it is our motivation to cover in one single volume this exciting domain of bioinorganic chemistry.

Keywords: bioinorganic chemistry · cluster · copper · great oxidation event · iron· iron-sulfur protein· sulfide · sulfur · trace element · transition metal complex

1. A BACKWARD GLANCE − BIOINORGANIC CHEMISTRY AND THE EXPLORATION OF TRANSITION METAL SULFUR SITES

As the guest-editors we had to come to grips with the question "What is so exciting about transition metal sulfur sites in biology" in getting together the list of potential topics and authors for this volume of the *METAL IONS IN LIFE SCIENCES* (MILS) series. Transition metal ions that are classically inorganic in nature, have had a long and distinguished history at the interface of chemistry and biology, which is nowadays referred to by most as bioinorganic chemistry, a vibrant discipline at the interface of chemistry and the biological sciences [1–25]. Significant advances in bioinorganic chemistry since the middle of last century have been achieved by (i) the application of powerful spectroscopic techniques for studies of both structural and dynamic aspects [1, 10, 12, 18], (ii) the widespread use of molecular engineering to create new biologically relevant structures, to manipulate very large biomolecules at will, and to design artificial metalloenzymes [16, 26–28], (iii) the rapid determination of high-resolution structures of proteins including the structures of large membrane-bound respiratory complexes [29–31], and (iv) more recently the tremendous progress of computational chemistry [32, 33].

In his article entitled "Bioinorganic Chemistry: A New Field or Discipline? Words, Meanings, and Reality" pioneering researcher Helmut Beinert wrote: "To the uninitiated it may seem that bioinorganic chemistry is a new field, although, reports on metals bound to proteins or enzymes date way back into the 19th century and may probably be found in earlier centuries if we replace the terms *proteins* and *enzymes* with *animal or plant tissues.* The term *organic* meant that

it has to do with life compared to the counterpart *inorganic* then referred to lifeless matter" [11]. Obviously, the subject of bioinorganic chemistry promised to be a fruitful area of research for many years to come. As a consequence, the interest in bringing together scientists from various disciplines grew rapidly. Prime examples, to name a few, are: the symposium on *Copper in Biological Systems*, held in 1965 at Arden House, New York, USA [10], the *Manziana* copper conferences initiated by Bruno Mondovì, Jack Peisach, and Bo Malmström until 1995 close to Rome, Italy [11–13], the *International Conferences on Biological Inorganic Chemistry* (ICBIC) founded by Ivano Bertini and colleagues in 1983 in Florence, Italy [34], since then taking place in different countries around the world until today, and last but not least the Gordon Research Conferences *Metals in Biology* (GRC MIB). Already initiated in 1962 and still active on an annual basis in Ventura, USA, GRC MIB addresses long-standing, central challenges and emerging trends in the field of bioinorganic chemistry. Clearly, it turned into one of the most influential meetings over the years, allowing close interactions among biologists, biochemists, inorganic chemists, and physical chemists, to provide an excellent platform for discussing the latest developments (Figure 1).

Metalloenzymes catalyze the most challenging and consequential chemical reactions on Earth, including the ATP driven $N_2 \rightarrow NH_3$ reduction by nitrogenase [35], the transformation of CO to CO_2 by carbon monoxide dehydrogenase as well as the splitting/formation of H_2 by hydrogenase [36–38], and the light-driven splitting of H_2O by the Mn_4CaO_5 cluster of the O_2-evolving complex in photosystem II [39]. The complex centers of redox-active transition metal ions and exogenous (in)organic ligands, that support these reactions, must be assembled and inserted into proteins by a sophisticated biosynthetic machinery, as elegantly shown for the active site cluster of nitrogenase or the molybdopterin cofactor [40–46], and discussed in chapter 7 for iron-sulfur proteins. Indeed, some of the most notable transformations in nature occur at transition metal centers coordinated to sulfur ligands.

Although the field has its roots in findings many centuries old, the current stream of research stems from studies of proteins and enzymes hosting transition metal ions at their active sites. The metal cores of these biomolecules offer a unique opportunity to probe their structure and function at functional groups that can readily be distinguished from the surrounding organic environment by spectroscopic techniques. Consequently, biophysical chemists were at the front of bioinorganic chemistry, taking advantage of the power of their methodologies on a host of interesting problems [1, 11, 18]. These include UV/Vis, continuous wave and pulsed electron paramagnetic resonance (EPR) techniques, Mössbauer (for iron) spectroscopy, X-ray absorption (XAS) including near-edge spectroscopy (XANES) and extended X-ray absorption fine structure (EXAFS), magnetic circular dichroism (MCD), and resonance Raman (rR) spectroscopy. Combined with methods for studying enzyme mechanisms, rapid-freeze (EPR) and stopped-flow (UV/Vis) kinetics experiments, the ability to get macromolecular X-ray structures at synchrotrons in record times, and the power of site-directed mutagenesis, cloning, polymerase chain reaction, and similar advances that fueled the molecular biology revolution. In summary, these physical and biochemical

Figure 1. Participants of the Gordon Research Conference *Metals in Biology*, held in February 1980 in Santa Barbara, USA. Conference chairs Jim Fee and Harry Gray; the photograph shows several pioneering researchers of bioinorganic chemistry mentioned in this volume of MILS on transition metal sulfur sites: William Blumberg, Robert Bray, Barbara Burgess, Stephen Cramer, John Dawson, David Dooley, John Enemark, Hans Freeman, John Groves, Brian Hoffmann, Andrew Johnson, Steve Lippard, David Lowe, Bo Malmström, Jack Peisach, Israel Pecht, Larry Que, Robert Scott, Barry Smith, Thomas Spiro, Edward Solomon, Edward Stiefel, and Joan Valentine. Courtesy Michelle Knapp, Gordon Research Conferences incorporated.

techniques brought deeper insights into a host of problems in bioinorganic chemistry by the turn of the century [1].

The copper field represents a beautiful example in our opinion, where the advent of EPR spectroscopy clearly broke the ice. With this technique much more sense could be made of the metal-to-protein stoichiometry and the UV/Vis absorption spectra that had been available so far. Suddenly, those designations, such as CuA and CuB in cytochrome *c* oxidase (COX) or type 1, 2, or 3 copper in multi-copper oxidases assumed distinct character [11]. The early work of Keilin and Mann [47], and Tissières [48], conclusively demonstrated that laccase (later classified as a blue oxidase [49]), the oxidizing enzyme present in the latex of the Indo-Chinese lacquer tree, is a copper-dependent enzyme. Furthermore, it was assumed that laccase contained Cu(II) in the resting state, and that catalytic activity was associated with a change in valency of the metal to Cu(I). This assumption was confirmed in 1959 by Malmström and coworkers applying the EPR technique [50]. Two decades later, the three-dimensional structure of the 'blue' or 'type 1' copper protein plastocyanin isolated and purified from the poplar tree was determined at a resolution of 2.7 Å by the group of Hans Freeman in Sidney, Australia [51]. The Cu atom is embedded in a highly distorted tetrahedral coordination geometry and is coordinated by a cysteine thiol group, a methionine thioether group, and two histidine imidazole groups forming a $CuS_{Cys}S_{Met}2N_{His}$ center with unique electronic structural properties [52]. Similarly, the purple CuA center present in nitrous oxide reductase (N_2OR) and predicted to be also present in COX, was characterized by Antholine, Zumft, and Kroneck (as reviewed by Kroneck most recently [54]) approximately three decades ago, using multi-frequency EPR spectroscopy in the range 2.5–35 GHz. The characteristic equidistant 7-line pattern of the EPR spectrum, due to Cu hyperfine interaction with an intensity ratio of 1:2:3:4:3:2:1 for individual lines, indicated the presence of a novel Cu site (Figure 2). Its binuclear, mixed-valent structure, $[Cu^{1.5+}\text{-}(Cys)_2\text{-}Cu^{1.5+}]$, with $S = 1/2$ and two bridging cysteine thiolate residues, was later confirmed by X-ray crystallography [53, 54].

Both redox enzymes N_2OR and COX are constituents of important biological processes and employ CuA as electron transfer center. N_2OR is the terminal reductase in a respiratory chain converting N_2O to N_2 in denitrifying bacteria, and COX is the terminal oxidase of the aerobic respiratory chain of certain bacteria and eukaryotic organisms. They are involved in complex multi-electron and multi-proton transfer reactions of kinetically inert gaseous molecules. Notably, COX also pumps protons across the membrane ($H_i^+ \rightarrow H_o^+$), with up to 4 protons per O_2 molecule (Equations 1, 2) [54].

$$N_2O + 2H^+ + 2e^- \rightarrow N_2 + H_2O \quad (1)$$

$$4\ \text{Fe-cyt}^{2+} + 8H_i^+ + O_2 \rightarrow 4\text{Fe-cyt}^{3+} + 4H_o^+ + 2\ H_2O \quad (2)$$

From such detailed studies we learned that Nature has evolved highly sophisticated ways of controlling the relatively flexible stereochemistry of transition metal ions. Early on, the structure and reactivity of active sites in metalloproteins

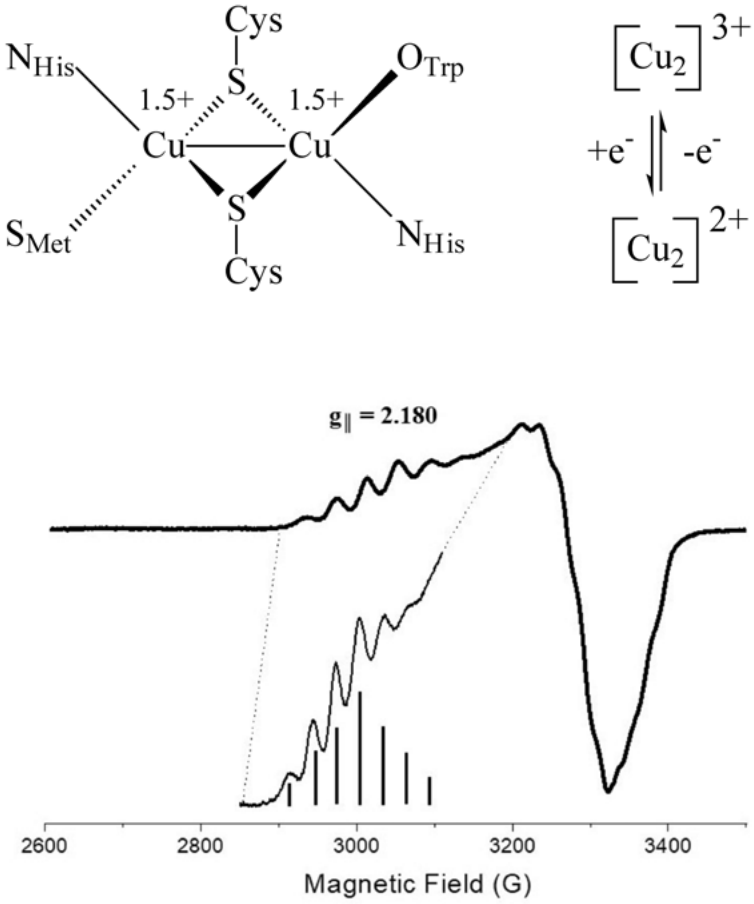


Figure 2. Structural model of the purple mixed-valent, electron transfer center CuA in nitrous oxide reductase from *Pseudomonas stutzeri*, $[Cu_A^{1.5+}\text{-}(Cys)_2\text{-}Cu_A^{1.5+}]$ (S = 1/2), and its characteristic 7-line EPR spectrum, microwave frequency 9.31 GHz, temperature of 10 K; the hyperfine splitting observed in the $g_{\parallel}$ region results from the interaction of one unpaired electron with two Cu nuclei, $^{63,65}Cu$ nuclear spin I = 3/2 [54].

turned out to be different from anything previously encountered in simple metal complexes, which inspired numerous inorganic chemists to expand their vision of metal ion reactivity [2, 7].

All the metal centers (both mono-nuclear and multi-nuclear) discussed in this volume harbor sulfur ligands in their coordination sphere, e.g., sulfide (S^{2-}), cysteine (RS^-), methionine ($RSCH_3$), the molybdopterin cofactor, or a combination of these sulfur ligands (see Chapter 2). We will focus on the most exciting, in our view, transition metal-sulfur sites discovered in living organisms, their complex three-dimensional structures and catalytic activities. These include the type 1 Cu mononuclear center in blue copper proteins (see Chapter 3), the purple mixed-

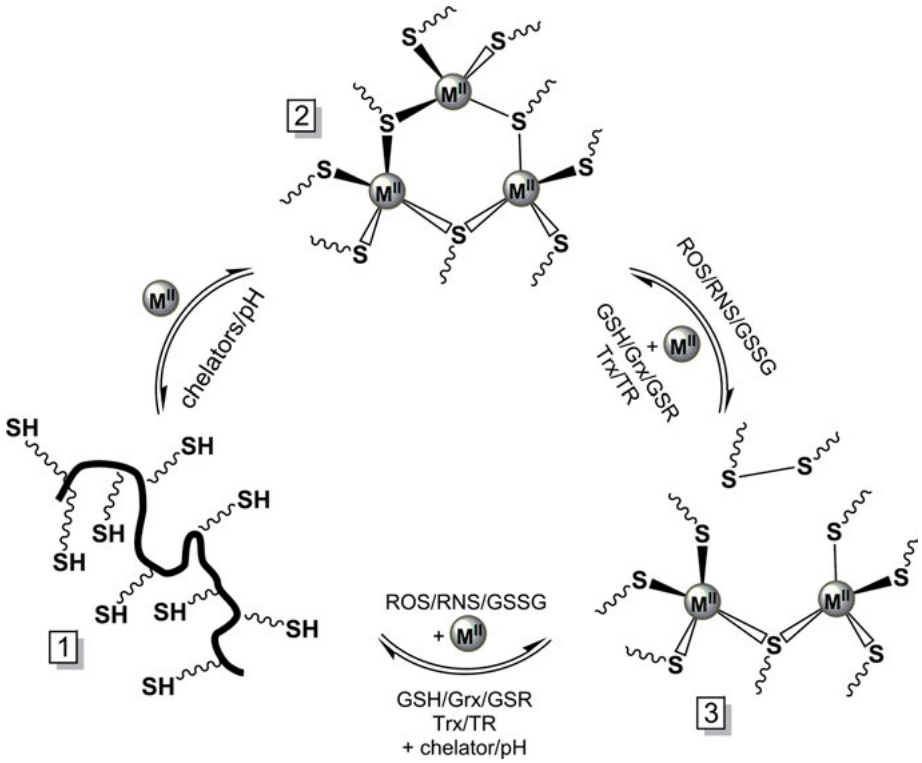


Figure 3. Distribution of metallothionein (MT) species *in vivo* using the vertebrate ß-type metal-thiolate cluster as model. The reduced, metal-free and largely unstructured apo-MT (1) forms a characteristic metal-thiolate cluster structure (2) in the presence of d^{10} metal ions, here Zn(II) or Cd(II). Metal release back to (1) is achieved by a drop in pH or by suitable metal ion chelators. Due to the high thermodynamic stability of the clusters, the latter process is mainly kinetically driven *in vivo*. Conditions of oxidative stress (reactive oxygen/nitrogen species, ROS/RNS, or higher levels of glutathione disulfide, GSSG) can generate disulfide bridges accompanied by (partial) metal ion release (3). It is generally considered that partially oxidized sub-metalated species retain part of their three-dimensional fold and hence allow reversibility of the process in presence of reductants (glutathione, GSH; glutaredoxin, Grx; glutathione-disulfide reductase, GSR; thioredoxin, Trx; thioredoxin reductase, TR; or alike) and metal ions. Also the more complex transition from (1) to (3) or *vice versa* is feasible.

valent [$Cu^{1.5+}$-$(Cys)_2$- $Cu^{1.5+}$] CuA present in COX and N_2OR (see Chapter 4), the unique tetranuclear copper-sulfide catalytic center of N_2OR (see Chapter 5), the heme-thiolate site in cytochrome P_{450} with the strong oxidants compound I and II as key intermediates (see Chapter 6), the iron-sulfur proteins with bound inorganic (S^{2-}) and organic (Cys^{-}) sulfur (see Chapter 7), the pterin dithiolene cofactor (Moco) coordinated to either molybdenum or tungsten (see Chapter 8), the [8Fe-7S] P-cluster and the catalytic FeMo cofactor of nitrogenase, which hosts the unique [Mo-7Fe-9S-C]-homocitrate moiety (see Chapter 9), the siroheme-

[4Fe-4S] coupled center involved in the six-electron reduction of sulfite (SO_3^{2-}) to hydrogen sulfide (H_2S) (see Chapter 10), or the catalytic NiFeS sites of hydrogenases and CO dehydrogenase (see Chapter 11). Finally, Zn-finger domains will be introduced in Chapter 12. While their amino acid composition can vary between the individual domains, a shared feature is the coordination of a structural Zn(II) ion by a different combination of cysteine and histidine ligands [55–57].

Because of the vast literature accumulated most recently on metallothionein (MT), this important class of sulfur-rich metal-binding proteins will not be covered in a special chapter. Instead, we refer to very informative and authoritative reviews by experts in this dynamic research area [58–63]. MTs are a ubiquitous class of proteins with an extremely high cysteine content, which can reach up to 30 % of the total amino acid content in the vertebrate forms (Figure 3). The sheer number of potential metal-coordinating ligands in the form of soft thiolate groups strongly suggest a pronounced metal binding ability, and indeed, thioneins, the metal-free forms of the proteins, soak-up certain metal ions like a sponge. Due to their thiophilicity, metal ions with the d^{10} electronic configuration, such as Zn(II), Cu(I), Cd(II), and Hg(II) are the preferred binding partners, as stated by Eva Freisinger, guest-editor of a special issue of the *International Journal of Molecular Sciences* published in 2017 (https://www.mdpi.com/journal/ijms/special_issues/metallothioneins).

2. FROM ANOXIC TO OXIC ENVIRONMENTS – SULFUR AND THE BIOAVAILABILITY OF TRANSITION METALS

The availability of transition metal ions found in many biological catalysts, without which there would be no life, has changed over geological time and varies between habitats but is held within vital limits in cells [64]. Oxygen is hailed as the *Elixir of Life*, a wonder molecule with many advantages as well as many disadvantages. However, everybody agrees, oxygen (O_2) is important, stop breathing and you will be dead in a few minutes. In the beginning of the Earth, approximately four billions years ago, there was not much oxygen around. The air contained probably 1 ppm of O_2 compared to just less than 21 % (208500 ppm) today, in principle a pollution without parallel in the history of life on Earth. For the first single-celled organisms that lived on the early Earth, O_2 was anything but live-giving, it was a poison that could kill, even at trace levels [65, 66]. Anaerobes, "oxygen-hating microorganisms", can use inorganic nitrogen and sulfur compounds such as the oxyanions nitrate (NO_3^-) or sulfate (SO_4^{2-}), to conserve energy by converting them to dinitrogen (N_2) or ammonia (NH_3), and to elemental sulfur (S^0) or hydrogen sulfide (H_2S) (Equations 3–6) [67–70].

$$2\ NO_3^- + 10\ e^- + 12\ H^+ \rightarrow N_2 + 6\ H_2O \quad (3)$$

$$NO_3^- + 8e^- + 9\ H^+ \rightarrow NH_3 + 3\ H_2O \quad (4)$$

$$SO_4^{2-} + 6\ e^- + 8\ H^+ \rightarrow S^0 + 4\ H_2O \quad (5)$$

$$SO_4^{2-} + 8\ e^- + 10\ H^+ \rightarrow H_2S + 4\ H_2O \quad (6)$$

In contrast, aerobes, humans, animals and plants, depend on O_2 as terminal electron acceptor to conserve energy (respiration, ATP synthesis) [71, 72]. The initial increase of O_2 in the atmosphere, the great oxidation event (GOE), and its relationship with life remains a matter of debate and intensive research [73]. The generally accepted theory is that the emergence of organisms capable of performing oxygenic photosynthesis – probably ancestors of today's cyanobacteria – contributed to this. The Earth's early atmosphere was formed by outgassing of gases, such as H_2O, CO_2, SO_2, H_2S, and HCl, trapped in the interior of the Earth. This still can be observed today during volcanic eruptions. Probably, CO, N_2, NO, H_2, HCN, acetylene (C_2H_2), and traces of O_2 were also present by the time of the origin of life [74–80]. The biogeochemical cycling of the basic elements for life (H, C, N, O, P, S) and the biogeochemistry of the essential transition metals caught the interest of researchers in view of their importance for the Earth and their impact on our environment and climate [81].

Transition metal ions function as cofactors, or as part of cofactors in enzymes, and as structural elements in proteins. Essential life processes, e.g., photosynthesis, respiration, or nitrogen fixation, strictly depend on transition metal ions and their ability to catalyze multi-electron, multi-proton transformations. Other important life processes, such as proteolysis and the equilibration of carbon dioxide (CO_2) and bicarbonate (HCO_3^-) are hydrolytic reactions that also require catalysis by metalloenzymes. The essential transition metal ions for terrestrial organisms include vanadium to zinc of the first-row transition metal series, and molybdenum and tungsten in the second- and third-row series, many of these metals are coordinated to sulfur ligands [2–4, 82–84]. There exists a long-term link between atmospheric O_2 and the geochemical cycles of carbon and sulfur, as pointed out by Ebelmen as early as 1845 (Equation 7) [85].

$$2\ Fe_2O_3 + 8\ SO_4^{2-} + 16\ H^+ \quad \rightleftharpoons \quad 15\ O_2 + 4\ FeS_2 + 8\ H_2O \tag{7}$$

In reaction (7) (right to left), FeS_2 (pyrite) is oxidized during weathering of the continents. The process from left to right represents several reactions: (i) photosynthesis and initial burial of organic matter, (ii) early bacterial reduction of sulfate (SO_4^{2-}) to hydrogen sulfide (H_2S), with organic matter serving as the reductant, and (iii) the precipitation of FeS_2 via the reaction of H_2S with Fe_2O_3 [70]. To reconstruct the redox history of the planet, geochemists like to choose sulfur, certainly one of the most versatile elements due to its reactivity in different reduction and oxidation states. The inorganic sulfur compounds of biological relevance which occur in the biological sulfur cycle are elemental sulfur (S^0), sulfate (SO_4^{2-}), sulfite (SO_3^{2-}), thiosulfate ($S_2O_3^{2-}$), sulfide (S^{2-}), polythionates (e.g., tetrathionate, $S_4O_6^{2-}$), and polysulfides (S_n^{2-}) (see Chapter 2). Many biogeochemical processes fractionate the multiple stable isotopes of sulfur in telltale ways, helping researchers to understand the redox evolution of the oceans and the atmosphere [86–88]. For example, to illuminate the history of the emergence of O_2-producing (oxygenic) photosynthesis, the behavior of the ancient manganese cycle was examined using scientific drill cores through an early Paleoproterozoic succession (2.415 Ga) preserved in South Africa. It was found that the original Mn-oxide

phases were not produced by reactions with O_2, pointing to a different high-potential oxidant. These results suggest that the oxidative branch of the Mn cycle predates the rise of O_2 in the atmosphere, and support the hypothesis that the water-oxidizing complex of photosystem II evolved from a former transitional photosystem capable of single-electron oxidation reactions of Mn [89].

It has been postulated that O_2-evolving photosynthesis altered the solubility of metals on a global scale, and it has been hypothesized that the selection of metal centers in metalloenzymes has been influenced by the availability of metals through geological time, in particular as a result of large differences in the solubility of metal sulfides [20, 90–92]. General restrictions on the availability of elements as free metal cations (M^+, M^{2+}) in the oceans are insolubility and complex ion formation with anions, such as sulfide, sulfate, silicate, hydroxide/oxide, or carbonate. The general order of insolubility of salts in the series of divalent transition metal ions (M^{2+}), is $Cu^{2+} > Ni^{2+} > Zn^{2+} > Co^{2+} > Fe^{2+} > Mn^{2+}$ which holds for sulfides and oxides [90–92]. Probably, the early ocean had high Mg^{2+} and Fe^{2+} concentrations, and the high Fe^{2+} concentration reduced somewhat the amount of sulfide due to precipitation as pyrite, FeS_2, which could enter into weathering reactions (Equation 7). Furthermore, the residual sulfide greatly restricted the free ion concentrations of Cu and Zn, probably due to precipitation as insoluble sulfides. Other M^{2+} concentrations in the environment can be estimated from their solubility products. Of particular interest is the small difference in the solubility products of Ni(II) sulfide and Ni(II) hydroxide, making nickel a particular important element for early life organisms [92]. The environment was electron-rich to some extent, and gases such as H_2 and CO_2, and transition metal sulfides were available. Special catalysts were needed to handle these gases, featuring a rich biochemical role for nickel [93]. In some sense it looks as if nickel, similar to cobalt, was most useful when metabolism was based on such chemicals, but after the advent of O_2 their value diminished [91].

Molecular oxygen did not immediately accumulate in the atmosphere and oceans, due largely to the presence of abundant Fe(II) compounds, which reacted with O_2 to precipitate massive banded Fe(III) formations (BIFs). As larger concentrations of O_2 began to build up in the atmosphere, redox reactions of metal sulfides with O_2 could occur, such as the transformation of insoluble MoS_2 (Mo(IV), molybdenum disulfide) to soluble MoO_4^{2-}(Mo(VI), molybdate) (Equation 8) [71, 82].

$$2\ MoS_2 + 7\ O_2 + 2\ H_2O \quad \rightleftharpoons \quad 2\ MoO_4^{2-} + 4\ SO_2 + 4\ H^+ \tag{8}$$

This reaction may have allowed life to proliferate by making the highly soluble molybdate ion available for incorporation into critical molybdenum enzymes, such as nitrate reductase ($NO_3^- \rightarrow NO_2^-$) and sulfite oxidase ($SO_3^{2-} \rightarrow SO_4^{2-}$). MoO_4^{2-} itself is catalytically inactive in biological systems and has to be complexed by a special scaffold, either the ubiquitous pterin dithiolene cofactor (Moco) (see Chapter 8), or the catalytic cofactor of bacterial nitrogenase, which hosts the unique [Mo-7Fe-9S-C]-homocitrate moiety (see Chapter 9).

3. OUTLOOK AND FUTURE DIRECTIONS

Recently, deep-time evolutionary connections within the oxidoreductase class of enzymes were analyzed with regard to protein structure similarities. In metalloenzymes, structural similarity between proteins may be a result of convergence on a limited repertoire of metal-binding protein topologies. Most likely, a small contingent of modular structures was incorporated repeatedly in oxidoreductases across the tree of life, with ferredoxin as the first likely module, giving rise to a number of specialized iron-sulfur proteins. Independently, other important redox proteins evolved, such as the cytochrome *c* featuring as axial ligands histidine and methionine to the heme iron center, the blue type 1 Cu mononuclear site, and the mixed-valent CuA binuclear center featuring Cu ligands histidine, cysteine, and methionine, which give access to an increased variety of metabolic substrates and redox potentials [23]. In this context, the concept of a last universal common ancestor of all cells (LUCA) is central to the study of early evolution and life's origin. To get further insight, all clusters and phylogenetic trees for 6.1 million protein coding genes from sequenced prokaryotic genomes were examined to reconstruct the microbial ecology of LUCA which led to the identification of protein families that trace to LUCA by phylogenetic criteria. These proteins are not universally distributed, thus they can shed light on LUCA's physiology. LUCA's biochemistry is replete with iron-sulfur clusters and radical reaction mechanisms, its cofactors reveal dependence upon transition metals, flavins, S-adenosyl methionine, coenzyme A, ferredoxin, molybdopterin, corrins, as well as selenium, and its genetic code required nucleoside modifications and S-adenosyl methionine-dependent methylations. LUCA inhabited a geochemically active environment rich in H_2, CO_2, and iron. The data support the theory of an autotrophic origin of life involving the Wood–Ljungdahl pathway (also called the reductive acetyl-coenzyme A pathway, with H_2 as electron donor, and CO_2 as electron acceptor) in a hydrothermal setting [94].

In summary, the bioinorganic chemistry of transition metals and sulfur is a vibrant field of research as impressively documented by discoveries of novel metal sites with remarkable chemical activities [95]. For example, Jeoung and Dobbek reported the double-cubane $[Fe_8S_9]$-cluster $[\{Fe_4S_4(SCys)_3\}2(\mu_2\text{-}S)]$. The enzyme containing this cluster can reduce small molecules including acetylene (C_2H_2), azide (N_3^-), and hydrazine (N_2H_4), it belongs to a class of metalloenzymes akin in fold, metal clusters, and reactivity to nitrogenase [96]. In 2015, Hausinger and associates opened a new chapter of nickel biochemistry when they described for the first time a tethered nicotinic acid-derived Ni complex in lactate racemase from *Lactobacillus plantarum* (Figure 4) [97].

This novel organometallic cofactor, a Ni(II) pyridinium-3,5-bisthiocarboxylic acid mononucleotide, carries a Ni–C bond and is named nickel-pincer nucleotide (NPN) [98–102]. Rosenzweig and coworkers isolated the protein PmoD from methane-oxidizing microorganisms which transform the greenhouse gas methane (CH_4) to methanol (CH_3OH) using the copper-dependent enzyme particulate methane monooxygenase. PmoD is critical for Cu-dependent growth on methane and forms a Cu center that exhibits the spectroscopic (UV/Vis, EPR) features

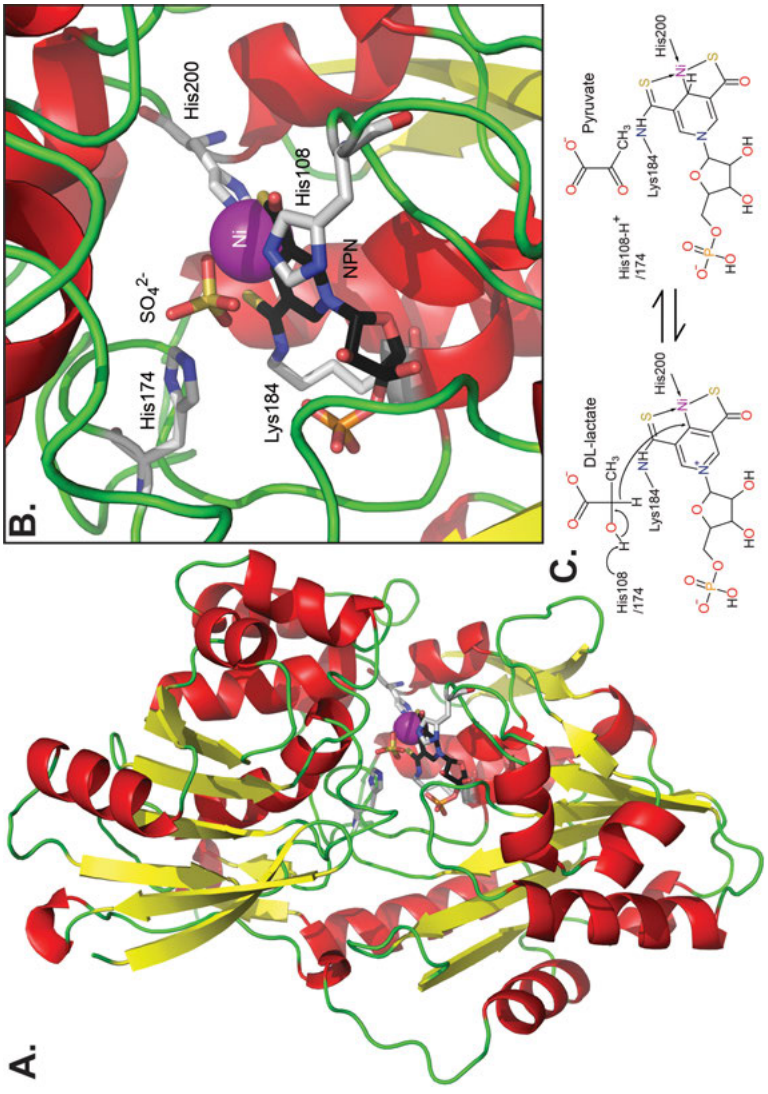


Figure 4. Crystal structure and proposed reaction mechanisms of lactate racemase from *Lactobacillus plantarum* which harbors the novel organometallic cofactor, nickel-pincer nucleotide (NPN). The organic portion of this cofactor, pyridinium-3-thioamide-5-thiocarboxylic acid mononucleotide, is covalently tethered to Lys184 of the enzyme and serves as a tridentate ligand to nickel [95]. (**A**) Overall three-dimensional structure (PDB ID 5HUQ). (**B**) Catalytic site with NPN, (**C**) Proposed hydride transfer catalytic mechanism. Taken from B. Desguin, *J. Bacteriol. Parasitol.* **2018**, *9*, 335. doi:10.4172/2155–9597.1000335 [102].

of a well-defined mixed-valent CuA site found in COX and N_2OR, however, it uses a previously unobserved ligand set derived from a cupredoxin homodimer (see Chapter 4) [103].

Last but not least, another exciting development which gained a lot of interest in recent years, relates to the emerging field of rare earth element biochemistry and the role of lanthanides as biologically essential metals. A statement which was unthinkable until recently, says bioinorganic chemist Lena Daumann. She and microbiologist Huub Op den Camp lead a team of researchers who want to understand why nature has chosen lanthanides to have a catalytic role in alcohol dehydrogenase enzymes as they are found in methanotrophic and methylotrophic bacteria. These enzymes oxidize the alcohol to the corresponding aldehyde in a 2-electron, 2-proton transfer reaction at the catalytic PQQ (2,7,9-tricarboxypyrroloquinoline quinone) cofactor which, depending on the organism, coordinates either a divalent Ca^{2+} ion or a trivalent lanthanide ion (Ln^{3+}, Ce^{3+}, Eu^{3+}) [104–107].

ACKNOWLEDGMENTS

The authors are grateful for continuous financial support by CONACYT and DGAPA-UNAM (MEST), and Deutsche Forschungsgemeinschaft and Universität Konstanz (PK).

ABBREVIATIONS AND DEFINITIONS

ATP	adenosine 5′-triphosphate
BIF	banded ferric iron (Fe^{3+}) formation
COX	cytochrome *c* oxidase
EPR	electron paramagnetic resonance
EXAFS	extended X-ray absorption fine structure
Fe-cyt$^{2+/3+}$	cytochrome reduced/oxidized
Ga	billion years (gigayears = 10^9 years)
GOE	great oxidation event
LUCA	last universal common ancestor
MCD	magnetic circular dichroism
NPN	nickel-pincer nucleotide
N_2OR	nitrous oxide reductase
PQQ	2,7,9-tricarboxypyrroloquinoline quinone
rR	resonance Raman
XANES	X-ray absorption near-edge spectroscopy
XAS	X-ray absorption spectroscopy

REFERENCES

1. S. J. Lippard, *Nature Chem. Biol.* **2006**, *2*, 504–507.
2. *Copper Proteins*, Ed. T. G. Spiro, John Wiley & Sons, Chichester, UK, 1981.
3. *Iron-Sulfur Proteins*, Ed. T. G. Spiro, John Wiley & Sons, Chichester, UK, 1982.
4. W. Kaim, B. Schwederski, A. Klein, *Bioinorganic Chemistry: Inorganic Elements in the Chemistry of Life. An Introduction and Guide*, 2nd edn., John Wiley & Sons, Chichester, UK, 2013.
5. R. J. P. Williams, *Nature* **1990**, *343*, 213–214.
6. R. J. P. Williams, *Coord. Chem. Rev.* **1990**, *100*, 573–610.
7. R. H. Holm, P. Kennepohl, E. I. Solomon, *Chem. Rev.* **1996**, *96*, 2239–2314.
8. A. Butler, *Science* **1998**, *281*, 207–209.
9. J. S. Valentine, T. V. O'Halloran, *Curr. Op. Chem. Biol.* **1999**, *3*, 129–130.
10. *The Biochemistry of Copper*, Eds J. Peisach, P. Aisen, W. E. Blumberg, Academic Press, New York and London, 1966.
11. H. Beinert, *J. Biol. Chem.* **2002**, *277*, 37967–37972.
12. H. Beinert, *Coord. Chem. Rev.* **1977**, *23*, 119–129.
13. H. Beinert, *J. Inorg. Biochem.* **1996**, *64*, 79–100.
14. F. M. M. Morel, N. M. Price, *Science* **2003**, *300*, 944–947.
15. H. B. Gray, *Proc. Natl. Acad. Sci. USA* **2003**, *100*, 3563–3568.
16. R. J. P. Williams, *Chem. Commun.* **2003**, 1109–1113.
17. J. N. Rauch, J. M. Pacyna, *Global Biogeochem. Cycles* **2009**, *23*, GB2001, doi:10.1029/2008GB003376.
18. S. J. Lippard, *J. Am. Chem. Soc.* **2010**, *132*, 14689–14693.
19. Y. Lu, *Proc. Natl. Acad. Sci. USA* **2010**, *107*, 1811–1812.
20. R. J. P. Williams, R. E. M. Rickaby, *Evolution's Destiny: Co-evolving Chemistry of the Environment and Life*, Royal Chemical Society Publishing, 2012.
21. K. L. Bren, R. Eisenberg, H. B. Gray, *Proc. Natl. Acad. Sci. USA* **2015**, *112*, 13123–13127.
22. L. J. Robbins, S. V. Lalonde, N. J. Planavsky, C. A. Partin, C. T. Reinhard, B. Kendall, C. Scott, D. S. Hardisty, B. C. Gill, D. S. Alessi, C. L. Dupont, M. A. Saito, S. A. Crowe, S. W. Poulton, A. Bekker, T. W. Lyons, K. O. Konhauser, *Earth-Science Rev.* **2016**, *163*, 323–348.
23. H. Raanan, D. H. Pike, E. K. Moore, P. G. Falkowski, V. Nanda, *Proc. Natl. Acad. Sci. USA* **2018**, *115*, 1280–1285.
24. M. A. Zoroddu, J. Aaseth, G. Crisponi, S. Medici, M. Peana, V. M. Nurchi, *J. Inorg. Biochem.* **2019**, *195*, 120–129.
25. J. Liu, S. Chakraborty, P. Hosseinzadeh, Y. Yu, S. Tian, I. Petrik, A. Bhagi-Damodaran, Y. Lu, *Chem. Rev.* **2014**, *114*, 4366–4469.
26. Y. Lu, *Metalloprotein Design & Engineering*, in *Encyclopedia of Inorganic Chemistry*, 2nd edn, Vol. V, Ed. R. B. King, John Wiley & Sons, Chichester, 2005, pp. 3159–3192.
27. J. Rittle, M. J. Field, M. T. Green, F. A. Tezcan, *Nature Chem.* **2019**, *11*, 434–441.
28. A. Deliz Liang, J. Serrano-Plana, R. L. Peterson, T. R. Ward, *Acc. Chem. Res.* **2019**, *52*, 585–595.
29. I. Schlichting, J. Berendzen, K. Chu, A. M. Stock, S. A. Maves, D. E. Benson, R. M. Sweet, D. Ringe, G. A. Petsko, S. G. Sligar, *Science* **2000**, *287*, 1615–1622.
30. M. Sommerhalter, R. L. Lieberman, A. C. Rosenzweig, *Inorg. Chem.* **2005**, *44*, 4770–4778.
31. M. Akram, A. Dietl, U. Mersdorf, S. Prinz, W. Maalcke, J. Keltjens, C. Ferousi, N. M. de Almeida, J. Reimann, B. Kartal, M. S. M. Jetten, K. Parey, T. R. M. Barends, *Sci. Adv.* **2019**, *5*: eaav4310; DOI: 10.1126/sciadv.aav4310.

32. S. Sinnecker, F. Neese, *Theoretical Bioinorganic Spectroscopy. Atomistic Approaches in Modern Biology*, *Topics in Current Chemistry*, **2006**, *268*, 47–83.
33. M. R. A. Blomberg, T. Borowski, F. H. Rong-Zhen Liao, P. E. M. Siegbahn, *Chem. Rev.* **2014**, *114*, 3601–3658.
34. L. Banci, C. Luchinat, *J. Biol. Inorg. Chem.* **2014**, *19*, 487–489.
35. C. C. Lee, M. W. Ribbe, Y. Hu, *Met. Ions Life Sci.* **2014**, *14*, 147–176.
36. J.-H. Jeoung, J. Fesseler, S. Goetzl, H. Dobbek, *Met. Ions Life Sci.* **2014**, *14*, 34–69.
37. V. C.-C. Wang, S. W. Ragsdale, F. A. Armstrong, *Met. Ions Life Sci.* **2014**, *14*, 71–97.
38. A. Parkin, *Met. Ions Life Sci.* **2014**, *14*, 99–124.
39. J. Yano, J. Kern, V. K. Yachandra, H. Nilsson, S. Koroidov, J. Messinger, *Met. Ions Life Sci.* **2015**, *15*, 13–43.
40. P. C. Dos Santos, D. R. Dean, Y. Hu, M. W. Ribbe, *Chem. Rev.* **2004**, *104*, 1159–1174.
41. G. Schwarz, *Cell. Mol. Life Sci.* **2005**, *62*, 2792–2810.
42. G. Schwarz, R. R. Mendel, M. W. Ribbe, *Nature* **2009**, *460*, 839–847.
43. C. Iobbi-Nivol, S. Leimkühler, *Biochim. Biophys. Acta* **2013**, *827*, 1086–1101.
44. R. R. Mendel, *J. Biol. Chem.* **2013**, *288*, 13165–13172.
45. R. R. Mendel, S. Leimkühler, *J. Biol. Inorg. Chem*. **2015**, *20*, 337–347.
46. J. Krausze, T. W. Hercher, D. Zwerschke, M. L. Kirk, W. Blankenfeldt, R. Mendel, T. Kruse, *Biochem. J.* **2018**, *475*, 1739–1753.
47. D. Keilin, T. Mann, *Nature* **1939**, *143*, 23–24.
48. A. Tissières, *Nature* **1948**, *162*, 340–341.
49. R. Malkin, B. G. Malmström, *Adv. Enzymol.* **1970**, *33*, 177–243.
50. B. G. Malmström, R. Mosbach, T. Vänngård, *Nature* **1959**, *183*, 321–322.
51. P. M. Colman, H. C. Freeman, J. M. Guss, M. Murata, V. A. Norris, J. A. M. Ramshaw, M. P. Venkatappa, *Nature* **1978**, *272*, 319–324.
52. H. B. Gray, E. I. Solomon, *Electronic Structures of Blue Copper Centers in Protein,* in *Copper Proteins*, Ed. T. G. Spiro, John Wiley & Sons, Chichester, UK, 1981, pp. 1–39.
53. W. G. Zumft, P. M. H. Kroneck, *Adv. Microb. Physiol.* **2006**, *52*, 107–227.
54. P. M. H. Kroneck, *J. Biol. Inorg. Chem.* **2018**, *23*, 27–39.
55. J. M. Berg, *J. Biol. Chem.* **1990**, *265*, 6513–6516.
56. K. Kluska, J. Adamczyk, A. Krężel, *Coord. Chem. Rev.* **2018**, *367*, 18–64.
57. C. Abbehausen, *Metallomics* **2019**, *11*, 15–28.
58. *Metallothioneins and Related Chelators*, Vol. 5 of *Met. Ions Life Sci*., Eds A. Sigel, H. Sigel, R. K. O. Sigel, Royal Society of Chemistry, Cambridge, 2009.
59. E. Freisinger, *Chimia* **2010**, *64*, 217–224.
60. M. Vašák, *J. Biol. Inorg. Chem.* **2011**, *16*, 975–976.
61. A. Krężel, W. Maret, *Int. J. Mol. Sci.* **2017**, *18*, 1237; doi:10.3390/ijms18061237.
62. A. Ziller, L. Fraissinet-Tachet, *Metallomics* **2018**, *10*, 1549–1559.
63. J. S. Calvo, V. M. Lopez, G. Meloni, *Metallomics* **2018**, *10*, 1777–1791.
64. K. J. Waldron, J. C. Rutherford, D. Ford, N. J. Robinson, *Nature* **2009**, *460*, 823–830.
65. N. Lane, *Oxygen, the Molecule that Made the World*, Oxford University Press, Oxford, UK, 2002.
66. D. E. Canfield, *Oxygen: A Four Billion Year History*, Princeton University Press, Princeton, USA, 2014.
67. R. R. Eady, S. V. Antonyuk, S. S. Hasnain, *Curr. Op. Chem. Biol*. **2016**, *31*, 103–112.
68. R. M. Martínez-Espinosa, J. A. Cole, D. J. Richardson, N. J. Watmough, *Biochem. Soc. Trans.* **2011**, *39*, 175–178; doi:10.1042/BST0390175.
69. K. Parey, G. Fritz, U. Ermler, P. M. H. Kroneck, *Metallomics*, **2013**, *5*, 302–317.
70. L. L. Barton, M.-L. Fardeau, G. D. Fauque, *Met. Ions Life Sci.* **2014**, *14*, 237–277.
71. P. M. H. Kroneck, *Met. Ions Biol. Syst.* **2005**, *43*, 1–10.

72. M. E. Sosa Torres, J. P. Saucedo-Vázquez, P. M. H. Kroneck, *Met. Ions Life Sci.* **2015**, *15*, 1–12.
73. A. M. Satkoski, N. J. Beukes, W. Li, B. L. Beard, C. M. Johnson, *Earth and Planetary Science Letters* **2015**, *430*, 43–53.
74. D. E. Canfield, *Annu. Rev. Earth Planet. Sci.* **2005**, *33*, 1–36.
75. J. F. Kasting, M. T. Howard, *Phil. Trans. R. Soc. B* **2006**, *361*, 1733–1742.
76. H. D. Holland, *Geochim. Cosmochim. Acta* **2002**, *66*, 3811–3826.
77. A. D. Anbar, Y. Duan, T. W. Lyons, G. L. Arnold, B. Kendall, R. A. Creaser, A. J. Kaufman, G. W. Gordon, C. Scott, J. Garvin, R. Buick, *Science* **2007**, *317*, 1903–1906.
78. A. D. Holland, *Geochim. Cosmochim. Acta* **2009**, *73*, 5241–5255.
79. J. F. Kasting, *Chemical Geology* **2013**, *362*, 13–25.
80. T. Geisberger, P. Diederich, T. Steiner, W. Eisenreich, P. Schmitt-Kopplin, C. Huber, *Life* **2019**, *9*, 50; doi:10.3390/life9020050.
81. V. Smil, *Cycles of Life: Civilization and the Biosphere*, Scientific American Library, W. H. Freeman and Company, New York, USA, 1997.
82. E. Stiefel, *Met. Ions Biol. Syst.* **2002**, *39*, 1–29.
83. H. Beinert, R. H. Holm, E. Münck, *Science* **1997**, *277*, 653–659.
84. D. C. Rees, J. B. Howard, *Science* **2003**, *300*, 929–930.
85. J. J. Ebelmen, *Ann. Rev. Mines.* **1845**, *12*, 627–654.
86. L. R. Kump, *Perspective* **2012**, 410–411; DOI: 10.2113/gselements. 8.6.410F.
87. M. Sub Sim, D. T. Wang, G. M. Zane, J. D. Wall, T. Bosak, S. Ono, *Frontiers in Microbiology* **2013**, *4*, 171, doi: 10.3389/fmicb.2013.00171.
88. J. K. G. Prince, R. H. Rainbird, *Geology* **2019**, *47*, 1–5.
89. J. E. Johnson, S. M. Webb, K. Thomas, S. Ono, J. L. Kirschvink, W. W. Fischer, *Proc. Natl. Acad. Sci. USA* **2013**, *110*, 11238–11243.
90. M. A. Saito, D. M. Sigman, F. M. M. Morel, *Inorg. Chim. Acta* **2003**, *356*, 308–318.
91. R. J. P. Williams, J. J. R. Fraústo da Silva, *The Natural Selection of the Chemical Elements*, Clarendon Press, Oxford, UK, 1996.
92. J. J. R. Fraústo da Silva, R. J. P. Williams, *The Biological Chemistry of the Elements.* Oxford University Press, Oxford, UK, 2001.
93. H. Küpper, P. M. H. Kroneck, *Met. Ions Life Sci.* **2007**, *2*, 31–62.
94. M. C. Weiss, F. L. Sousa, N. Mrnjavac, S. Neukirchen, M. Roettger, S. Nelson-Sathi, W. F. Martin, *Nature Microbiology* **2016**, *1*, 16116. doi: 10.1038/nmicrobiol.2016.116.
95. R. P. Hausinger, *Curr. Op. Struct. Biol*. **2019**, *59*, 1–8.
96. J.-H. Jeoung, H. Dobbek, *Proc. Natl. Acad. Sci. USA* **2018** , *115*, 2994–2999.
97. B. Desguin, T. Zhang, P. Soumillion, P. Hols, J. Hu, R. P. Hausinger, *Science* **2015**, *349*, 66–69.
98. B. Desguin, P. Soumillion, P. Hols, R. P. Hausinger, *Proc. Natl. Acad. Sci. USA* **2016**, *113*, 5598–5603.
99. T. Xu, M. D. Wodrich, R. Scopelliti, C. Corminboeuf, X. Hu, *Proc. Natl. Acad. Sci. USA* **2017**, *114*, 1242–1245.
100. J. A. Rankin, R. C. Mauban, M. Fellner, B. Desguin, J. McCracken, J. Hu, S. A. Varganov, R. P. Hausinger, *Biochemistry* **2018**, *57*, 3244–3251.
101. B. Desguin, M. Fellner, O. Riant, J. Hu, R. P. Hausinger, P. Hols, P. Soumillion, *J. Biol. Chem.* **2018**, *293*,12303–12317.
102. B. Desguin, *J. Bacteriol. Parasitol.* **2018**, *9*, 335. doi:10.4172/2155-9597.1000335
103. O. S. Fisher, G. E. Kenney, M. O. Ross, S. Y. Ro, B. E. Lemma, S. Batelu, P. M. Thomas, V. C. Sosnowski, C. J. DeHart, N. L. Kelleher, T. L. Stemmler, B. M. Hoffman, A. C. Rosenzweig, *Nature Comm*. **2018**, *9*, 4276 |DOI: 10.1038/s41467-018-06681-5.
104. L. J. Daumann, *Angew. Chem. Int. Ed.* **2019**, *58*, 2–11.

105. B. Jahn, A. Pol, H. Lumpe, T. R. M. Barends, A. Dietl, C. Hogendoorn, H. J. M. Op den Camp, L. J. Daumann, *ChemBioChem*. **2018**, *19*, 1147–1153.
106. H. Lumpe, A. Pol, H. J. M. Op den Camp, L. J. Daumann, *Dalton Trans.* **2018**, *47*, 10463–10472.
107. P. Kalimuthu, L. J. Daumann, A. Pol, H. J. M. Op den Camp, P. V. Bernhardt, *Chem. Eur. J.* **2019**, *25*, 1–10.

2

Sulfur, the Versatile Non-metal

Martha E. Sosa Torres,[*1] ***Alfonso Rito Morales,***[1]
Alejandro Solano Peralta,[1] ***and Peter M. H. Kroneck***[2]

[1]Departamento de Química, Inorgánica y Nuclear, Facultad de Química,
Universidad Nacional Autónoma de México, Ciudad Universitaria, 04510, CDMX, México
<mest@unam.mx>
<ponchori@comunidad.unam.mx>
<alexsolano@unam.mx>

[2]Department of Biology, University of Konstanz, D-78457 Konstanz, Germany
<peter.kroneck@uni-konstanz.de>

Abstract: The non-metallic chemical element sulfur, $^{32}_{16}S$, referred to in Genesis as brimstone and identified as element by Lavoisier, is the tenth most abundant element in the universe and the fifth most common element on Earth. Important inorganic forms of sulfur in the

Metal Ions in Life Sciences, Volume 20 Guest Editors: Martha E. Sosa Torres and Peter M. H. Kroneck
Series Editors: Astrid Sigel, Eva Freisinger, and Roland K. O. Sigel

https://doi.org/10.1515/9783110589757-002

biosphere are elemental sulfur (S_8), sulfate (SO_4^{2-}), and sulfide (S^{2-}), sulfite (SO_3^{2-}), thiosulfate, ($S_2O_3^{2-}$), and polythionates ($S_3O_6^{2-}$; $S_4O_6^{2-}$). Because of its wide range of stable oxidation states, from +6 to –2, sulfur plays important roles in central biochemistry as a structural and redox-active element and is intimately related to life on Earth. Unusual reaction pathways involving sulfur compounds become possible by the specific properties of this element. Sulfur occurs in all the major classes of biomolecules, including enzymes, proteins, sugars, nucleic acids, vitamin cofactors, and metabolites. The flexibility of these biomolecules follows from its versatile chemistry. The best known sulfur mineral is perhaps pyrite (Fool's gold), with the chemical formula, FeS_2. Sulfur radical anions, such as $[S_3]^{\bullet-}$, are responsible for the intense blue color of lapis lazuli, one of the most desired and expensive artists' materials. In the microbial world, inorganic sulfur compounds, e.g., elemental sulfur and sulfate, belong to the most important electron acceptors. Studies on microbial sulfur metabolism revealed many novel enzymes and pathways and advanced the understanding on metabolic processes used for energy conservation, not only of the microbes, but of biology in general. Transition metal sulfur complexes display intriguing catalytic activities, they provide surfaces and complex cavities in metalloenzymes that activate inert molecules such as H_2, CO, N_2 or N_2O, and they catalyze the transformations of numerous organic molecules. Both thiamine diphosphate- (ThDP) and *S*-adenosyl-L-methionine- (SAM) dependent enzymes belong to Nature's most powerful catalysts with a remarkable spectrum of catalytic activities. In conclusion, given sulfur's diverse properties, evolution made an excellent choice in selecting sulfur as one the basic elements of life.

Keywords: adenosine 5′-phosphosulfate · biotin · coenzyme M · cyclo-octasulfur · hydrogen sulfide · iron-sulfur · lapis lazuli · pyrite · *S*-adenosyl-L-methionine · sulfide · thiamine · thiol

1. INTRODUCTION: THE MANY FACES OF SULFUR – FROM BRIMSTONE TO FOOL'S GOLD TO LAPIS LAZULI

Evolution has made intriguing choices in selecting the basic elements of life. Although metal ions, such as sodium, potassium, magnesium, calcium, manganese, iron, or copper, are central to all of biochemistry because of their functions as cofactors, as part of cofactors in enzymes, and as structural elements in proteins, a group of elements clustered in the early part of the periodic table – carbon, nitrogen, phosphorus, oxygen, and sulfur – are represented in the molecules of life [1]. The ocean represents a major reservoir of sulfur on Earth, with large quantities in the form of dissolved sulfate (SO_4^{2-}) and sedimentary minerals, e.g., gypsum ($CaSO_4 \times 2H_2O$) and pyrite (FeS_2). Sulfate is the most stable form of sulfur on today's oxic Earth. Weathering and leaching of rocks and sediments are its main sources to the ocean. The natural release of volatile organic sulfur compounds from the ocean, mainly as dimethyl sulfide, $(CH_3)_2S$, transports sulfur from the ocean to terrestrial regions, and it also affects atmospheric chemistry and the climate system [2].

The non-metallic chemical element sulfur, $^{32}_{16}S$, referred to in Genesis as brimstone and correctly identified as an element in 1777 by Antoine Lavoisier, is the tenth most abundant element in the universe and the fifth most common element on Earth [3]. Sulfur has two electrons in the outer s orbital and four electrons in the p orbitals (s^2p^4 configuration), and it is one of the most reactive of the elements within the Periodic Table. Pure sulfur, also called cyclo-octasul-

fur (cyclo-S_8), is the most common allotrope of sulfur [4]. It is a tasteless, odorless, crystalline solid, bright yellow in color, a poor conductor of electricity, and quite insoluble in water (5 μg/L at 298 K) [5]. Noteworthy is also its tendency to form chain and ring polymers, reflecting the remarkable strength of the S−S bond (266 kJ mol^{-1}) [6]. Although studied for centuries, sulfur remains at the center of extensive research by chemists, biochemists, microbiologists, and geochemists. This includes research on improving energy efficiency, environmentally friendly uses for oil refinery waste products, development of polymers with unique optical and mechanical properties, and materials produced for biological applications [7].

Sulfur occurs in all major classes of biomolecules, including enzymes, proteins, sugars, nucleic acids, vitamin cofactors, and metabolites. The flexibility of sulfur-containing biomolecules follows from the versatile chemistry of this element. As members of the same periodic group, sulfur and oxygen share similarities in chemical reactivity. However, sulfur's position in the periodic table endows its compounds with distinct properties that are advantageous to biological systems. For example, thiols (R-SH) are superior nucleophiles compared with alcohols (R-OH) and also serve as versatile activating groups in thioester biochemistry. Disulfide bonds (RS–SR) are more stable than peroxide bonds (RO–OR), and biology has taken advantage of this stability by using disulfide bonds as structural features of proteins. Sulfur can adopt a wider variety of oxidation states than oxygen, for instance in the oxidation of cysteine to sulfenic (R-SOH), sulfinic (R-SO_2H), and sulfonic (R-SO_3H) acids. Sulfur is less electronegative than oxygen, permitting it to more readily adopt a positive charge, which is used by a number of enzyme cofactors. Sulfur is big enough to be able to populate 3d orbitals, which distinctly sets it off from oxygen, and it has a much lower tendency to form hydrogen bonds than oxygen. Sulfur can occur in formal valencies from –2 to +6, i.e., it seems equally prone to donate electrons to reach the electron configuration of neon as it is to accept electrons to adopt the argon configuration. Today we know, sulfur forms highly covalent bonds, as has recently been confirmed by sulfur K-edge X-ray absorption spectroscopy [8, 9]. There is little energy change involved for sulfur to go from being an electrophile as, say, in disulfide, to being a nucleophile as in a thiol. Thus, sulfur makes and breaks bonds easily; elementary sulfur occurs in rings of eight and in open sulfur chains, as we find them in polysulfides; and it is easy to exchange sulfur isotopes into these structures.

Most sulfur-containing minerals are metal sulfides, and the best known is perhaps the mineral pyrite, also called Fool's gold, with the chemical formula Fe(II) disulfide, FeS_2 ($Fe^{2+}S_2^{2-}$), containing clear S–S bonds [3, 10]. Recently, one of the most peculiar deep-sea hydrothermal-vent gastropods was discovered, referred to as the 'scaly-foot gastropod' (*Chrysomallon squamiferum*) [11, 12]. This gastropod has been found at depths of about 2.4–2.8 km, and it attracted much attention because of its hundreds of black metallic sclerites covering the foot (Figure 1).

The shell is of a unique construction: the outer layer consists of iron sulfides, the middle layer is equivalent to the organic periostracum found in other gastro-

Figure 1. Deep-sea hydrothermal-vent gastropod *Chrysomallon squamiferum* (scaly-foot gastropod) from the Kairei vent field; the outer layer of the shell consists of iron sulfides, the middle layer is equivalent to the organic periostracum found in other gastropods, and the innermost layer is made of aragonite ($CaCO_3$). Reproduced from [12]; open-access article distributed under the terms of the Creative Commons Attribution License.

pods, and the innermost layer is made of aragonite ($CaCO_3$). Features of this composite material are in the focus of researchers for possible use in protective applications [13]. The foot is also unusual, being armored at the sides with iron-mineralized sclerites. The soft tissue core of each sclerite is covered in conchiolin, which is in turn covered with pyrite (FeS_2) and greigite (Fe_3S_4), making it the only extant metazoan known to utilize iron sulfide as part of its skeleton. The scales, coated with iron sulfides, are heavily colonized by ε- and δ-proteobacteria, and it was thought that the microbes perhaps participate in the mineralization of the sclerites. However, S and Fe isotopic analyses indicate that both S and Fe in the sclerites originate from hydrothermal fluids rather than from bacterial metabolites. The magnetic properties of the sclerites are not optimized for magnetoreception and instead support use of the magnetic iron sulfide minerals as structural elements. In any case, this novel metazoan-microbial collaboration illustrates the great potential of organismal adaptation in chemically and physically challenging deep-sea environments [14–17].

There is also a *blue side* of sulfur. When burned, elemental sulfur emits a blue flame with formation of sulfur dioxide (SO_2), which has a stinging smell. When dissolved in ionic liquids, the solution develops an intense blue color, due to the presence of the $[S_3]^{\cdot-}$ radical anion [18–28]. Electron paramagnetic resonance (EPR) played a key role in identifying and characterizing the $[S_3]^{\cdot-}$ radicals in ultramarine blue materials, and Gardener and Fraenkel reported the initial

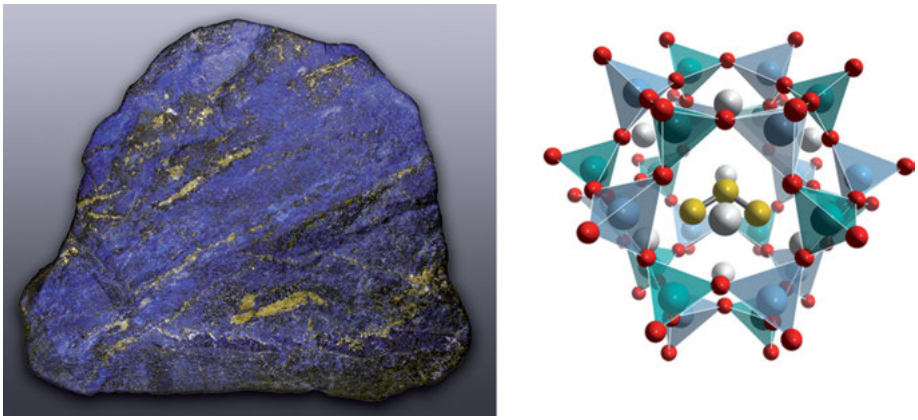


Figure 2. Lapis lazuli, the deep blue metamorphic rock used as a semi-precious stone since antiquity. **Left**: Lapis lazuli with pyrite from Afghanistan in its natural state; reproduced with permission by H. Grobe, Creative Commons CC-BY-SA-2.5; https://en.wikipedia.org/wiki/Lapis_lazuli. **Right**: Chemical structure of lazurite, with frameworks of alternating silica and alumina tetrahedra creating large cages; white spheres are Na^+ cations, entrapped in the center the yellow $S_3^{\cdot-}$ radical anion responsible for the blue color. Reproduced from https://chemicalstructure.net/portfolio/lazurite/.

observation of an EPR spectrum and attributed it to a sulfur-centered radical [18]. It has been used as a reference material for EPR spectroscopy, since its g factor = 2.0029 is well separated from the free electron value, $g = 2.0023$, and since the EPR signal does not saturate readily, it is a useful microwave power monitor [21, 24, 25]. Sulfur radical anions are also responsible for the intense blue color of lapis lazuli, a metamorphic rock used as a semi-precious stone. As early as the 7th millennium BCE, lapis lazuli was mined in Afghanistan, and it was used in the funeral mask of Tutankhamun [29]. As one of the most desired and expensive artists' materials throughout history, there has long been interest in studying natural lapis lazuli and get a deeper insight in the traditional method of extracting its blue component, lazurite [30]. Lapis lazuli is a complex rock whose composition is defined by the presence of the mineral lazurite $(Na,Ca)_8(AlSiO_4)_6(SO_4,S,Cl)_2$, which is responsible for its overall blue hue (Figure 2). Inclusions of several other minerals are also common, including pyrite and calcite ($CaCO_3$) in varying amounts. Lazurite belongs to the group of sodalite minerals which contain frameworks of alternating silica and alumina tetrahedra creating large cages. Extra-framework ions (Ca^{2+}, K^+, Na^+, Cl^-, OH^-, SO_4^{2-}, S_n^{2-}), and neutral species (H_2O) are entrapped within these cages. Lazurite's blue color is attributed to sulfur polyanion radicals trapped in the cage structure. Variations in color appear to be related to ratios between various sulfur species: the trisulfur radical $[S_3]^{\cdot-}$ is mainly responsible for the blue color, but contributions from disulfur $[S_2]^{\cdot-}$ and tetrasulfur $[S_4]^{\cdot-}$ radicals can shift the color towards yellow, green, or red, respectively [22–24, 26–28, 30].

Until recently, geochemists believed that inside Earth, only two forms of molecules contained sulfur: (i) sulfides (based on H_2S or S^{2-}), and (ii) sulfates (based

on H_2SO_4 or SO_4^{2-}). The detection of $[S_3]^{\cdot-}$ during these experiments indicates that sulfur must be considerably more mobile in hydrothermal fluids in the Earth's crust than was previously thought. This is because, unlike sulfides and sulfates, which attach to minerals as soon as they appear in fluids, $[S_3]^{\cdot-}$ proves to be extremely stable in the aqueous phase. In other words, below ground these ions must flow for long distances in dissolved form, taking with them the noble metals to which they may be bound [31–33].

In view of the immensely rich literature on the chemistry, biochemistry, and geochemistry of sulfur, we suggest for introduction the article by iron-sulfur pioneer Helmut Beinert entitled "A tribute to sulfur" published in 2000 [34]. For further insight into the topic of transition metals and sulfur and the biogeochemical cycle of sulfur, broader perspectives and more inclusive of the work of experts in this dynamic research area are available in very informative and authoritative comments and reviews [35–55].

In this chapter we will introduce the reader to several important aspects of sulfur biochemistry, microbiology, and geochemistry, and we will also take a glimpse at its unique coordination chemistry which will be discussed in greater detail in the following chapters of this volume.

2. SULFUR AND ENERGY CONSERVATION

Like oxygen and selenium, sulfur belongs to group 16 of the periodic table of the elements, notably it is about 1000 times less abundant in nature than oxygen. Because of its wide range of stable oxidation states, sulfur can play important roles in central biochemistry as a structural and redox-active element and is intimately related to life on Earth. The three most abundant forms of sulfur in the biosphere are elemental sulfur (S_8; formal oxidation state 0), sulfate (SO_4^{2-}; +6) and sulfide (S^{2-}; –2). Other prominent oxyanions of sulfur are sulfite (SO_3^{2-}; +4), dithionite ($S_2O_4^{2-}$; +3), thiosulfate, ($S_2O_3^{2-}$; +5/–1) and polythionates, such as trithionate ($S_3O_6^{2-}$; 0/+5) and tetrathionate ($S_4O_6^{2-}$; 0/+5) [56, 57]. Interconversions of these species constitute their biogeochemical cycles which are sustained by complex biological processes, with microbes playing a prominent role. Numerous studies suggest that the ability to reduce sulfate was developed early during prokaryotic evolution [58–63] (see Section 2.3 and Chapter 10 of this book). Unusual reaction pathways involving sulfur apparently become possible by the specific properties of this element [34].

2.1. Metal Sulfides and Early Life

How did life begin? Naturally, there can hardly be a bigger question. Although heavily disputed, many of the scientists studying the origin of life are confident today that they are on the right track – and they have the experiments to back up their confidence [64–77].

In his article entitled "Origin of Life: RNA World versus Autocatalytic Anabolism", published in 2006 [78], chemist Günter Wächtershäuser wrote: The deep past of the earth is unobservable. Therefore, the problem of the origin of life, the emergence of the first evolvable entity, which is the primordial ancestor of all extant organisms, can only be solved by a theory. Theories on the origin of life are scientific rather than myth, if they have empirical significance: empirical biological significance by providing evolutionary explanations for extant facts of biology and/or empirical chemical significance by predicting unknown but testable chemical reactions. Preferably, they also should have geological significance by being compatible with geological theories on the early history of the earth, which themselves, however, need to be scrutinized for their power to explain and predict facts of geology and chemistry. Only two theories on the origin of life appeared detailed enough for an evaluation by the above criteria: (i) the *RNA world theory* which assumes that the first organism was a "living" RNA-like molecule replicating in a prebiotic broth of activated nucleotides, and (ii) the *autocatalytic anabolist theory* (also called *iron-sulfur world theory*) which assumes that life began on minerals with an anabolic metabolism of synthetic, autocatalytic carbon fixation cycles [78]. According to Wächtershäuser, the earliest form of life, termed "pioneer organism", originated in a volcanic hydrothermal flow, at high pressure and high temperature. It had a composite structure of a mineral base with catalytic transition metal centers. These metal centers catalyzed autotrophic carbon fixation pathways generating non-polymer organic compounds from gases, such as carbon monoxide (CO), carbon dioxide (CO_2), hydrogen cyanide (HCN), and hydrogen sulfide (H_2S) (Equation 1). The reaction of FeS and H_2S led to the formation of pyrite (FeS_2) and dihydrogen and has been demonstrated under mild volcanic conditions [79–81].

$$FeS + H_2S \rightarrow FeS_2 + 2H^+ + 2e^- (= H_2) \quad \Delta G^0 = -41.9\ \text{kJ/Lmol} \qquad (1)$$

The organic compounds were retained on or in the mineral base as ligands of the transition metal centers with a flow retention time in correspondence with their mineral bonding strength thereby defining an autocatalytic "surface metabolism", producing more complex organic compounds, more complex pathways, and more complex catalytic centers. Notably, the biogeochemistry associated with deep-sea hydrothermal environments in the present and ancient deep oceans is driven by the unique microbial and chemical processes without energy input from sun light. Reductants, e.g., H_2S and H_2, are used as energy sources which are emitted from hydrothermal vents. Iron sulfides are continuously precipitated in such environments. Thus, their surface, hosting traces of transition metals, such as Ni, Cu, W, and Mo, has played a critical role in prebiotic organic synthesis and early evolution of energy metabolisms in ancient Earth.

Of course we do not know for sure how life started. However, if it was at high temperature in deep-sea vents, clearly iron and sulfur had to be of prime importance. There was plenty of them around and both belong to the chemically most versatile elements, sulfur more so than iron. The properties of both elements are

all mirrored in the behavior of Fe-S clusters, with iron, as a fitting partner, though not quite as versatile [42, 82–86]. Recently it was reported that particles of the Fe(II),Fe(III) sulfide mineral greigite can reduce CO_2 under ambient conditions to simple organic chemicals, such as methanol (CH_3OH), formic acid (HCOOH), acetic acid (CH_3COOH), and pyruvic acid ($CH_3COCOOH$) [85, 86]. Greigite (formula $Fe^{2+}Fe_2^{3+}S_4$) is the sulfur equivalent of the iron oxide magnetite, Fe_3O_4. The structural similarity between Ni-doped greigite and the Ni,Fe-S clusters present in biological enzymes has led to suggestions that greigite minerals could have acted as catalyst for the origin of life [87–90]. Note that the cubic Fe_4S_4 unit of greigite is observed in the [4Fe-4S] units of many important iron-sulfur enzymes involved in transformations of CO, H_2, or N_2 (see Chapters 7, 8, 10, and 11 of this book).

2.2. Sulfur Respiration

Given a choice of elements to reconstruct the redox history of the Earth, geochemists will choose sulfur. Many biogeochemical processes fractionate the multiple stable isotopes of sulfur in telltale ways, leaving their imprint in the sedimentary record [91–95]. The availability and speciation of sulfur in the early biosphere must have played an important role, the different oxidation states of this element (S^{2-}, S^0, $S_2O_3^{2-}$, SO_3^{2-}, or SO_4^{2-}) can all act as an important source of energy for different types of sulfur-metabolizing organisms.

In the microbial world, elemental sulfur (S^0) belongs to the most important inorganic electron acceptors. Studies on microbial sulfur metabolism revealed many novel enzymes and pathways and advanced the understanding on metabolic processes used for energy conservation in extreme environments, not only of the microbes, but of biology in general. Major processes include aerobic oxidation as well as anaerobic reduction of S^0 [96–98]. The microbial oxidation of inorganic sulfur compounds and elemental sulfur to sulfate is one of the major reactions in the global sulfur cycle. Notably the process of biological sulfur oxidation is important in bioleaching operations for the industrial bioleaching of metal sulfides or heavy metal recovery from industrial waste [98–101]. Based on the 1.7 Å crystal structure of the sulfur oxygenase, a reaction scheme is proposed for the aerobic S^0 oxidation in the thermoacidophilic archaeon *Acidianus ambivalens* (Figure 3). It includes the O_2-dependent disproportionation of S^0 to produce sulfite and hydrogen sulfide (Equations 2–4), thiosulfate is likely produced from a non-enzymatic reaction (Equation 5) [102, 103].

$$S^0 + O_2 + H_2O \rightarrow HSO_3^- + H^+ \quad \text{(oxygenation)} \quad (2)$$

$$3S^0 + 3H_2O \rightarrow HSO_3^- + 2HS^- + 3H^+ \quad \text{(disproportionation)} \quad (3)$$

$$4S^0 + O_2 + 4H_2O \rightarrow 2HSO_3^- + 2HS^- + 4H^+ \quad \text{(sum)} \quad (4)$$

$$S^0 + HSO_3^- \rightarrow S_2O_3^{2-} + H^+ \quad \text{(non-enzymatic)} \quad (5)$$

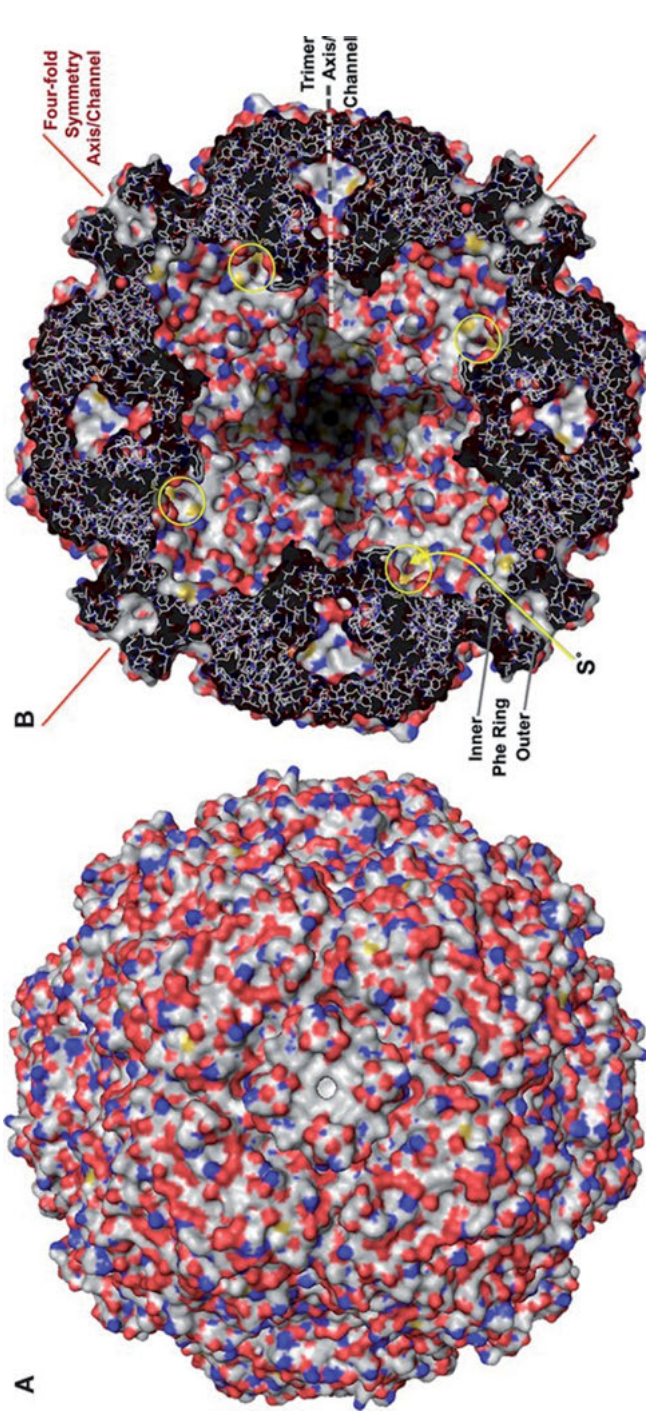


Figure 3. Crystal structure of the sulfur oxygenase of thermoacidophilic archaeon *Acidianus ambivalens,* PDB ID 2CB2, 2YAW, 2YAX, 2YAV [102, 103]. (**A**) Surface representation of the holoenzyme centered at the chimney-like structure at the fourfold symmetry axis; the spherical, hollow, cytoplasmic enzyme is composed of 24 identical subunits with an active site pocket, each comprising a mononuclear non-heme iron site and a cysteine persulfide. (**B**) Representation of the protein structure and the inner surface of the holoenzyme sliced at the center of the fourfold symmetry axes; the position of the inner and outer phenylalanine rings are indicated, also the active site pores (yellow circles), and the approximate position of the trimer symmetry axis, which is tilted out of plane. Reproduced with permission from [103]; copyright 2011, Veith, Urich, Seyfarth, Protze, Frazão, and Kletzin. This is an open-access article subject to an exclusive license agreement between the authors and Frontiers Media SA, which permits unrestricted use, distribution, and reproduction in any medium, provided the original authors and source are credited.

A cysteine persulfide, CyS-S$^-$, and a low-potential mononuclear non-heme iron site ligated by a 2-His-1-carboxylate facial triad (His86, His90, Glu114) in a pocket constitute the active site. It is proposed that a cysteine-bound sulfane sulfur chain is the substrate of disproportionation and oxygenation at the iron site and, further, that dissolved and linear sulfur species are the actual substrates and not the regular S_8 ring.

Coming to anaerobic sulfur reduction, the ability to reduce elemental sulfur to H_2S (Equation 6) using H_2 or organic substrates as electron donors is widespread among bacteria and archaea [44].

$$S^0 + H_2 \rightarrow H_2S \quad (6)$$

There exists a highly specialized group of obligate anaerobes which use dissimilatory sulfur reduction as the main, if not the sole, catabolic process [104]. True dissimilatory sulfur-reducing bacteria were discovered by Pfennig and Biebl [105]. These bacteria preferentially develop together with sulfur-excreting phototrophic green-sulfur bacteria. It was demonstrated for *Desulfuromonas acetoxidans* that the energy-conserving reduction of elemental sulfur to hydrogen sulfide was stoichiometrically linked to the oxidation of acetate to CO_2. This process is called sulfur respiration in analogy to the reduction of dioxygen (O_2) to water (H_2O), and it was shown that sulfur respiration was coupled to phosphorylation [39, 106–112]. As expected from their metabolic activities, from preliminary work on these organisms, and from the sheer appearance of their extracts, the sulfur-reducing bacteria proved to be a rich source of iron-sulfur proteins and cytochromes as shown by low temperature EPR studies [113–115]. Elemental sulfur is particularly important for the metabolism of extremely thermophilic archaea, which show an optimal growth temperature higher than 80 °C. Thomas Brock was the first person to observe such extremophiles, and his pioneering work initiated investigation of thermophilic bacteria in volcanic areas [40, 116].

In view of its low solubility in water S^0 cannot be easily consumed by microbes. So-called "hydrophilic" or "colloidal" sulfur, and organic trisulfides, RS-S^0-SR, have been reported to produce H_2S with considerable velocities in the presence of enzyme preparations obtained from sulfur-reducing bacteria [117, 118]. Notably, elemental sulfur is readily converted to polysulfide in aqueous solution by reaction of elemental sulfur with sulfide, the product of sulfur respiration (Equation 7), and the sulfur-reducing microorganism *Wolinella succinogenes* was shown to grow on polysulfide as electron acceptor converting it to hydrogen sulfide (Equation 8) [39, 119, 120].

$$nS^0 + HS^- \rightarrow S^{2-}_{n+1} + H^+ \quad (7)$$

$$S^{2-}_{n+1} + 2e^- + 2H^+ \rightarrow S^{2-}_n + H_2S \quad (8)$$

Recently, the 2.4 Å structure of the polysulfide reductase (Psr) from bacterium *Thermus thermophilus* was resolved (PDB ID 2VPZ), revealing how the integral membrane protein complex recognizes and reduces its unique substrate S^{2-}_{n+1} [121]. The quinone-coupled reduction of polysulfide represents a process important in extreme environments, usually volcanic or geothermically active, such as

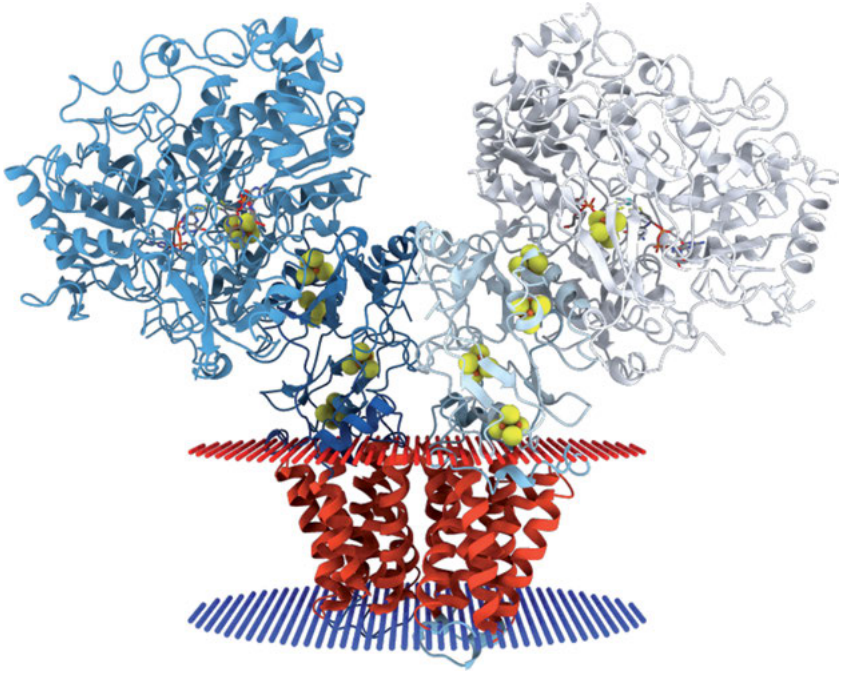


Figure 4. Crystal structure (2.4 Å) of the polysulfide reductase (Psr) from bacterium *Thermus thermophilus*; ribbon representation of the PsrABC dimer viewed parallel to the membrane, subunit C in red; each catalytic PsrA subunit hosts one Mo molybdopterin guanine dinucleotide (bis-MGD) cofactor and five [4Fe-4S] clusters; in the catalytic cycle of Psr, menaquinol (MQ) is reduced on the periplasmic side of the membrane, releasing two protons and electrons; the electrons are transported via the Fe-S clusters to the active-site Mo, where polysulfide is reduced with the evolution of hydrogen sulfide; image from the RCSB PDB of PDB ID 2VPZ [121].

deep-sea vents and hot springs. Psr is a key energy-conserving enzyme of the *T. thermophilus* respiratory chain, using S^{2-}_{n+1} as the terminal electron acceptor and pumping protons across the membrane via a previously unknown mechanism. The enzyme belongs to the well-studied family of molybdenum- or tungsten-containing oxidoreductases, the PsrA subunit hosts the catalytic Mo center ligated by the molybdopterin guanine dinucleotide cofactor (bis-MGD), five cubane-type [4Fe-4S] clusters serve as electron transfer agents (Figure 4). FAD and NAD(P)H-dependent coenzyme A disulfide reductases/polysulfide reductases have been proposed to be important constituents in the sulfur-reducing hyperthermophiles *Pyrococcus horikoshii* and *Pyrococcus furiosus*. The enzyme from *P. horikoshii* was reported to reduce efficiently a range of disulfides (RS-SR), persulfides (RS-SH), and polysulfide compounds. Its role in the reduction of elemental sulfur *in vivo* remains unclear, it might be a part of a sulfur-dependent antioxidant system. The 2.7 Å structure (PDB ID 4FX9) reveals a relatively restricted substrate channel leading into the sulfur-reducing side of the FAD isoalloxazine ring, suggesting how this enzyme class may select for specific disulfide substrates [122].

2.3. Microbial Activation of Sulfate

Dissimilatory reduction of sulfate to hydrogen sulfide, with sulfite as crucial intermediate, is one of the oldest and most prominent microbial energy-conserving pathways on Earth. As life may have originated in hot environments, the occurrences of sulfate-reducing hyperthermophilic archaea and deep-branching thermophilic bacteria indicate an early origin of this process. Isotopic data suggest that dissimilatory sulfate reduction began over 3 billion years ago but acquired global significance only after sulfate concentrations had considerably increased in the Precambrian oceans approximately 2.35 billion years ago. Sulfate-reducing microorganisms are ubiquitous and play an imperative role in the global cycling of carbon and sulfur [57–59, 123–131].

Sulfate is an energy-rich molecule that is, however, chemically inert. To use this energy source, microorganisms have to invest ATP [58]. Three enzymes are the key players in the dissimilatory reduction of sulfate to hydrogen sulfide: (i) ATP sulfurylase, (ii) adenosine 5′-phosphosulfate reductase, and (iii) dissimilatory sulfite reductase (Equations 9–11). Because of its low redox potential ($E^{0\prime}$ = –516 mV), sulfate cannot be directly reduced by H_2 or organic acids, it has to be converted to adenosine 5′-phosphosulfate (APS) in a reaction catalyzed by ATP sulfurylase (Figure 5). Hereby, the redox potential $E^{0\prime}$ (APS/AMP + HSO_3^-) is shifted to –60 mV. The formation of APS is endergonic and probably driven by the subsequent hydrolysis of pyrophosphate and the energetically favorable APS reduction. Therefore, the activation of sulfate to APS is assumed to consume two ATP equivalents. APS reductase, a FAD, [4Fe–4S] enzyme, converts APS to sulfite and AMP, followed by the six-electron reduction of sulfite to hydrogen sulfide as the final step ($E^{0\prime}$ (HSO_3^-/HS^-)= –116 mV).

$$\text{Mg-ATP} + SO_4^{2-} \rightarrow \text{Mg-PP}_i + \text{APS} \quad (9)$$

$$\text{APS} + 2e^- \rightarrow \text{AMP} + HSO_3^- \quad (10)$$

$$HSO_3^- + 6e^- + 6H^+ \rightarrow HS^- + 3H_2O \quad (11)$$

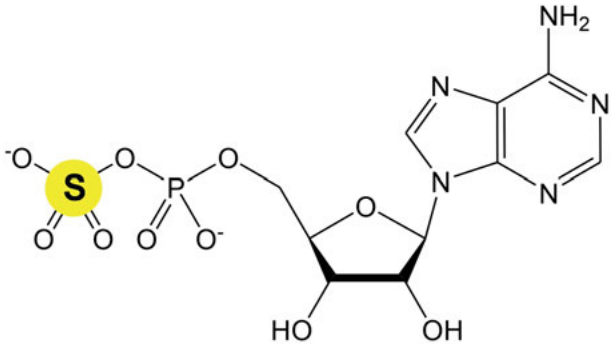


Figure 5. Schematic representation of adenosine 5′-phosphosulfate (APS).

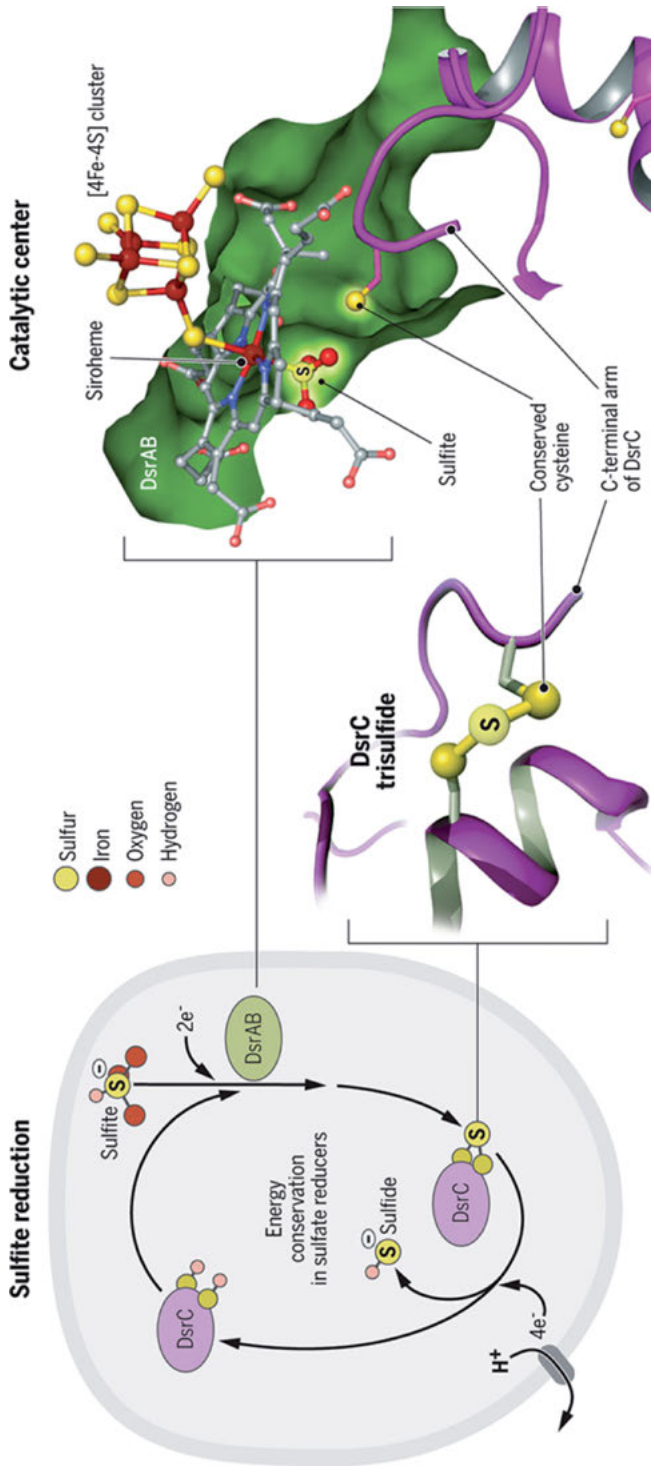


Figure 6. Mechanism of dissimilatory sulfite reductase (DsrAB) and energy coupling by the DsrC protein trisulfide in sulfate-reducing bacteria. **Left**: Dissimilatory reduction of sulfite by DsrAB leads to the formation of the DsrC trisulfide; it links the exergonic reaction $SO_3^{2-} \rightarrow H_2S$ to energy conservation in sulfate-reducing bacteria. **Middle**: Model of the DsrC trisulfide in the structure of *Archaeoglobus fulgidus* DsrC (PDB ID 1SAU). **Right**: The coupled siroheme–[4Fe-4S] catalytic site of DsrAB binds SO_3^{2-} through its sulfur atom to the siroheme Fe center (PDB ID 3MM5); the C-terminal arm of the DsrC protein extends with the conserved cysteine toward the bound sulfite to pick up the reduced sulfur-oxygen intermediate. Reproduced with permission from [132] (G. Fritz, P. M. H. Kroneck, *Science* **2015**, *350*, 1476–1477); copyright 2015, Science, AAAS.

This multi-electron, multi-proton transfer process is catalyzed by dissimilatory sulfite reductase which contains the coupled siroheme-[4Fe-4S] cofactor at the active site (see Chapter 10). In total, the equivalent of 2/3 ATP will be generated in this pathway. Two molecules ATP, out of three, are required for the activation of sulfate and 1/3 ATP for the transport of the sulfate into the cell. Recently, a protein-based trisulfide was identified as key intermediate that couples the reduction of sulfite to energy conservation (Figure 6) [132–134].

3. SULFUR AND METAL COORDINATION

Because of the restricted scope of this chapter, and in view of the following chapters written by renowned experts in the field of transition metal sulfur sites in biological systems, only a short overview and selected bibliography of this topic will be given. It is well known that sulfur-containing compounds can act as a poison for catalysts because of their strong coordinating and adsorptive properties, which cause them to block reactive metal sites. On the other hand, numerous transition metal sulfides display intriguing catalytic activities in their own right. They provide surfaces (solid state) and complex cavities (metalloenzymes) that activate small molecules such as H_2, CO, N_2 or N_2O, and they catalyze the reductive transformations of organic molecules [31, 135–141].

The chemistry of the transition metals has been exploited extensively in both biology and industry. In many applications, the metal is coordinated by sulfur, either in the form of a sulfur-containing organic ligand or in the form of a variety of inorganic sulfur-donor groups. Sulfur coordination is required for the functioning of numerous biological catalytic metal centers. These metal sites encompass different transition metals, such as V, Fe, Ni, Cu, Zn, Mo or W, and both mononuclear and homo- as well as hetero-polynuclear architectures have been found. The role of sulfur involves the modulation of the activity of the metal, but often the sulfur ligand itself can be involved in substrate binding, acid-base activity, or redox processes crucial to catalysis. The industrial interest in metal sulfur compounds includes applications in catalysis, anticorrosion, lubrication, antioxidancy, and battery technology [141].

Reports on metal ions bound to proteins or enzymes, in terms of the coordination compounds of the 1913 Nobel laureate Alfred Werner [142], date way back into the 19th century and may probably be found in earlier centuries if the terms “proteins” and “enzymes” will be replaced with “animal or plant tissues”. Naturally, the major polymers in living matter are carbon compounds, whereas the transition metals are usually only present in traces. However, there is no life without transition metals [143]. Obviously, there is a long trail of discovery of metals in proteins or in other components of living organisms, metals that were required for their structure or function, although details of what the functions implied were initially missing and hard to come by [144]. Since the early days of *Bio(inorganic) chemistry*, or *Inorganic Biochemistry*, when the nature of the metal constituents of the respiratory enzyme (*Atmungsferment*) cytochrome *c* oxi-

dase were still a matter of debate between Otto Warburg and David Keilin [145], ever more complex metal-dependent proteins have been purified and crystallized, physical techniques and attendant theories can probe deeply into the intricacies of electronic structure and structural dynamics, chemical syntheses of protein and metal coordination units become more sophisticated, and biochemical syntheses with site-directed mutagenesis disclose functions of specific amino acid residues. Basically, proteins, although somewhat complicated, reduce to a spatially correlated multidentate ligand. The proteins act as a unique ligand, its structure and environment can modulate important properties such as electronic structure, redox potential, and detailed stereochemistry [41].

Proteins coordinate metal ions (M) with nitrogen, oxygen, and sulfur *endogenous* biological ligands; the amino acids cysteine (Cys; RS^-) and methionine (Met; RS-CH_3) provide thiolate and thioether sulfur for M coordination. Bonding depends on the M-SR or M-SR-CH_3 angle. For RS^-, the dominant valence orbitals involved in bonding are the two sulfur p orbitals which are perpendicular to the C–S bond. These orbitals are degenerate in the free base, but split in energy into π and pseudo-σ levels upon bonding to a metal with a M-S-C angle less than 180°. The π orbital is perpendicular to the M-S-C plane, while the pseudo-σ orbital is in the plane with its lobe pointed toward the metal ion but not necessarily along the bond axis unless the M-S-C angle is close to 90°; the M-S-C angles of metal thiolates are mainly in the range of 100–120°. RS-CH_3 ligands are often found to bind the metal ion below the molecular plane and approximately 40° off the plane normal; two valence orbitals (b_1, a_1) are involved in bonding, which are p orbitals on the sulfur atom perpendicular to and largely in the ligand plane, respectively. Sulfide (S^{2-}), like water (H_2O), hydroxide (HO^-), and oxide (O^{2-}) belong to the class of exogenous ligands. The frequent occurrence of sulfido bridges (as well as of oxo bridges) and terminal ligands arises from the enhanced acidities of their conjugate acids when coordinated to sufficiently oxidized metal ions, together with the considerable bond strengths associated with binding these intensely nucleophilic dianions, with the most prominent examples observed in the bridging units Fe(III)-S-Fe(III)/(II) of iron-sulfur clusters and terminal oxo-sulfido groups in molybdenum and tungsten proteins [41, 146–148]; (see Chapters 7–11 of this book).

Generally, in metalloproteins two types of metal sulfur sites are found: (i) mononuclear centers with specific protein or cofactor ligation, such as the blue type 1 copper electron transfer center, $CuS_{Cys}(N_{His})_2S_{Met}$, in plastocyanin [149], the $CuN_{His}(S_{Met})_2$ site in Cu trafficking protein CusF [150], or the Cu(I) S_{Met}-X_3-S_{Met} binding motif in α-synuclein [151], and (ii) polynuclear sites with μ-sulfido bridges connecting two or more metals, e.g., the centers in iron-sulfur proteins, [2Fe-2S] and [4Fe-4S] [82, 152], or the four-metal cubane [Ni,3Fe-4S] cluster in anaerobic carbon monoxide dehydrogenase [153]; hereby the iron atoms are coordinated both by the bridging sulfur atoms (also called *inorganic* or *acid-labile* sulfur) and by cysteine thiolate sulfurs (also called *organic sulfur*) [82, 154–156].

A common structural motif present in electron transfer protein active sites, e.g., the blue type 1 Cu center in plastocyanin, the purple mixed-valence CuA

center [$Cu^{1.5+}$-$Cu^{1.5+}$] in cytochrome oxidase and nitrous oxide reductase, or the new type of CuA site observed in methane- and ammonia-oxidizing bacteria, and the Fe protein cytochrome *c*, is the presence of an axial metal–S_{Met} bond. This structural feature gained major interest in recent years in view of its significant influence on the redox potential of the electron transfer center [157–163].

4. SULFUR IN COFACTORS AND COENZYMES

Next to essential amino acids cysteine (Cys) and methionine (Met), biotin (vitamin B_7) and thiamine (vitamin B_1) (Figure 7) are important organosulfur compounds needed for biochemical functioning. Homocysteine, $(H_2N)(HOOC)CH$-CH_2-CH_2-SH [164], a homologue of cysteine and ligand of molybdenum or vanadium, in nitrogenase [165], and taurine, H_2N-CH_2-CH_2-SO_3H, an important source of hydrogen sulfide via anaerobic respiration of sulfite in the human gut [166], are other sulfur-containing acids, but are not part of the primary structure of proteins.

Many enzymes with important roles in cellular metabolism use prosthetic groups ending with a sulfhydryl (-SH) moiety, e.g., coenzyme A and lipoic acid [167]. The tripeptide glutathione, γ-glutamyl-cysteinyl-glycine, carries reducing equivalents for cellular repair of oxidation through its sulfhydryl group. In methanogenesis, a multistep biochemical transformation of carbon dioxide (CO_2) to methane (CH_4), several organosulfur cofactors are involved including the immediate precursor to methane, methyl-coenzyme M, CH_3-$SCH_2CH_2SO_3^-$ (CH_3-S-CoM). The Ni enzyme methyl-coenzyme M reductase harbors the Ni(I)-hydrocorphin coenzyme F-430 as a prosthetic group and catalyzes the reversible reduction of CH_3-S-CoM with coenzyme B, HS-CoM, to CH_4 and the heterodisulfide CoM-S-S-CoB [168–173].

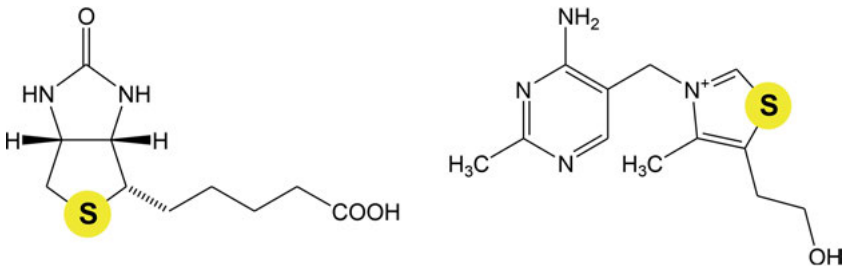


Figure 7. Schematic representation of the organosulfur compounds biotin (vitamin B_7; left) and thiamine (vitamin B_1; right).

4.1. Thiamine Diphosphate, Essential Coenzyme in All Forms of Life

Thiamine diphosphate (ThDP) is an essential organosulfur molecule that is required for processes of general metabolism amongst all organisms, it is the only naturally occurring thiamine derivative that is found as an enzyme cofactor. It consists of a sulfur-containing thiazolium nucleus, a 4′-aminopyrimidine ring, and a diphosphate moiety that is required to provide tight binding to the target enzyme via a Mg^{2+} cation (Figure 8). The electrophilic nature of the positively charged thiazolium ring is key to the stabilization of carbanion states of the cofactor, the ThDP-ylide, and of covalent substrate-ThDP conjugates in terms of a so-called 'Umpolung – or polarity inversion' mechanism [174–176]. This makes ThDP an effective participant in the reversible catalytic cleavage of C–C bonds of vicinal dicarbonyl or α-hydroxyketone groups, exhibiting a chemical ability not possessed by protein functional groups.

Since the fundamental experiments published by Ronald Breslow in 1958 on the mechanism of ThDP action [177], an enormous number of complex organic syntheses catalyzed by ThDP-dependent enzymes with high chemo-, regio-, and stereoselectivity was reported. They are exceedingly multifunctional biocatalysts, form a vast and diverse class of proteins, and catalyze a wide variety of enzymatic reactions including the formation or cleavage of C–S, C–O, and C–N bonds [178–186]. Most importantly, ThDP-dependent enzymes are involved in the making

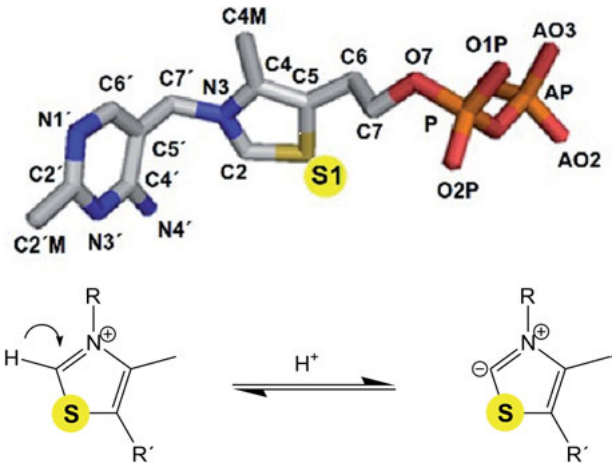


Figure 8. Stick representation of thiamine diphosphate (ThDP) in the typical V-conformation found in structures of ThDP-dependent enzymes; the N4′ amino nitrogen of the pyrimidine ring comes close to the thiazolium ring allowing formation of the ThDP-ylide ion; the C2-H moiety of the thiazole ring can act as an acid forming a carbanion; the tetravalent nitrogen just adjacent to C2 stabilizes the negative charge, making the reaction much more favorable; the diphosphate group binds to a Mg^{2+} cation, which forms a complex with the active site cavity [184], C (grey), N (blue), S (yellow), O (red), P (orange).

and breaking of C–C bonds, their synthetic applications increased exponentially since the 1990s thanks to the progress in molecular biology giving access to large amounts of highly pure enzymes and variants [187–190]. Because of the ability to form asymmetric C–C bonds, ThDP-dependent enzymes are versatile catalysts for important transformations and exhibit a remarkable substrate spectrum. Although very diverse in sequence and domain organisation, they share two common protein domains, the pyrophosphate and the pyrimidine domain. For the comprehensive and systematic comparison of protein sequences and structures the **T**hiamine diphosphate-dependent **E**nzyme **E**ngineering **D**atabase (TEED) was established in 2010 [191]. ThDP-dependent enzymes share a conserved fold [182, 183], in which paired active sites (~25 Å apart) are arranged in the interface between two subunits. Note that ThDP-dependent enzymes are typically obligate dimers, however, some associate further into higher oligomers. Activation of the organosulfur cofactor depends on the binding of its 4′-aminopyrimidine ring to the active site cavity that enables it to access catalytically relevant imino and pyrimidinium tautomers (Figure 8), and the adoption of a V-shaped conformation when the two cofactor rings are bent over a conserved hydrophobic residue. This brings the exocyclic N4′ of the pyrimidine ring close to the C2 atom of the sulfur-containing thiazolium ring, causing a significant decrease of the pK_a of this C-H moiety to facilitate its deprotonation and formation of a highly nucleophilic ylide species [174–176, 192].

4.2. *S*-Adenosyl-L-Methionine, Key Player in the Radical SAM Pathway

Originally, free radicals had been regarded as rather rare species in enzymatic catalysis [193]. The most convincing evidence for the participation of radicals in enzymatic reactions was obtained by magnetic resonance techniques, such as the flavin semiquinone [194], the quinone cofactors in electron transfer reactions, and the tyrosyl radical in ribonucleotide reductases [195]. More recently, two important classes of dioxygen-independent enzymatic reactions were recognized to proceed through radical-based mechanisms, those catalyzed by *S*-adenosyl-L-methionine/Fe-S enzymes and the adenosylcobalamin-dependent enzymes (which do not contain sulfur and therefore are not discussed here) [45, 46, 196–213]. *S*-adenosyl-L-methionine (SAM or AdoMet) is formed by the reaction of L-methionine and ATP as first reported by Cantoni in 1952 [214]. A metabolite present in all living cells, SAM plays a central role in cellular biochemistry as a precursor to methylation, aminopropylation, and transsulfuration pathways. It is well known as the methyl donor for the majority of methyltransferases that modify DNA, RNA, histones and other proteins, dictating replicational, transcriptional, and translational fidelity, mismatch repair, chromatin modelling, epigenetic modifications and imprinting, which are all topics of great interest and importance in cancer research and aging [215]. SAM acts as the precursor in the biosynthesis of the polyamines spermidine and spermine, of the metal ion chelating compounds nicotianamine and phytosiderophores, and of the gaseous plant hormone ethylene

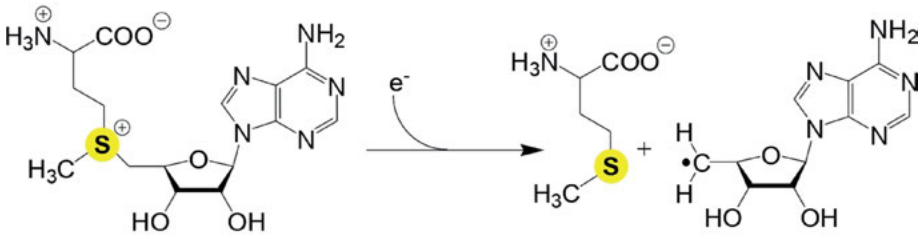


Figure 9. Reductive cleavage of *S*-adenosyl-L-methionine (SAM) to L-methionine and the 5′-deoxyadenosyl radical.

[199]. Decades of research on the biochemical and molecular roles of SAM in cellular metabolism provided an extensive foundation for its use in clinical studies, including those on depression, dementia, vacuolar myelopathy, liver disease, and osteoarthritis [197]. SAM has long been regarded as a source of methyl groups (CH_3) for DNA methylation, biosynthesis of hormones and neurotransmitters, and regulation of signal transduction (cell's *methyl iodide*). The methylation reaction of biomolecules can proceed readily based on the chemical reactivity of nucleophiles with the electrophilic sulfonium ion (Figure 9), and this property of SAM was regarded as the principal reason for its existence [46, 198].

Notably, a new role for SAM – according to biochemist Perry Frey "A Wolf in Sheep's Clothing, or a Rich Man's Adenosylcobalamin?" [198, 210] – evolved in parallel to the emergence of novel types of redox, regulatory, and enzymatic roles for iron-sulfur clusters, mediating two electron redox processes, coupling proton and electron transfer, and catalyzing disulfide reduction and reductive cleavage of SAM via sulfur-based cluster chemistry (Figure 10) [198, 208–213, 216–219].

Since then, an increasing number of radical SAM/Fe-S enzymes, e.g., lysine 2,3-aminomutase, pyruvate formate lyase activase, anaerobic ribonucleotide reductase, biotin synthase, lipoyl synthase, and benzylsuccinate synthase to name a few, were discovered. The bioinformatics study from 2001 [196] revealed that the superfamily of radical SAM enzymes contained over 600 members, many catalyzing reactions that were rich in novel chemical transformations. Meanwhile, the superfamily has grown immensely, and new details about the scope of reactions and biochemical pathways in which its members participate have emerged [204]. A recent analysis of the **S**tructure-**F**unction **L**inkage **D**atabase (SFLD) (http://sfld.rbvi.ucsf.edu/django/superfamily/29/) [220] suggests that there are more than 113,000 SAM radical enzymes [210]. SAM serves as the free radical initiator. Basically, it is cleaved to the 5′-deoxyadenosyl radical (Figure 9), which propagates radical formation by abstracting hydrogen atoms, either from substrate molecules to form radical intermediates, or from glycyl residues of enzymes to activate them [202, 203]. These diverse transformations involving SAM are a consequence of the diverse reactions that can be supported at the cofactor's sulfur center, as highlighted in the biosynthesis of biotin (vitamin B_7). This challenging reaction, which requires the insertion of one sulfur atom between two

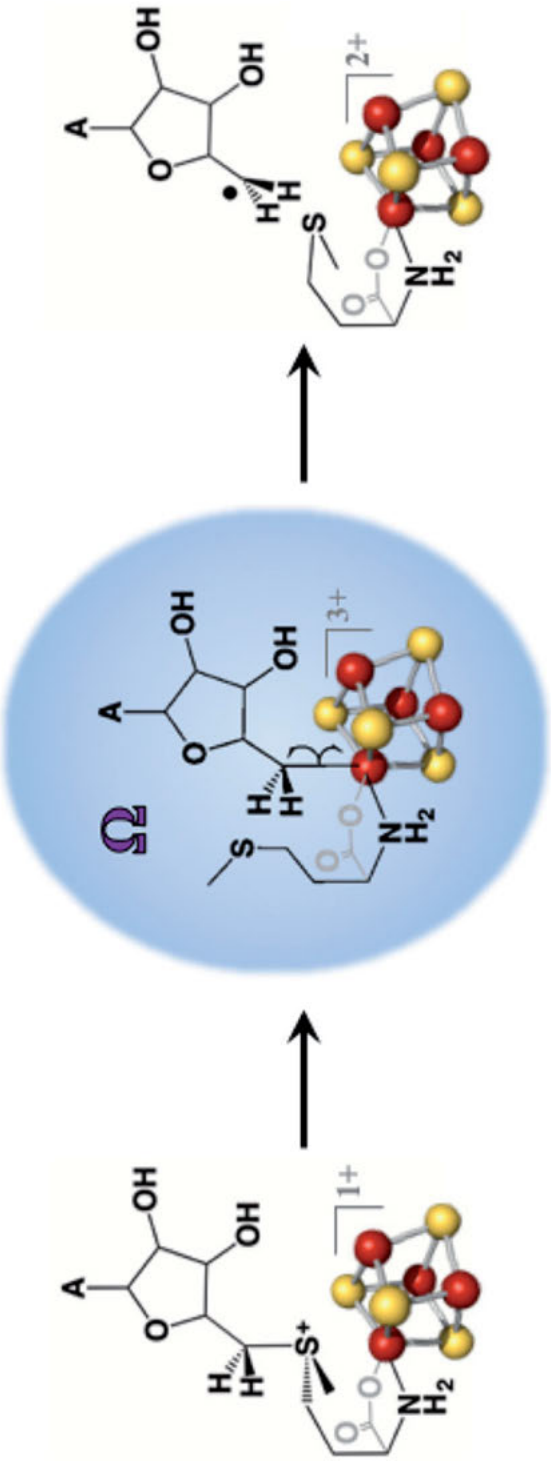


Figure 10. [4Fe-4S] cluster-assisted cleavage of *S*-adenosyl-L-methione (SAM). The amino and carboxylate groups of SAM bind to the unique iron of the catalytic Fe-S cluster, placing the sulfonium in close proximity to the cluster; the initiating enzymatic cleavage of SAM generates an organometallic intermediate Ω prior to liberation of the 5′-deoxyadenosyl radical, which initiates radical chemistry on the substrate. Adapted with permission from [211] (W. E. Broderick, B. M. Hoffman, J. B. Broderick, *Acc. Chem. Res.* **2018**, *51*, 2611–2619); copyright 2018, American Chemical Society.

non-activated carbon atoms of dethiobiotin, is catalyzed by sulfur transferase biotin synthase, a homodimer in solution, with one [2Fe-2S] and one [4Fe-4S] cluster, and SAM in each monomer (PDB ID 1R30) [221]. Lipoyl synthase (PDB ID 5EXK, 5EXJ) is another remarkable member of the radical SAM superfamily. In addition to SAM and the regular [4Fe-4S] cluster, this enzyme employs an auxiliary [4Fe-4S] cluster to catalyze the attachment of two sulfur atoms as the final step of lipoic acid synthesis [222–225]. Lipoic acid is an eight-carbon, straight-chain fatty acid containing sulfhydryl groups at atoms C6 and C8 which allow reversible formation of a cyclic disulfide. As a cofactor lipoic acid is required for the oxidative decarboxylation of α-keto acids. It is covalently attached to the enzyme by an amide bond to a terminal lysine residue, e.g., in the pyruvate dehydrogenase complex [226, 227].

In conclusion, radical SAM enzymes are powerful catalysts, in terms of the sheer number of spectacular transformations catalyzed, the radical SAM superfamily is hard to beat. A number of superb review articles recently written on the subject are recommended to acquire a deeper appreciation of this important class of enzymes [209–213].

5. OUTLOOK AND FUTURE DIRECTIONS

Although sulfur has been explored extensively since its identification as element by Lavoisier in 1777, sulfur and its compounds remain in the focus of active research in chemistry, biochemistry, geochemistry, microbiology, and medicine. Clearly, evolution made an excellent choice in selecting sulfur as one basic element of life in view of its astounding chemical properties. Notably, the field of geomicrobiology has experienced an extraordinary growth in recent years. Microorganisms have been studied in all kinds of extreme environments on Earth, ranging from crystalline rocks from the deep subsurface, ancient sedimentary rocks and hypersaline lakes, to dry deserts and deep-ocean hydrothermal vent systems [96]. For example, the anaerobic oxidation of the greenhouse gas methane, that exists in huge amounts in sea-floor sediments across the globe, to carbon dioxide by a consortium of methanotrophic archaea, has been shown to be coupled to the process of sulfate reduction to hydrogen sulfide, with elemental sulfur as a key intermediate [228–230]. Advances in oxygen and sulfur isotope biogeochemistry make significant contributions to our understanding of these complicated processes linking the biogeochemical cycles of the elements, e.g., carbon and sulfur [231–236].

New enzymes have been discovered, functionally and structurally characterized, such as the membrane-bound sulfane reductase from thermophilic *Pyrococcus furiosus* [237], the formyl-methanofuran dehydrogenase from *Methanothermobacter wolfeii*, a key enzyme of methanogenesis carrying an electron supplying core with a record high of 46 (!) electronically coupled [4Fe-4S] clusters [238], or the human DNA primase, where a [4Fe-4S] cluster appears to act as an on/off redox switch for DNA binding [239]. Not too surprising, a new catalytic activity by a radical SAM enzyme, the reductive cleavage of sulfoxides and sulfones, has

been found most recently [240], and in diphthamide biosynthesis an usual cleavage of the SAM C–S bond leading to the capture of the 3-amino-3-carboxypropyl radical, and not to the 5′-deoxyadenosyl radical, could be detected [241, 242].

The investigation of the biosynthesis and physiological role of persulfides, such as cysteine hydropersulfide (CysSSH) or glutathione persulfide (GSSH), has attracted lots of interest by researchers in recent years [50]. These sulfur compounds are found in prokaryotes, eukaryotic cells, and mammalian tissues. The chemical properties and abundance of these species suggest a pivotal role for compounds containing an -SSH group in cell-regulatory processes, persulfides can function as potent antioxidants and cellular protectants, and as redox signaling intermediates [243]. Furthermore, the important life-supporting role of hydrogen sulfide (H_2S), which evolved from microorganisms to plants, invertebrates, vertebrates, and finally to mammals, has to be mentioned. For a long time, H_2S was only known for its toxicity and environmental hazard, the appreciation of its physiology started by the discovery of endogenous H_2S production in mammalian cells and gained momentum by typifying this gasotransmitter (next to NO and CO) with a variety of important functions in cardiovascular, neuronal, immune, renal, respiratory, gastrointestinal, reproductive, liver, and endocrine systems [244–246].

As sulfur is quite abundant in nature, nontoxic, environmentally benign, there is an increasing interest in industrial applications. For example, sulfur has a ultrahigh theoretical capacity and theoretical energy density, which distinguishes it from other cathode materials. Consequently, the area of lithium-sulfur batteries gained significant momentum over the past few decades and substantial progress could be made. However, there still exist quite a few scientific and technical problems to be solved which currently prevent lithium-sulfur batteries to become commercialized [247, 248]. In addition, there exist exciting applications of sulfur radicals in organic synthesis, polymer chemistry, and in materials science, as most recently reviewed [249].

Last but not least, as there is a rich chemistry and biochemistry for selenium, like oxygen and sulfur a member of group 16 of the Periodic Table, the excellent review by Reich and Hondal entitled “Why Nature Chose Selenium” must be mentioned here. It describes important discoveries of the biological processes that selenium participates in, and a point-by-point comparison of the chemistry of selenium with the atom it replaces in biology, sulfur. The in-depth analysis reveals that redox chemistry is the largest chemical difference between Se and S. This difference is very large for both one-electron and two-electron redox reactions, and much of this difference is due to the inability of selenium to form π bonds of all types [250].

ACKNOWLEDGMENTS

The authors are grateful for continuous financial support by DGAPA-UNAM, Facultad de Química-UNAM and CONACYT (MEST, ARM, ASP), and by the Deutsche Forschungsgemeinschaft and the Universität Konstanz (PK).

ABBREVIATIONS AND DEFINITIONS

AMP	adenosine 5′-monophosphate
APS	adenosine 5′-phosphosulfate
ATP	adenosine 5′-triphosphate
Cys	cysteine
Dsr	dissimilatory sulfite reductase
EPR	electron paramagnetic resonance
FAD	flavin adenine dinucleotide
His	histidine
Met	methionine
MGD	molybdopterin guanine dinucleotide
NAD(P)H	nicotinamide adenine dinucleotide (phosphate)
PP_i	pyrophosphate
Psr	polysulfide reductase
SAM	*S*-adenosyl-L-methionine
SFLD	Structure-Function Linkage Database
TEED	Thiamine diphosphate-dependent Enzyme Engineering Database
ThDP	thiamine diphosphate
XAS	X-ray absorption spectroscopy

REFERENCES

1. Editorial, *Nature Chemical Biology* 2006, **2**, 169; doi.org/10.1038/nchembio0406-169.
2. S. M. Sievert, R. P. Kiene, H. Schulz-Vogt, *Oceanography* **2007**, *20*, 117–123.
3. N. N. Greenwood, A. Earnshaw, *Chemistry of the Elements*, 2nd ed., Butterworth-Heinemann, Oxford, 1997.
4. B. Meyer, *Chem. Rev.* **1976**, *76*, 367–388.
5. J. Boulegue, *Phosphorus and Sulfur and the Related Elements* **1978**, *5*, 127–128, doi: 10.1080/03086647808069875.
6. R. Steudel, *Angew. Chem.* **1975**, *14*, 655–702.
7. D. A. Boyd, *Angew. Chem. Int. Ed.* **2016**, *55*, 15486–15502.
8. R. T. Raines, *Nature Struct. Biol.* **1997**, *4*, 425–427.
9. K. Rose, S. Shadle, T. Glaser, S. deVries, A. Cherepanov, G. W. Canters, B. Hedman, K. O. Hodgson, E. I. Solomon *J. Am. Chem. Soc.* **1999**, *121*, 2353–23639.
10. M. S. Schmøkel, L. Bjerg, S. Cenedese, M. R. V. Jørgensen, Y.-S. Chen, J. Overgaard, B. B. Iversen, *Chem. Sci.* **2014**, *5*, 1408–1421.
11. A. D. Rogers, P. A. Tyler, D. P. Connelly, J. T. Copley, R. James, R. D. Larter, K. Linse, R. A. Mills, A. Naveira Garabato, R. D. Pancost, D. A. Pearce, N. V. C. Polunin, C. R. German, T. Shank, P. H. Boersch-Supan, B. J. Alker, A. Aquilina, S. A. Bennett, A. Clarke, R. J. J. Dinley, A. G. C. Graham, D. R. H. Green, J. A. Hawkes, L. Hepburn, A. Hilario, V. A. I. Huvenne, L. Marsh, E. Ramirez-Llodra, W. D. K. Reid, C. N. Roterman, C. J. Sweeting, S. Thatje, K. Zwirglmaier, *PLoS Biol.* **2012**, *10*(1): e1001234.
12. K. Nakamura, H. Watanabe, J. Miyazaki, K. Takai, S. Kawagucci, T. Noguchi, S. Nemoto, T. Watsuji, T. Matsuzaki, T. Shibuya, K. Okamura, M. Mochizuki, Y. Orihashi, T. Ura, A. Asada, D. Marie, M. Koonjul, M. Singh, G. Beedessee, M. Bhikajee, K. Tamaki, *PLoS ONE* **2012**, *7*, e32965. doi:10.1371/journal.pone.0032965.

13. H. Yao, M. Dao, T. Imholt, J. Huang, K. Wheeler, A. Bonilla, S. Suresh, C. Ortiz, *Proc. Natl. Acad. Sci. USA* **2010**, *107*, 987–992.
14. A. Warén, S. Bengtson, S. K. Goffredi, C. L. Van Dover, *Science*, **2003**, *302*, 1007.
15. C. Chen, K. Linse, J. T. Copley, A. D. Rogers, *J. Molluscan Studies*, **2015**, *81*, 322–334. doi:10.1093/mollus/eyv013.
16. S. K. Goffredi, A. Warén, V. J. Orphan, C. L. Van Dover, R. C. Vrijenhoek, *Appl. Environ. Microbiol.* **2004**, *70*, 3082–3090.
17. Y. Suzuki, R. E. Kopp, T. Kogure, A. Suga, K. Takai, S. Tsuchida, N. Ozaki, K. Endo, J. Hashimoto, Y. Kato, C. Mizota, T. Hirata, H. Chiba, K. H. Nealson, K. Horikoshi, J. L. Kirschvink, *Earth Planet. Sci. Lett.* **2006**, *242*, 39–50.
18. D. M. Gardener, G. K. Fraenkel, *J. Am. Chem. Soc.* **1955**, *77*, 6399–6400.
19. C. J. Allegre, Y. Bottinga, *Nature* **1974**, *252*, 31–32.
20. D. Reinen, G.-G. Lindner, *Chem. Soc. Rev.*, **1999**, *28*, 75–84.
21. G. R. Eaton, S. S. Eaton, J. W. Stoner, R. W. Quine, G. A. Rinard, A. I. Smirnov, R. T. Weber, J. Krzysteks, A. K. Hassans, L.-C. Brunei, A. Demortieró, *Appl. Magn. Reson.* **2001**, *21*, 563–570.
22. N. Gobeltz-Hautecoeur, A. Demortier, B. Lede, J. P. Lelieur, C. Duhayon, *Inorg. Chem.* **2002**, *41*, 2848–2854.
23. R. Steudel, *Top. Curr. Chem.* **2003**, *231*, 127–152.
24. D. Arieli, D. E. W. Vaughan, D. Goldfarb, *J. Am. Chem. Soc.* **2004**, *126*, 5776–5788.
25. R. Linguerri, N. Komiha, J. Fabian, P. Rosmus, *Z. Phys. Chem.* **2008**, *222*, 163–176.
26. E. Boros, M. J. Earle, M. A. Gîlea, A. Metlen, A.-V. Mudring, F. Rieger, A. J. Robertson, K. R. Seddon, A. A. Tomaszowska, L. Trusov, J. S. Vy, *Chem. Commun.* **2010**, *46*, 716–718.
27. M. E. Fleet, X. Liu, *Spectrochim. Acta B* **2010**, *65*, 75–79.
28. K. Raulin, N. Gobeltz, H. Vezin, N. Touati, B. Ledé, A. Moissette, *Phys. Chem. Chem. Phys.* **2011**, *13*, 9253–9259.
29. *Illustrated Guide to the Egyptian Museum in Cairo*, Eds A. Bongioanni, M. Sole Croce, **2001**.
30. M. Ganio, E. S. Pouyet, S. M. Webb, C. M. Schmidt Patterson, M. S. Walton, *Pure Appl. Chem.* **2018**, *90*, 463–475.
31. C. E. Manning, *Science* **2011**, *331*, 1018–1019.
32. G. S. Pokrovski, L. Dubrovinsky, *Science* **2011**, *331*, 1052–1054.
33. G. S. Pokrovski, M. A. Kokh, D. Guillaume, A. Y. Borisova, P. Gisquet, J.-L. Hazemann, E. Lahera, W. Del Net, O. Proux, D. Testemale, V. Haigis, R. Jonchière, A. P. Seitsonen, G. Ferlat, R. Vuilleumier, A. M. Saitta, M.-C. Boiron, J. Dubessy, *Proc. Natl. Acad. Sci. USA* **2015**, *212*, 13484–13489.
34. H. Beinert, *Eur. J. Biochem.* **2000**, *267*, 5657–5664.
35. C. N. R. Rao, K. P. R. Pisharody, *Progr. Solid State Chem.* **1976**, *10*, 207–270.
36. G. Wächtershäuser, *Microbiol. Rev.* **1988**, *52*, 452–484.
37. J. I. Toohey, *Biochem. J.* **1989**, *264*, 625–632.
38. M. Rakowski Dubois, *Chem. Rev.* **1989**, *89*, 1–9.
39. R. Schauder, A. Kröger, *Arch. Microbiol.* **1993**, 159, 491–497.
40. E. A. Bonch-Osmolovskaya, *FEMS Microbiol. Rev.* **1994**, *15*, 65–77.
41. R. H. Holm, P. Kennepohl, E. I. Solomon, *Chem. Rev.* **1996**, *96*, 2239–2314.
42. H. Beinert, R. H. Holm, E. Münck, *Science* **1997**, *277*, 653–659.
43. E. I. Stiefel, *Pure Appl. Chem.* **1998**, *70*, 889–896.
44. R. Hedderich, O. Klimmek, A. Kröger, R. Dirmeier, M. Keller, K. O. Stetter, *FEMS Microbiol. Rev.* **1999**, *22*, 353–381.
45. J. Cheek, J. B. Broderick, *J. Biol. Inorg. Chem.* **2001**, *6*, 209–226.
46. P. A. Frey, *Annu. Rev. Biochem.* **2001**, *70*, 121–148.

47. M. Fontecave, S. Ollagnier-de-Choudens, E. Mulliez, *Chem. Rev.* **2003**, *103*, 2149–2166.
48. D. C. Rees, J. B. Howard, *Science* **2003**, *300*, 929–931.
49. G. Wächtershäuser, *Phil. Trans. R. Soc. B* **2006**, *361*, 1787–1808.
50. E. G. Mueller, *Nature Chem. Biol.* **2006**, *2*, 185–194.
51. M. Fontecave, *Nature Chem. Biol.* **2006**, *2*, 171–174.
52. D. Rickard, G. W. Luther, *Rev. Mineral. Geochem.* **2006**, *61*, 421–504.
53. D. Rickard, G. W. Luther, *Chem. Rev.* **2007**, *107*, 514–562.
54. N. Sahai, H. Kaddour, P. Dalai, *Elements* **2016**, 12, 389–395.
55. C. M. Jäger, A. K. Croft, *Chem. Bio. Eng. Rev.* **2018**, *5*, 143–162.
56. S. M. Sievert, R. P. Kiene, H. Schulz-Vogt, *Oceanography* **2007**, *20*, 117–1236.
57. J. Simon, P. M. H. Kroneck, *Adv. Microb. Physiol.* **2013**, *62*, 49–117.
58. K. Parey, G. Fritz, U. Ermler, P. M. H. Kroneck, *Metallomics* **2013**, *5*, 302–317.
59. L. L. Barton, M.-L. Fardeau, G. D. Fauque, *Met. Ions Life Sci.* **2014**, *14*, 237–277.
60. W. E. Martin, F. L. Sousa, N. Lane, *Science* **2014**, *344*, 1092–1093.
61. A. P. Nutman, V. C. Bennett, C. R. L. Friend, M. J. Van Kranendonk, A. R. Chivas, *Nature* **2016**, *537*, 535–538.
62. K. Wasmund, M. Mußmann, A. Loy, *Environm. Microbiol. Rep.* **2017**, *9*, 323–344.
63. K. Anantharaman, B. Hausmann, S. P. Jungbluth, R. S. Kantor, A. Lavy, L. A. Warren, M. S. Rappé, M. Pester, A. Loy, B. C. Thomas, J. F. Banfield, *The ISME Journal* **2018**, 12, 1715–1728.
64. www.bbc.com/earth/story/20161026-the-secret-of-how-life-on-earth-began. Accessed May 9, 2019.
65. M. J. Russell, J. Hall, *J. Geol. Soc. London* **1997**, *154*, 377–402.
66. J. L. Bada, A. Lazcano, *Science* **2002**, *296*, 1982–1983.
67. J. L. Bada, A. Lazcano, *Science* **2003**, *300*, 745–746.
68. W. Martin, M. J. Russell, *Phil. Trans. R. Soc. London B* **2003**, *358*, 59–85.
69. W. Martin, M. J. Russell, *Phil. Trans. R. Soc. London B* **2007**, *362*, 1887–1925.
70. J. L. Bada, B. Fegley, Jr., S. L. Miller, A. Lazcano, H. J. Cleaves, R. M. Hazen, J. Chalmers, *Science* **2007**, *315*, 937–938.
71. G. Wächtershäuser, C. Huber, *Science* **2007**, *315*, 938–939.
72. J. P. Schrum, T. F. Zhu, J. W. Szostak, *Cold Spring Harb. Perspect. Biol.* **2010**, doi: 10.1101/cshperspect.a002212.
73. M. Müller, M. Mentel, J. J. van Hellemond, K. Henze, C. Woehle, S. B. Gould, R.-Y. Yu, M. van der Giezen, A. G. M. Tielens, W. F. Martin, *Microbiol. Mol. Biol. Rev.* **2012**, *76*, 444–495.
74. F. L. Sousa, T. Thiergart, G. Landan, S. Nelson-Sathi, I. A. C. Pereira, J. F. Allen, N. Lane, W. F. Martin, *Phil. Trans. R. Soc. London B* **2013**, *368*: 20130088.
75. J. D. Kim, S. Senn, A. Harel, B. I. Jelen, P. G. Falkowski, *Phil. Trans. R. Soc. London B* **2013**, *368*: 20120257; doi: org/10.1098/rstb.2012.0257.
76. M.-P. Bassez, *Orig. Life Evol. Biosph.* **2017** *47*, 453–480, doi: 10.1007/s11084–017–9534–5.
77. N. Kitadai, S. Maruyama, *Geoscience Frontiers* **2018**, *9*, 1117–1153.
78. G. Wächtershäuser, *Origin of Life: RNA World versus Autocatalytic Anabolism*, in Vol. 1 of *The Prokaryotes*, Eds M. Dworkin, S. Falkow, E. Rosenberg, K. H. Schleifer, E. Stackebrandt, Springer, New York, NY, 2006, pp. 275–283.
79. P. Taylor, T. E. Rummery, D. G. Owen, *J. Inorg. Nucl. Chem.* **1979**, *41*, 1683–1687.
80. G. Wächtershäuser, *System. Appl. Microbiol.* **1988**, *10*, 207–210.
81. E. Drobner, H. Huber, G. Wächtershäuser, D. Rose, K. O. Stetter, *Nature* **1990**, *346*, 742–744.
82. H. Beinert, *J. Biol. Inorg. Chem.* **2000**, *5*, 2–15.
83. D. C. Rees, J. B. Howard, *Science* **2003**, *300*, 929–931.

84. W. Nitschke, S. E. McGlynn, E. J. Milner-White, M. J. Russell, *Biochim. Biophys. Acta* **2013**, *1827*, 871–881.
85. A. Roldan, N. Hollingsworth, A. Roffey, H.-U. Islam, J. B. M. Goodall, C. R. A. Catlow, J. A. Darr, W. Bras, G. Sankar, K. B. Holt, G. Hogarth, N. H. de Leeuw, *Chem. Commun.* **2015**, *51*, 7501–7504.
86. Y. Han, G. Gonnella, N. Adam, A. Schippers, L. Burkhardt, S. Kurtz, U. Schwarz-Schampera, H. Franke, M. Perner, *Sci. Rep.* **2018**, *8*, 10386. doi: 10.1038/s41598–018-28613–5.
87. W. Heinen, A. M. Lauwers, *Origins of Life and Evolution of the Biosphere* **1996**, *26*, 131–150.
88. M. J. Russell, W. Martin, *Trends Biochem. Sci.* **2004**, *29*, 358–363.
89. R. M. Hazen, *Elements*, **2005**, *1*, 135–137.
90. N. Sahai, H. Kaddour, P. Dalai, *Elements* **2016**, *12*, 389–394.
91. P. Philippot, M. VanZuilen, K. Lepot, C. Thomazo, J. Farquhar, M. J. Van Kranendonk, *Science* **2007**, *317*, 1534–1537.
92. D. T. Johnston, *Earth-Science Rev.* **2011**, *106*, 161–183.
93. A. S. Bradley, W. D. Leavitt, D. T. Johnston, *Geobiology* **2011**, *9*, 446–457.
94. L. R. Kump, *Elements* **2012**, *8*, 410–411.
95. R. Tostevin, A. V. Turchyn, J. Farquhar, D. T. Johnston, D. L. Eldridge, J. K. B. Bishop, M. McIlvin, *Earth Planet. Sci. Lett.* **2014**, *396*, 14–21.
96. H. Dong, B. Yu, *Episodes* **2007**, *30*, 202–216.
97. L. H. Gregersen, D. A. Bryant, N.-U. Frigaard, *Front. Microbiol.* **2011**, *2*, article 116, 1–14, doi: 10.3389/fmicb.2011.00116.
98. Y. Liu, L. L. Beer, W. B. Whitman, *Environ. Microbiol*. **2012**, *14*, 2632–2644.
99. D. E. Rawlings, D. B. Johnson, *Microbiology* **2007**, *153*, 315–324.
100. M. Vera, A. Schippers, W. Sand, *Appl. Microbiol. Biotechnol.* **2013**, *97*, 7529–7541.
101. C. Janosch, F. Remonsellez, W. Sand, M. Vera, *Microorganisms* **2015**, *3*, 707–724.
102. T. Urich, C. M. Gomes, A. Kletzin, C. Frazão, *Science* **2006**, *311*, 996–1000.
103. A. Veith, T. Urich, K. Seyfarth, J. Protze, C. Frazão, A. Kletzin, *Front. Microbiol.* **2011**, Mar 7; 2:37. doi: 10.3389/fmicb.2011.00037.
104. F. Widdel, T. A. Hansen, *The Dissimilatory Sulfate- and Sulfur-Reducing Bacteria*, in Vol. 1 of *The Prokaryotes*, 2nd ed., Eds A. Balows, H. Trüper, M. Dworkin, W. Harder, K. H. Schleifer, Springer-Verlag, New York, Berlin, Heidelberg, 1992, pp. 583–624.
105. N. Pfennig, H. Biebl, *Arch. Microbiol.* **1976**, *110*, 3–12.
106. H. Biebl, N. Pfenning, *Arch. Microbiol.* **1977**, *112*, 115–117.
107. R. K. Thauer, W. Jungermann, K. Decker, *Bact. Rev.* **1977**, *41*, 100–180.
108. N. A. Gebhardt, R. K. Thauer, D. Linder, P. M. Kaulfers, N. Pfennig, *Arch. Microbiol.* **1985**, *141*, 392–398.
109. J. Paulsen, A. Kröger, R. K. Thauer, *Arch. Microbiol.* **1986**, *144*, 78–83.
110. J. M. Macy, I. Schröder, R. K. Thauer, A. Kröger, *Arch. Microbiol.* **1986**, *144*, 147–150.
111. K. O. Stetter, G. Fiala, G. Huber, R. Huber, A. Segerer, *FEMS Microbiol. Rev.* **1990**, *75*, 117–124.
112. W. Schumacher, P. M. H. Kroneck, N. Pfennig, *Arch. Microbiol.* **1992**, *158*, 287–293.
113. R. Bache, P. M. H. Kroneck, H. Merkle, H. Beinert, *Biochim. Biophys. Acta* **1983**, *722*, 417–426.
114. A. Zöphel, M. C. Kennedy, H. Beinert, P. M. H. Kroneck, *Eur. J. Biochem.* **1991**, 195, 849–856.
115. K. Ma, R. N. Schicho, R. M. Kelly, M. W. W. Adams, *Proc. Natl. Acad. Sci. USA* **1993**, *90*, 5341–5344.

116. T. D. Brock, *Thermophilic Microorganisms and Life at High Temperatures*. Springer-Verlag, Berlin, Heidelberg, New York, 1978.
117. A. Zöphel, M. C. Kennedy, H. Beinert, P. M. H. Kroneck, *Arch. Microbiol.* **1988**, *150*, 72–77.
118. G. D. Fauque, *Meth. Enzymol.* **1994**, *243*, 353–367.
119. R. Schauder, E. Müller, *Arch. Microbiol.* **1993**, *160*, 377–382.
120. G. D. Fauque, O. Klimmek, A. Kröger, *Meth. Enzymol.* **1994**, *243*, 367–383.
121. M. Jormakka, K. Yokoyama, T. Yano, M. Tamakoshi, S. Akimoto, T. Shimamura, P. Curmi, S. Iwata, *Nature Struct. Mol. Biol.* **2008**, *15*, 730–737.
122. S. Herwald, A. Y. Liu, B. E. Zhu, K. W. Sea, K. M. Lopez, M. H. Sazinsky, E. J. Crane, *Biochemistry* **2013**, *52*, 2764–2773.
123. M. Schidlowski, J. M. Hayes, I. R. Kaplan, in *Earth' Earliest Biosphere, Its Origin and Evolution*, Ed. J. W. Schopf, Princeton University Press, Princeton, USA, 1983, pp. 149–186.
124. E. M. Cameron, *Nature* **1982**, *296*, 145–148.
125. M. W. W. Adams, *Annu. Rev. Microbiol.* **1993**, *47*, 627–658.
126. P. M. Matias, I. A. Pereira, C. M. Soares, M. A. Carrondo, *Prog. Biophys. Mol.Biol.* **2005**, *89*, 292–329.
127. R. Rabus, T. A. Hansen, F. Widdel, *Dissimilatory Sulfate- and Sulfur-Reducing Prokaryotes*, in Vol. 2 of *The Prokaryotes*, Eds M. Dworkin, S. Falkow, E. Rosenberg, K. H. Schleifer, E. Stackebrandt, Springer, New York, NY, 2006, pp. 659–768.
128. M. G. Klotz, D. A. Bryant, T. E. Hanson, *Front. Microbiol.* **2011**, Dec 2; 2:241. doi: 10.3389/fmicb.2011.00241.
129. D. Wacey, M. R. Kilburn, M. Saunders, J. Cliff, M. D. Brasier, *Nature Geosci.* **2011**, *4*, 698–702.
130. M. S. Sim, D. T. Wang, G. M. Zane, J. D. Wall, T. Bosak, S. Ono, *Front. Microbiol.* **2013**, June 25; 4:171. doi: 10.3389/fmicb.2013.00171.
131. R. Rabus, S. S. Venceslau, L. Wöhlbrand, G. Voordouw, Judy D. Wall, I. A. C. Pereira, *Adv. Microb. Physiol.* **2015**, *66*, 55–321.
132. G. Fritz, P. M. H. Kroneck, *Science* **2015**, *350*, 1476–1477.
133. A. A. Santos, S. S. Venceslau, F. Grein, W. D. Leavitt, C. Dahl, D. T. Johnston, I. A. C. Pereira, *Science* **2015**, *350*, 1541–1545.
134. A. G. Duarte, T. Catarino, G. F. White, D. Lousa, S. Neukirchen, C. M. Soares, F. L.Sousa, T. A. Clarke, I. A. C. Pereira, *Nature Commun.* **2018**, *9*, Article number: 5448. doi: 10.1038/s41467-018-07839-x.
135. J.-H. Jeoung, J. Fesseler, S. Goetzl, H. Dobbek, *Met. Ions Life Sci.* **2014**, 14, 37–69.
136. V. C.-C. Wang, S. W. Ragsdale, F. A. Armstrong, *Met. Ions Life Sci.* **2014**, 14, 71–97.
137. A. Parkin, *Met. Ions Life Sci*. **2014**, *14*, 99–124.
138. C. C. Lee, M. W. Ribbe, Y. Hu, *Met. Ions Life Sci*. **2014**, *14*, 147–176.
139. L. K. Schneider, A. Wüst, A. Pomowski, L. Zhang, O. Einsle, *Met. Ions Life Sci*. **2014**, *14*, 177–210.
140. S. C. Lee, W. Lo, R. H. Holm, *Chem. Rev.* **2014**, *114*, 3579–3600.
141. E. I. Stiefel, *Transition Metal Sulfur Chemistry: Biological and Industrial Significance and Key Trends*, in *Transition Metal Sulfur Chemistry*, Eds E. I. Stiefel, K. Matsumoto, ACS Symposium Series, Vol. 653, American Chemical Society, Washington, DC, 1996, pp. 2–38.
142. G. B. Kauffman, *Alfred Werner, Founder of Coordination Chemistry*, Springer Verlag, New York, 1966.
143. R. J. P. Williams, *Chem. Commun.* **2003**, 1109–1113.
144. H. Beinert, *J. Biol. Chem.* **2002**, *277*, 37967–37972.
145. B. G. Malmström, *Cytochrome c Oxidase and Related Enzymes*, in *Biological Oxidations*, Eds H. Sund, V. Ullrich, Springer, Berlin, Heidelberg, 1983, pp. 189–200.

146. E. I. Solomon, R. H. Holm, *Chem. Rev.* **2004**, *104*, 347–348.
147. Y. Lu, *Angew. Chem. Int. Ed.* **2006**, *45*, 5588–5601.
148. S. Groysman, R. H. Holm, *Biochemistry* **2009**, *48*, 2310–2320.
149. E. I. Solomon, L. B. LaCroix, D. W. Randall, *Pure Appl. Chem.*, **1998**, *70*, 799–808.
150. Y. Xue, A. V. Davis, G. Balakrishnan, J. P. Stasser, B. M. Staehlin, P. Focia, T. G. Spiro, J. E. Penner-Hahn, T. V. O'Halloran, *Nature Chem. Biol.* **2008**, *4*, 107–109.
151. I. Gentile, H. A. Garro, S. Delgado Ocaña, N. Gonzalez, T. Strohäker, D. Schibich, L. V. Quintanar, L. Sambrotta, M. Zweckstetter, C. Griesinger, M. Menacho Márqueza, C. O. Fernández, *Metallomics* **2018**, *10*, 1383–1389.
152. J. Meyer, *J. Biol. Inorg. Chem.* **2008**, *13*, 157–170.
153. E. C. Wittenborn, M. Merrouch, C. Ueda, L. Fradale, C. Léger, V. Fourmond, M.-E. Pandelia, S. Dementin, C. L. Drennan, *eLife* **2018**, *7*:e39451. doi: 10.7554/eLife.39451.
154. L. E. Mortenson, R. C. Valentine, J. E. Carnahan, *Biochem. Biophys. Res. Commun.*, **1962**, *7*, 448–452.
155. D. C. Blomstrom, E. Knight, Jr., W. D. Phillips, J. F. Weiher, *Proc. Natl. Acad. Sci. USA* **1964**, *51*, 1085–1092.
156. W. Lovenberg, K. McCarthy, *Biochem. Biophys. Res. Commun.* **1968**, *30*, 453–458.
157. J. Liu, S. Chakraborty, P. Hosseinzadeh, Y. Yu, S. Tian, I. Petrik, A. Bhagi, Y. Lu, *Chem. Rev.* **2014**, *114*, 4366–4469.
158. H. Jung Hwang, S. M. Berry, M. J. Nilges, Y. Lu, *J. Am. Chem. Soc.* **2005**, *127*, 7274–7275.
159. H. J. Wijma, I. MacPherson, O. Farver, E. I. Tocheva, I. Pecht, M. Ph. Verbeet, M. E. P. Murphy, G. W. Canters, *J. Am. Chem. Soc.* **2007**, *129*, 519–525.
160. J. Kang, H. Kino, M. Tateno, *Biochim. Biophys. Acta* **2011**, *1807*, 1314–1327.
161. T. Kroll, R. G. Hadt, S. A. Wilson, M. Lundberg, J. J. Yan, T.-C. Weng, D. Sokaras, R. Alonso-Mori, D. Casa, M. H. Upton, B. Hedman, K. O. Hodgson, E. I. Solomon, *J. Am. Chem. Soc.* **2014**, *136*, 18087–18099.
162. A. L. Le Sueur, R. N. Schaugaard, M.-H. Baik, M. C. Thielges, *J. Am. Chem. Soc.* **2016**, *138*, 7187–7193.
163. M. O. Ross, O. S. Fisher, M. N. Morgada, M. D. Krzyaniak, M. R. Wasielewski, A. J. Vila, B. M. Hoffman, A. C. Rosenzweig, *J. Am. Chem. Soc.* **2019**, *141*, 4678–4686.
164. J. D. Finkelstein, J. J. Martin, *Int. J. Biochem. Cell Biol.* **2000**, *32*, 385–389.
165. D. Sippel, M. Rohde, J. Netzer, C. Trncik, J. Gies, K. Grunau, I. Djurdjevic, L. Decamps, S. L. A. Andrade, O. Einsle, *Science* **2018**, *359*, 1484–1489.
166. S. C. Peck, K. Denger, A. Burrichter, S. M. Irwin, E. P. Balskus, D. Schleheck, *Proc. Natl. Acad. Sci. USA* **2019**, *116*, 3171–3176.
167. D. Nelson, M. Cox, *Lehninger Principles of Biochemistry*, 4th ed., W. H. Freeman and Company, New York, 2006.
168. U. Ermler, W. Grabarse, S. Shima, M. Goubeaud, R. K. Thauer, *Science* **1997**, *278*, 1457–1462.
169. R. K. Thauer, *Microbiology* **1998**, *144*, 2377–2406.
170. T. J. Lawton, A. C. Rosenzweig, *Science* **2016**, *352*, 892–893.
171. T. Wongnate, D. Sliwa, B. Ginovska, D. Smith, M. W. Wolf, N. Lehnert, S. Raugei, S. W. Ragsdale, *Science* **2016**, *352*, 953–958.
172. T. Wagner, J. Koch, U. Ermler, S. Shima, *Science* **2017**, *357*, 699–703.
173. R. K. Thauer, *Biochemistry* **2019**, *58*, 4269–4271.
174. D. Kern, G. Kern, H. Neef, K. Tittmann, M. Killenberg-Jabs, C. Wikner, G. Schneider, G. Hübner, *Science* **1997**, *275*, 67–70.
175. B. Shaanan, D. M. Chipman, *FEBS J.* **2009**, *276*, 2447–2453.
176. D. Meyer, P. Neumann, R. Ficner, K. Tittmann, *Nature Chem. Biol.* **2013**, *9*, 488–491.
177. R. Breslow, *J. Am. Chem. Soc.* **1958**, *80*, 3719–3726.

178. L. Bettendorff, F. Mastrogiacomo, J. Kish, T. Grisar, *J. Neurochem.* **1996**, *66*, 250–258.
179. A. Schellenberger, *Biochim. Biophys. Acta* **1998**, *1385*, 177–186.
180. G. Hübner, K. Tittmann, M. Killenberg-Jabs, J. Schäffner, M. Spinka, H. Neef, D. Kern , G. Kern, G. Schneider, C. Wikner, S. Ghisla, *Biochim. Biophys. Acta* **1998**, *1385*, 221–228.
181. F. J. Leeper, D. Hawksley, S. Mann, C. Perez Melero, M. D. H. Wood, *Biochem. Soc. Trans.* **2005**, *33*, 772–775.
182. R. G. Duggleby, *Acc. Chem. Res.* **2006**, *39*, 550–557.
183. R. A. W. Frank, F. J. Leeper, B. F. Luisi, *Cell. Mol. Life Sci.* **2007**, *64*, 892–905
184. R. Kluger, K. Tittmann, *Chem. Rev.* **2008**, *108*, 1797–1833.
185. Y. Liu, Y. Li, X. Wang, *Appl. Microbiol. Biotechnol.* **2016**, *100*, 8633–8649.
186. E. N. M. Jirgis, G. Bashiri, E. M. M. Bulloch, J. M. Johnston, E. N. Baker, *Structure* **2016**, *24*, 1167–1177.
187. U. Schörken, G. A. Sprenger, *Biochim. Biophys. Acta* **1998**, *1385*, 229–243.
188. M. Müller, D. Gocke, M. Pohl, *FEBS J.* **2009**, *276*, 2894–2904.
189. M. Müller, G. A. Sprenger, M. Pohl, *Curr. Opin. Chem. Biol.* **2013**, *17*, 261–270.
190. P. P. Giovannini, O. Bortolini, A. Massi, *Eur. J. Org. Chem.* **2016**, 4441–4459.
191. M. Widmann, R. Radloff, J. Pleiss, *BMC Biochemistry* **2010**, 11:9 doi: 10.1186/1471–2
192. F. Jordan, N. S. Nemeria, *Bioorg. Chem.*, **2014**, *57*, 251–262.
193. P. Hemmerich, H. Michel, C. Schug, V. Massey, *Structure and Bonding* **1982**, *48*, 93–124.
194. H. Beinert, *J. Am. Chem. Soc.* **1956**, *78*, 5323–5328.
195. J. A. Stubbe, W. A. van der Donk, *Chem. Rev.* **1998**, *98*, 705–762.
196. H. J. Sofia, G. Chen, B. G. Hetzler, J. F. Reyes-Spindola, N. E. Miller, *Nucleic Acids Res.* **2001**, *29*, 1097–1106.
197. T. Bottiglieri, *Am. J. Clin. Nutr.* **2002**, *76*, 1151S–1157S.
198. P. A. Frey, O. T. Magnusson, *Chem. Rev.* **2003**, *103*, 2129–2148.
199. S. Roje, *Phytochemistry* **2006**, *67*, 1686–1698.
200. W. A. M. Loenen, *Biochem. Soc. Trans.* **2006**, *34*, 330–333.
201. G. Layer, D. W Heinz, D. Jahn, W.-D. Schubert, *Curr. Op. Chem. Biol.* **2004**, *8*, 468–476.
202. S. C. Wang, P. A. Frey, *Biochemistry* **2007**, *46*, 12889–12895.
203. S. C. Wang, Perry A. Frey, *Trends Biochem. Sci.* **2007**, *32*, 101–110.
204. P. A. Frey, A. D. Hegeman, F. J. Ruzicka, *Crit. Rev. Biochem. Mol. Biol.* **2008**, *43*, 63–88.
205. S. J. Booker, T. L Grove, *F1000 Biol. Reports* **2010**, *2*:52; doi: 10.3410/B2–52.
206. E. N. G. Marsh, D. P. Patterson, L. Li, *ChemBioChem.* **2010**, *11*, 604–621.
207. D. P. Dowling, J. L. Vey, A. K. Croft, C. L. Drennan, *Biochim. Biophys. Acta* **2012**, *1824*, 1178–1195.
208. J. B. Broderick, B. R. Duffus, K. S. Duschene, E. M. Shepard, *Chem. Rev.* **2014**, *114*, 4229–4317.
209. B. J. Landgraf, E. L. McCarthy, S. J. Booker, *Ann. Rev. Biochem.* **2016**, *85*, 485–514.
210. J. Bridwell-Rabb, T. A. J. Grell, C. L. Drennan, *Ann. Rev. Biochem.* **2018**, 87, 555–584.
211. W. E. Broderick, B. M. Hoffman, J. B. Broderick, *Acc. Chem. Res.* **2018**, *51*, 2611–2619.
212. K. Yokoyama, E. A. Lilla, *Nature Prod. Rep.* **2018**, *35*, 660–694.
213. *Radical SAM Enzymes*, in Vol. 606 of *Methods in Enzymology*, Ed. V. Bandarian, 2018, pp. 1–522.
214. G. L. Cantoni, *J. Am. Chem. Soc*. **1952**, *74*, 2942–2943.

215. P. K. Chiang, R. K. Gordon, J. Tal, G. C. Zeng, B. P. Doctor, K. Pardhasaradhi, P. P. McCann, *FASEB J.* **1996**, *10*, 477–480.
216. J. B. Broderick, R. E. Duderstadt, D. C. Fernandez, K. Wojtuszewski, T. F. Henshaw, M. K. Johnson, *J. Am. Chem. Soc.* **1997**, *119*, 7396–7397.
217. R. Künzler, T. Pils, R. Kappl, J. Hüttermann, J. Knappe, *J. Biol. Chem.* **1998**, *273*, 4897–4903.
218. M. K. Johnson, *Curr. Op. Chem. Biol.* **1998**, *2*,173–181.
219. R. Grazina, S. R. Pauleta, J. J. G. Moura, I. Moura, in Vol. 3 of *Comprehensive Inorganic Chemistry II*, 2nd ed., Eds J. Reedijk, K. R. Poeppelmeier, Elsevier Ltd., Amsterdam, 2013, pp. 103–148. http://dx.doi.org/10.1016/B978-0-08-097774-4.00308-9.
220. E. Akiva, S. Brown, D. E. Almonacid, A. E. Barber 2nd, A. F. Custer, M. A, Hicks, C. C. Huang, F. Lauck, S. T. Mashiyama, E. C. Meng, D. Mischel, J. H. Morris, S. Ojha, A. M. Schnoes, D. Stryke, J. M. Yunes, T. E. Ferrin, G. L. Holliday, P. C. Babbitt, *Nucleic Acids Res.* **2014**, *42*, D521–530.
221. F. Berkovitch, Y. Nicolet, J. T. Wan, J. T. Jarrett, C. L. Drennan, *Science* **2004**, *303*, 76–79.
222. N. D. Lanza, S. J. Booker, *Biochim. Biophys. Acta* **2015**, *1853*, 1316–1334.
223. M. I. McLaughlin, N. D. Lanz, P. J. Goldman, K.-H. Leeb, S. J. Booker, C. L. Drennan, *Proc. Natl. Acad. Sci. USA* **2016**, *113*, 9446–9450.
224. A. C. Rosenzweig, *Science* **2017**, *358*, 307–308.
225. E. L. McCarthy, S. J. Booker, *Science* **2017**, *358*, 373–377.
226. L. J. Reed, B. G. DeBusk, I. C. Gunsalus, C. S. Hornberger, Jr., *Science* **1951**, *114*, 93–94.
227. L. J. Reed, *J. Biol. Chem.* **2001**, *276*, 38329–38336.
228. K. Knittel, A. Boetius, *Annu. Rev. Microbiol.* **2009**, *63*, 311–334.
229. S. B. Joye, *Nature* **2012**, *491*, 538–539.
230. J. Milucka, T. G. Ferdelman, L. Polerecky, D. Franzke, G. Wegener, M. Schmid, I. Lieberwirth, M. Wagner, F. Widdel, M. M. M. Kuypers, *Nature* **2012**, *491*, 541–546.
231. S. Scheller, H. Yu, G. L. Chadwick, S. E. McGlynn, V. J. Orphan, *Science* **2016**, *351*, 703–707.
232. E. Chi Fru, N. P. Rodríguez, C. A. Partin, S. V. Lalonde, P. Andersson, D. J. Weisse, A. El Albani, I. Rodushkin, K. O. Konhauser, *Proc. Natl. Acad. Sci. USA* **2016**, *113*, 4941–4946.
233. W. D. Leavitt, A. S. Bradley, A. A. Santos, I. A. C. Pereira, D. T. Johnston, *Front. Microbiol.* **2015**, *6*, 1392; doi: 10.3389/fmicb.2015.01392.
234. G. Antler, A. Pellerin, *Front. Microbiol.* **2018**, *9*:519; doi: 10.3389/fmicb.2018.00519.
235. M. Sub Sim, H. Ogata, W. Lubitz, J. F. Adkins, A. L. Sessions, V. J. Orphan, S. E. McGlynn, *Nature Commun.* **2019**, *10*, 44; doi: 10.1038/s41467-018-07878-4.
236. J. K. G. Prince, R. H. Rainbird, B. A. Wing, *Geology* **2019**, *47*, 375–379.
237. C.-H. Wu, G. J. Schut, F. L. Poole II, D. K. Haja, M. W. W. Adams, *J. Biol. Chem.* **2018**, *293*, 16687–16696.
238. T. Wagner, U. Ermler, S. Shima, *Science* **2016**, *354*, 114–117.
239. E. O'Brien, M. E. Holt, M. K. Thompson, L. E. Salay, A. C. Ehlinger, W. J. Chazin, J. K. Barton, *Science* **2017**, *355*, eaag1789; doi: 10.1126/science.aag1789.
240. M. Dong, Y. Zhang, H. Lin, *Biochemistry* **2018**, *57*, 3454–3459.
241. M. Dong, V. Kathiresan, M. K. Fenwick, A. T. Torelli, Y. Zhang, J. D. Caranto, B. Dzikovski, A. Sharma, K. M. Lancaster, J. H. Freed, S. E. Ealick, B. M. Hoffman, H. Lin, *Science* **2018**, *369*, 1247–1250.
242. D. Mandalapu, X. Ji, Q. Zhang, *Biochemistry* **2019**, *58*, 36–39.
243. T. Akaike, T. Ida, F.-Y. Wei, M. Nishida, Y. Kumagai, Md. M. Alam, H. Ihara, T. Sawa, T. Matsunaga, S. Kasamatsu, A. Nishimura, M. Morita, K. Tomizawa, A.

Nishimura, S. Watanabe, K. Inaba, H. Shima, N. Tanuma, M. Jung, S. Fujii, Y. Watanabe, M. Ohmuraya, P. Nagy, M. Feelisch, J. M. Fukuto, H. Motohashi, *Nature Commun.* **2018**, *8*, 1177; doi: 10.1038/s41467–017–01311-y.
244. R. Wang, *Physiol. Rev.* **2012**, *92*, 791–896.
245. L. L. Barton, N. L. Ritz, G. D. Fauque, H. C. Lin, *Dig. Dis. Sci.* **2017**, *62*, 2241–2257.
246. Y. Fu, J. Tang, G.-F. Yao, Z.-Q. Huang, Y.-H. Li, Z. Han, X.-Y. Chen, L.-Y. Hu, K.-D. Hu, H. Zhang, *Front. Plant Sci.* **2018**, *9*:1404; doi: 10.3389/fpls.2018.01404.
247. X. Fan, W. Sun, F. Meng, A. Xing, J. Liu, *Green Energy & Environment* **2018**, *3*, 2–19.
248. T. Ould Ely, D. Kamzabek, D. Chakraborty, M. F. Doherty, *ACS Appl. Energy Mater.* **2018**, *1*, 1783–1814.
249. R. S. Glass, *Top. Curr. Chem.* **2018**, *376*:22; doi: 10.1007/s41061S018-0197-0.
250. H. J. Reich, R. J. Hondal, *ACS Chem. Biol*. **2016**, *11*, 821–841.

3

The Type 1 Blue Copper Site: From Electron Transfer to Biological Function

Trinidad Arcos-López, Nils Schuth, and Liliana Quintanar

Department of Chemistry, Center for Research and Advanced Studies (Cinvestav),
Mexico City, D.F. 07360, México
<tarcos@cinvestav.mx>
<nils.schuth@cinvestav.mx>
<lilianaq@cinvestav.mx>

Metal Ions in Life Sciences, Volume 20 Guest Editors: Martha E. Sosa Torres and Peter M. H. Kroneck
Series Editors: Astrid Sigel, Eva Freisinger, and Roland K. O. Sigel

https://doi.org/10.1515/9783110589757-003

Abstract: Cupredoxins host in their scaffold one of the most studied and interesting metal sites in biology: the type 1 (T1) or blue Cu center. Blue Cu proteins have evolved to play key roles in biological electron transfer and have the ability to react with a wide variety of redox partners. The inner coordination sphere of T1 Cu sites conserves two histidines and one cysteine with a short Cu–S(Cys) bond as ligands in a trigonal arrangement, with a variable axial ligand that modulates the electronic structure and reactivity. The structural, electronic and geometric features of T1 Cu centers provide the basis for a site that can be optimized by the protein structure for each biological function. This chapter highlights the properties that make this unique Cu center in biology an efficient and tunable electron transfer site. The contributions of the first coordination shell and the high covalency of the Cu–S(Cys) bond in the T1 Cu site to its distinctive geometric and spectroscopic features are discussed, as well as the role of the protein scaffold in imposing an 'entatic' state with a distorted tetrahedral geometry that minimizes geometric changes upon redox cycling. The analysis of naturally occurring perturbed blue Cu sites provides further insights into how the protein scaffold can tune the properties of the T1 Cu site. Blue Cu sites display a wide range of reduction potentials, as these are tuned to be consistent with their physiologically relevant electron donors and acceptors. The different properties of the protein matrix that play important roles in fine-tuning the reduction potential of T1 Cu sites are also discussed, including the nature of the axial ligand and outer coordination sphere effects. These concepts are further illustrated by the discussion of examples of biosynthetic blue Cu proteins. Finally, the different features of the T1 Cu site that make it an optimal site for electron transfer (ET) are discussed, in terms of Markus theory for intra- and inter-molecular ET. The active site in multicopper oxidases is used as an example to illustrate the contributions of the anisotropic covalency of the blue Cu site to an efficient ET, while the diverse reactivity of the T1 Cu sites in these enzymes is discussed to dissect the different properties provided by the protein that help tune these unique sites for biological ET.

Keywords: blue copper · blue copper proteins · electron transfer · multicopper oxidases · type 1 copper

1. INTRODUCTION: BLUE COPPER SITES IN BIOLOGY

First-row transition metals are essential cofactors in metalloproteins as they facilitate a large variety of chemical processes due to their accessible range of redox-states as well as a flexible number of ligation partners. Numerous biological pathways depend on at least one metalloprotein. Thus, transition metals are essential for life. The evolution of metalloproteins reflects the bioavailability of metals, especially in the oceans. In the anoxic or euxinic (anoxic and sulfidic) ocean that existed in the early Precambrian era, copper was mostly available as insoluble Cu(I) salts, such as Cu_2S. It was after the great oxygenation event, 2.4–2.7 billion years ago, when an increase in atmospheric oxygen changed the bioavailability of copper, given the high solubility of Cu(II) salts. Thus, copper emerged as one of the last cofactors to be incorporated in biological systems [1, 2]. Though being a late addition for evolutionary consideration and toxic for most organisms, the versatility of functionalized copper in terms of properties, such as reactivity and reduction potential, resulted in a remarkable variety of organisms utilizing copper proteins, including bacteria, plants, and humans [3].

Several protein cofactors containing copper binding sites have been characterized and classified by their spectroscopic features, such as mononuclear type 1, mononuclear type 2, and binuclear type 3 Cu sites. Cu sites have also been classified by their function, including type 1 Cu and Cu_A sites as electron transfer centers; binuclear and trinuclear Cu sites for oxygen activation and reduction; and tetranuclear Cu_Z sites for N_2O reduction [4]. From this diversity of Cu sites that we know today, the first one to be utilized in proteins was the type 1 (T1) Cu center, also named blue Cu site [1, 2]. The unique spectral features of T1 copper sites, as compared to synthetic copper molecules, made blue Cu proteins an attractive research topic [4–7]. Oxidized T1 Cu sites are distinguished by an intense absorption maximum around 600 nm, resulting in the intense blue color characteristic of these sites, and a remarkably small hyperfine splitting in their electron paramagnetic resonance (EPR) spectra [8–10]. The unique spectral features of blue Cu sites are discussed further in Section 2.2.

The majority of T1 Cu sites are present in cupredoxins, also known as blue copper proteins, a family of proteins named according to its electron transfer function similar to ferredoxins [11]. Cupredoxins are a superfamily of small (<20 kDa) and soluble proteins containing a single T1 site. They can be found in all the three kingdoms Archaea, Bacteria, and Eukarya with a majority in plants in terms of number of unique proteins. In cupredoxins, both oxidized and reduced states are easily interchangeable and, in some cases, withstand oxidation in an oxidizing environment. Additionally, cupredoxins are capable to reduce/oxidize a variety of proteins, mostly cytochromes, with remarkable reduction/oxidation rates facilitated by hydrophobic patches interacting with the reduction/oxidation partner. Thus, earliest studies described blue copper proteins as electron transport systems [8, 9]. This proposed function is further supported by a rigid tertiary structure, which is neither perturbed by redox cycling, nor by the removal of the copper ion [12–14]. The versatility of cupredoxin oxidation/reduction partners is one of the mayor obstacles in conclusively identifying physiologically relevant redox partners. Additionally, blue Cu proteins (BCPs) often share their role as electron shuttle with cytochromes, such as cytochrome c_6 in photosynthesis, which are also potential donors/acceptors [15–20].

Numerous cupredoxins have been characterized and their amino acid sequences determined [21–23]. Additionally, a remarkable number of structures from X-ray diffraction and nuclear magnetic resonance have been reported. Sequence analysis from different small BCPs resulted in a phylogenetic tree with three branches separating at the trunk: (i) plastocyanin-like BCPs occurring in bacteria and plants, (ii) azurin-like BCPs exclusive to bacteria and (iii) phytocyanins only found in plants [11, 21, 24]. The plastocyanin-like branch is divided further into the subfamilies plastocyanins, pseudoazurins, and amicyanins, while the phytocyanin family includes stellacyanins, uclacyanins, plantacyanins, and early nodulin proteins, the latter uncapable to bind copper [22]. The azurin family shares common ancestors with multi-domain enzymes containing cupredoxin-like domains such as nitrite reductase and cytochrome oxidase [25]. Rusticyanins are proteins only found in the bacterium *Thiobacillus ferrooxidans* and share an even closer relation to multicopper oxidases and nitrite reductases, having the T1 Cu sites

with the highest reduction potentials of single domain cupredoxins [24, 26, 27]. Sequence identity between subfamilies is remarkably low (<20 %). However, all BCPs feature an eight-stranded Greek key β-barrel and extended hydrogen networks achieving the structural rigidity, though the actual order of β-sheets varies strongly between subfamilies [23, 28, 29].

Cupredoxins host in their scaffold one of the most studied and interesting metal sites in biology: the T1 or blue Cu center. BCPs fulfill an important function in biology as ET sites. The unique electronic structure and geometric properties of T1 Cu centers, as well as the particular structural features of each cupredoxin scaffold, provide these sites with key properties that make them optimal for ET, as discussed in the following sections.

2. THE BLUE COPPER CENTER: A MISTERY AT THE BIRTH OF BIOLOGICAL INORGANIC CHEMISTRY

2.1. The Type 1 Blue Copper Site: Born to Do Electron Transfer?

In 1939, Green et al. demonstrated the inhibitory effect of copper chelators on photosynthesis and respiration of the green water alga *Chlorella pyrenoidosa* [30]. The involvement of copper in photosynthesis was discovered in 1960 by Katoh and Takamiya, who first isolated plastocyanin (Pc) from the green algae *Chlorella ellipsoidea* [8, 31]. Around the same time, azurin (Az) was first isolated from *Bordetella* in 1963 [9] to become later one of the first blue copper proteins to be structurally characterized and sequenced [32, 33]. A few years later, the first phytocyanin was isolated as a very basic protein from cucumber seed and peel, and named the cucumber basic protein (CBP) or cucumber plantacyanin, though it had also been found in a variety of plants, e.g., spinach plantacyanin or spinach basic protein [34–36]. Initial characterization of isolated BCPs showed that their intense blue color would disappear upon addition of reducing reagents. It would take several years for the origin of the blue color to be discovered, while in some cases their physiological roles remain elusive even today, as is the case of azurin and phytocyanins. This is surprising, considering how widely expressed BCPs are in nature. For example, phytocyanins are a family of plant-specific BCPs that have been identified in a number of different plants including maize, cotton, tobacco, rice, cucumber, spinach, cabbage and, most importantly, in the small weed flower *Arabidopsis thaliana* [11, 37, 38]. The extensive efforts of sequencing the genome of *Arabidopsis thaliana* reveal phytocyanins as one of the largest protein families known in plants [39, 40]. In spite of their omnipresence in plants, and the fact that one of the members of this family, stellacyanin, has been extensively studied *in vitro*, the physiological role of phytocyanins remains unknown.

Bioinorganic chemists know that the blue Cu center is a site with unique features that makes it optimal for rapid long-range ET, as discussed in this chapter.

Yet, the challenge of elucidating the particular physiological role of each blue copper protein still stands, and it lies on the identification of their redox partners *in vivo*. In the case of plastocyanin, its role in photosynthesis is clear, and it is expressed in cyanobacteria, most green algae, and all higher plants [15, 16, 19, 41]. Localized in the inner space (lumen) formed by thylakoid membranes, Pc acts as an electron shuttle between photosystems II and I, large trans-membrane complexes formed by several proteins, involved in photosynthesis. The electron from water splitting in photosystem II reduces the cytochrome b_6f complex, the electron donor for Pc, which in turn reduces the $P700^+$ subunit of photosystem I [42, 43]. Pc resists oxidation by the oxic environment created by photosynthetic water splitting, suggesting that this blue copper protein evolved as an ET system that can keep copper redox cycling under aerobic conditions.

The fact that Pc is expressed in cyanobacteria, the first organisms on Earth to come in contact with oxygen, underscores the notion that the first biological role of copper incorporated in living organisms was that of ET to enable photosynthesis. Indeed, the Cu-binding cupredoxin fold superfamily evolved early and it is ubiquitous in eukaryotic proteomes, suggesting that their early role as electron transport proteins accompanied eukaryotic evolution [1]. Of course, as living organisms evolved to deal with increasing oxygen concentrations on our planet, copper became more bioavailable, and it was also incorporated into enzymes that can reduce or activate oxygen, including cytochrome *c* oxidase, which is a key player in respiration [44]. As a result of this evolution, today the role of T1 blue Cu centers (and Cu_A sites as well) as ET sites that play essential roles in both, photosynthesis and respiration, is undeniable. In fact, one could simply just say that life on our planet would not be the same without blue copper proteins that assure efficient electron transport. In the sections that follow, a summary of the unique structural, electronic, and geometric features of the blue Cu site is provided, to further highlight the properties that make this unique Cu center in biology an optimized site for electron transfer.

2.2. Origin of Unique Features of the Blue Copper Center: A Highly Covalent Matter

Since their discovery, the origin of the intense blue color characteristic of blue copper proteins puzzled bioinorganic chemists. Electronic absorption spectra of these proteins show an intense ($\varepsilon \sim 5000\ M^{-1}cm^{-1}$) maximum at 600 nm (Figure 1A). However, tetragonal and square planar Cu(II) complexes normally show weak d → d transitions ($\varepsilon \sim 40\ M^{-1}cm^{-1}$) in this region. This is because Cu(II) complexes adopt a preferred square planar geometry due to the Jahn-Teller distorting force present in the Cu(II) state. Square planar Cu(II) complexes have D_{4h} symmetry, and thus, all their ligand field (LF) transitions are parity forbidden, yielding very low intensity signals (Figure 1A and D). In order to explain the intense blue band centered at 600 nm, before the first blue Cu

protein was structurally characterized, early spectroscopic studies performed by Solomon, Hare, and Gray predicted a tetrahedral geometry for the T1 Cu center [45]. This was based on circular dichroism (CD) and magnetic circular dichroism (MCD) studies of *Rhus vernicifera* stellacyanin, bean plastocyanin, and *Pseudomonas aeruginosa* azurin. They predicted a flattened tetrahedral (D_{2d}) site with two His and one Cys as Cu ligands, and weak LF transitions appearing at lower energy (above 800 nm). The intense absorption band at 600 nm was deconvoluted as two ligand-to-metal charge transfer (LMCT) bands originating from the S(Cys) $\pi \rightarrow$ Cu $3d_{x^2-y^2}$ and S(Cys) $\sigma \rightarrow$ Cu $3d_{x^2-y^2}$ transitions [45]. The fact that the absorption spectrum of blue Cu sites reflected a distorted tetrahedral geometry led to the idea that the protein opposes the Jahn-Teller distortion that a normal Cu(II) complex would suffer from. By imposing a T_d geometry on the T1 Cu center, the protein would be activating the site for ET, enforcing a rack-induced or 'entatic' state (Section 3.1) [46, 47]. Today we know that the inner coordination sphere of T1 Cu sites is mostly conserved, with two Cu(II)–N(His) bonds at ~ 2 Å, which are common to Cu(II)-imidazole interactions [5], and a Cu(II)–S(Cys) bond at 2.1 Å that is shorter than most other Cu(II)–S thioether bonds (~ 2.5 Å). In those sites with a Met as an axial ligand, like the early studied bean Pc and *Pseudomonas aeruginosa* Az, a long axial Cu–S(Met) bond at 2.9 Å completes the coordination sphere (Figure 1C) [48, 49].

Another unique feature of blue Cu sites is that they display an unusually small metal hyperfine splitting in their EPR spectra. Cu(II) ions have a d^9 electronic configuration and are paramagnetic. The metal hyperfine coupling results from the interaction of the electron spin in the $3d_{x^2-y^2}$ orbital with the nuclear spin of ^{63,65}Cu (I = 3/2), which leads to the characteristic four line splitting pattern in the parallel region of Cu(II) EPR spectra. Normal Cu(II) complexes exhibit a large metal hyperfine coupling, with $A_{||}$ values ranging from 140 to 200 $\times 10^{-4}$ cm^{-1}; for $[CuCl_4]^{2-}$, $A_{||} = 164 \times 10^{-4}$ cm^{-1}. In contrast, T1 Cu sites display a metal hyperfine splitting that is decreased by more than a factor of two, as compared to normal Cu(II) complexes: for the blue site in Pc, $A_{||} = 63 \times 10^{-4}$ cm^{-1} (Figure 1B) [6, 7]. Initially, this decreased metal hyperfine interaction was thought to originate from the distorted T_d symmetry of the site, through mixing of the Cu $4p_z$ character into the $3d_{x^2-y^2}$ ground state of the complex; however, detailed X-ray absorption spectroscopy (XAS) studies determined that this was not the case [50, 51].

A decreased metal hyperfine splitting in blue Cu sites reflects on a significantly decreased interaction of the unpaired electron spin in the $3d_{x^2-y^2}$ orbital with the metal nuclear spin, which would be a result of having a very covalent site with the Cu electron spin delocalized into the ligand orbitals [52]. Indeed, the short Cu–S(Cys) bond distance reflects a strong and highly covalent site, as determined by electronic structure calculations using density functional theory (DFT). The ground state wavefunction for the T1 Cu site shows an important delocalization of the electron spin density into the $p\pi$ orbital of the sulfur atom of the Cys thiolate: 60 % S(p) and only 40 % Cu(d) character, as determined by early DFT studies [53]. The high covalency of the T1 Cu site has been confirmed by elegant

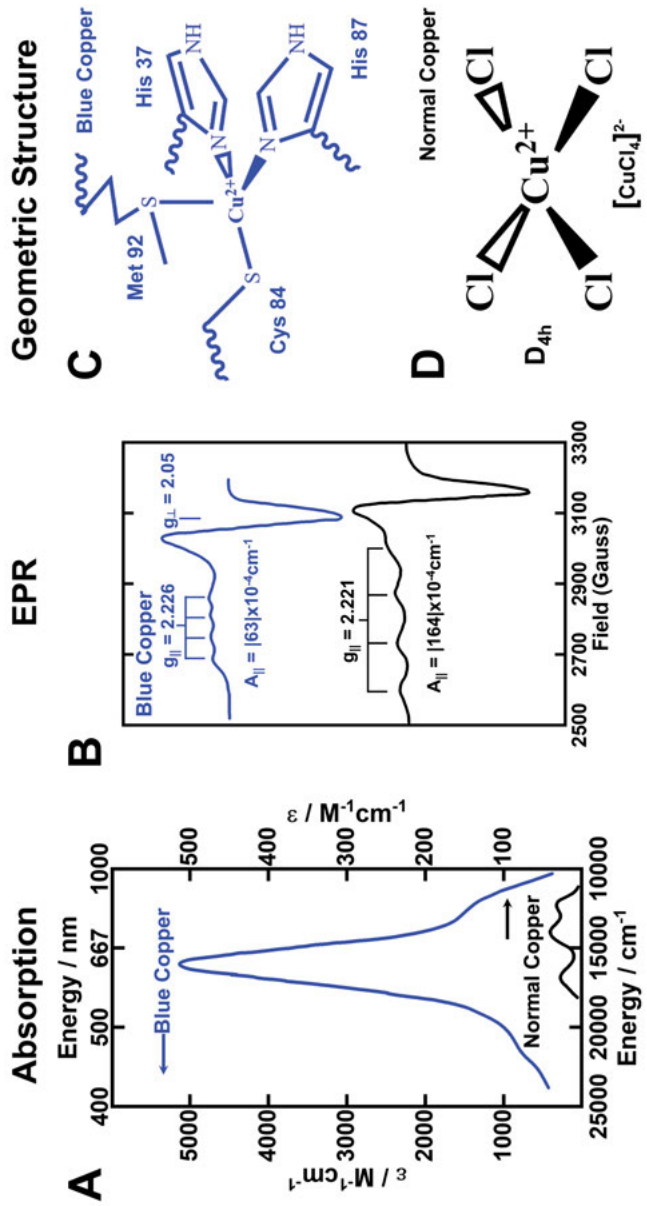


Figure 1. (**A**) Comparison of the electronic absorption, (**B**) X-band EPR spectra, and (**C** and **D**) geometric structures of plastocyanin (blue) and "normal" D_{4h} $[CuCl_4]^{2-}$ (black). Adapted with permission from [53], copyright 2004, American Chemical Society.

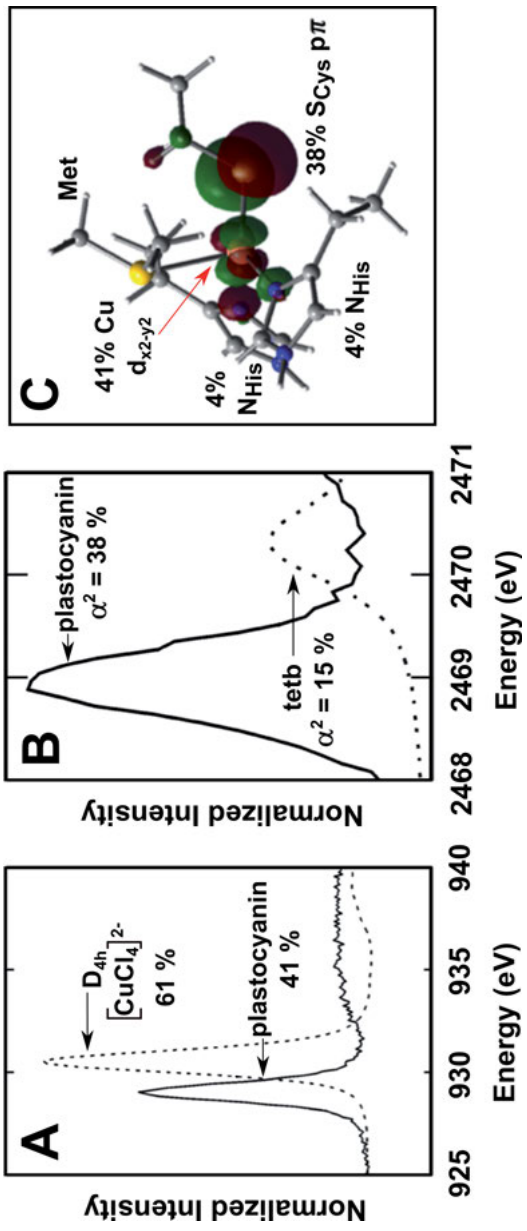


Figure 2. (**A**) Comparison of Cu L_3 edge intensities of D_{4h} $[CuCl_4]^{2-}$ or normal Cu (dotted line) and plastocyanin (solid line). (**B**) Comparison of sulfur K-edges of tetb (normal Cu/S model, dotted line) and plastocyanin (blue Cu, solid line). (**C**) Ground-state wave function (β-LUMO) of the blue Cu site in plastocyanin calculated by Xα-SW. Parts (A) and (B) adapted with permission from [53], copyright 2004 American Chemical Society. Part (C) adapted with permission from [6], copyright 2011 Elsevier.

XAS studies that compare the spectroscopic features of model complexes with that of a blue Cu site [51, 54, 55]. Metal L-edge XAS was used to quantify the $3d_{x^2-y^2}$ character in the highest occupied molecular orbital (HOMO): since the Cu(2p) to Cu(3d) transition is electric dipole-allowed, while the Cu(2p) to S(3p) transition is not, the intensity of the Cu(2p) to HOMO transition at 930 eV can directly report on the Cu(3d) character in the HOMO. Thus, a quantitative comparison of Cu L-edge XAS of a D_{4h} $[CuCl_4]^{2-}$ complex with the blue Cu site in Pc, allowed to determine that the latter has only ~41 % Cu $3d_{x^2-y^2}$ character (Figure 2A) [55]. Such low Cu $3d_{x^2-y^2}$ character underscores the high covalency of the blue Cu center, and is consistent with having a very small metal hyperfine splitting, as observed by EPR. On the other hand, ligand K-edge XAS can be used to probe the sulfur p character in HOMO, which would reflect the delocalization into the ligand S(p) from the Cys thiolate. The sulfur K-edge feature at 2470 eV reports on the S(1s) to HOMO transition. Since the S(1s) to S(3p) transition is electric dipole-allowed, while the S(1s) to Cu(3d) is not, the 2470 eV signal can directly report on the sulfur 3p character in the HOMO. A quantitative comparison of sulfur K-edge XAS of a Cu(II)-thiolate model complex with the blue Cu site in Pc, allowed to determine that the latter has ~38 % S(3p) character (Figure 2B) [51, 54]. These seminal XAS studies demonstrated the high covalency of the Cu(II)–S(Cys) bond, and provided the basis to explain the decreased metal hyperfine splitting that is characteristic of blue Cu sites.

The highly covalent Cu(II)–S(Cys) bond of T1 Cu centers also gives origin to the intense S(Cys) to Cu charge transfer bands responsible for the intense blue color of these sites, since the intensity of LMCT transitions reflect the overlap of the ligand donor and metal acceptor orbitals. A detailed MCD study of blue Cu sites at low temperature showed that indeed this is the case; however, the relative orientation of the Cu $3d_{x^2-y^2}$ and the S(Cys) 3p orbitals is different from the normal tetragonal Cu(II) complexes. Typically, the Cu $3d_{x^2-y^2}$ orbital would be oriented to maximize the σ overlap with the ligand donor orbital, the S(Cys) 3p orbital, that lies in the direction of the Cu–S bond; this is the case for normal tetragonal complexes and for the green Cu sites (Figure 3). Such relative orientation of the Cu $3d_{x^2-y^2}$ and the S(Cys) 3p orbitals yields normally a strong σ overlap leading to an intense S(Cys) σ → Cu $3d_{x^2-y^2}$ transition, and a weak π overlap that yields a low intensity S(Cys) π → Cu $3d_{x^2-y^2}$ transition. However, seminal MCD studies demonstrated that the T1 Cu center has the relative strengths of these interactions inverted, i.e., the σ overlap is small and the S(Cys) σ → Cu $3d_{x^2-y^2}$ transition has low intensity, while the strongest interaction comes from a large π overlap that yields a high intensity S(Cys) π → Cu $3d_{x^2-y^2}$ transition. This implies that the Cu $3d_{x^2-y^2}$ orbital in the T1 Cu site is oriented to maximize the π overlap with the S(Cys) 3p, i.e., is rotated 45º such that the $3d_{x^2-y^2}$ lobes bisect the Cu–S(Cys) bond (Figure 3).

In summary, the unique spectroscopic features of the T1 Cu site are due to a distorted tetrahedral geometry with a highly covalent Cu(II)–S(Cys) π bond. The distorted geometry imposed by the protein structure (‘entatic’ state) and

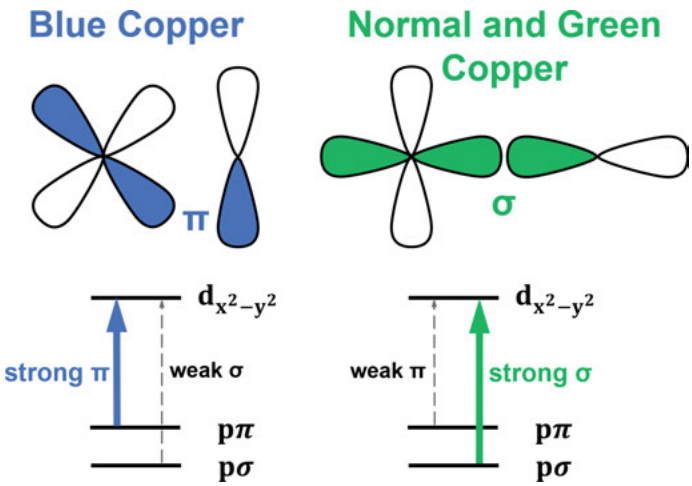


Figure 3. Comparison of schematic representations of Cu $d_{x^2-y^2}$ and S(Cys) p orbitals for a classic blue copper site, and normal and green copper sites, highlighting the relative intensities of the σ and π S(Cys) to Cu(II) charge transfer transitions in each case.

the unusually high covalency of the Cu–S thiolate bond provide the blue Cu site with unique tunable properties that make it optimal for ET function, as it will be discussed in Sections 3 and 4. Indeed, the protein scaffold plays an important role in tuning the unique geometric and electronic properties of the T1 Cu center, as evidenced by the naturally occurring perturbed blue Cu sites discussed below.

2.3. Perturbed Type 1 Copper Sites

The unique spectroscopic features of T1 Cu sites are due to the strong overlap between the metal $3d_{x^2-y^2}$ orbital and the S(Cys) 3p orbitals, resulting in a highly covalent site. The nature of the interaction between these orbitals allows for a classification of known T1 Cu sites into two sets [56]: (i) The classic blue copper sites, such as azurin and plastocyanin, with spectral features as described above; and (ii) the group of perturbed blue copper sites, including rusticyanin, (Rc), pseudoazurin, (PAz), CBP and nitrite reductases, (NiR). Most of the latter proteins have the same ligand set as Pc [S(Cys), 2 N(His), and S(Met)], but they exhibit perturbed spectral features: a rhombic EPR signal ($\Delta g_{\perp} = g_x - g_y > 0.01$) and an absorption maximum at 450 nm, in addition to the intense transition at 600 nm in their UV-visible absorption spectra (Figure 4A). The ratio of the two absorption bands ($R\varepsilon = \varepsilon_{450}/\varepsilon_{600}$) can be as low as 0.15 for the classic site in Pc, or as high as 1.2 for NiR; usually, perturbed blue copper sites display $R\varepsilon = \varepsilon_{450}/\varepsilon_{600} \geq 0.5$. The relative intensity of the absorption maximum at 450 nm has been associated with an enhanced pseudo σ-type of interaction (instead of a π-type) between the metal

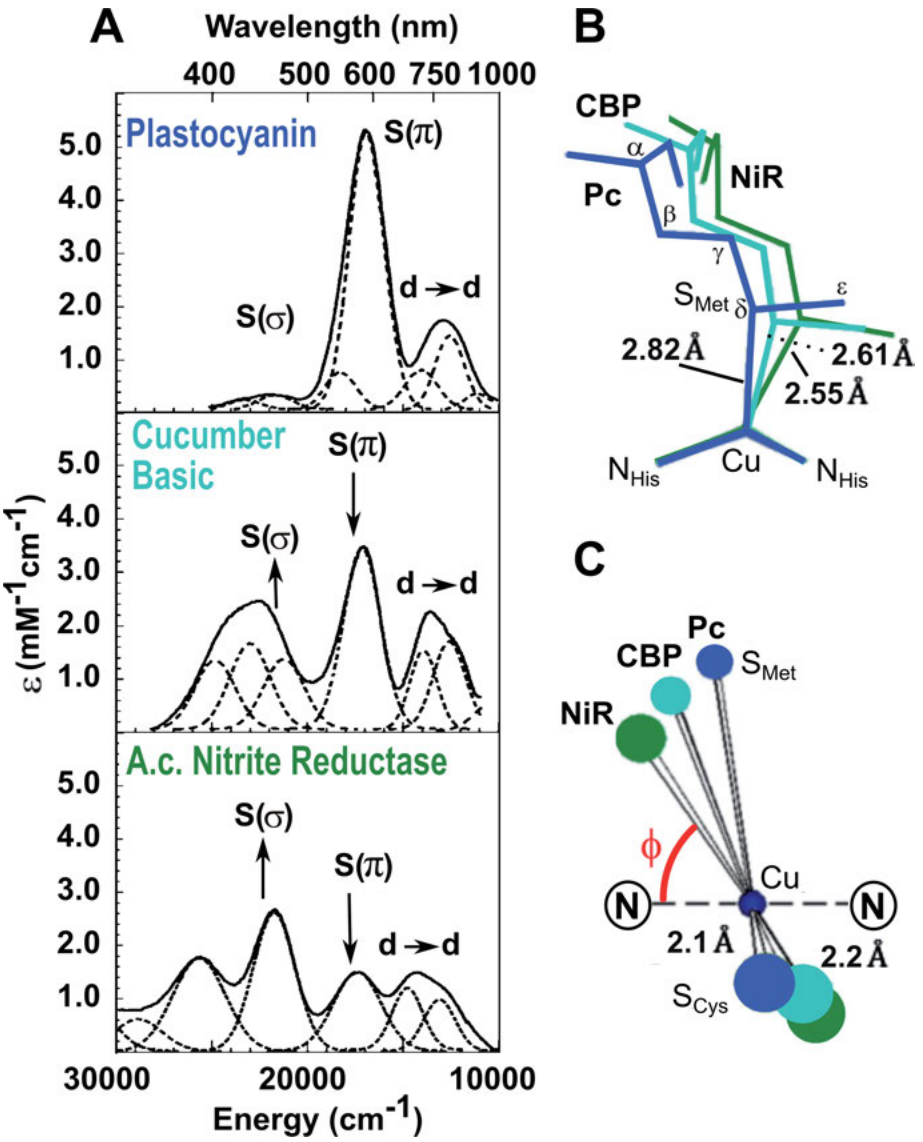


Figure 4. (**A**) Comparison of low-temperature absorption spectra of a series of blue copper proteins with the same ligand set [S(Cys), 2 N(His), S(Met)] Pc (classic blue site), CBP (perturbed blue site) and NiR (green site); showing increasing intensity in the ~450 nm (Cys pseudo-σ CT) band as the ~600 nm (Cys π CT) band decreases. (**B**) Comparison of crystal structures of the T1 Cu sites in these proteins, showing contraction of the Cu–S(Met) bond associated with elongation of the Cu–S(Cys) bond, and (**C**) variation of the dihedral angle ϕ between the planes N(His)-Cu-N(His) and S(Cys)-Cu-S(Met). Adapted with permission from [53], copyright 2004, American Chemical Society.

$3d_{x^2-y^2}$ and the ligand S(3p) orbitals, which varies according to the geometry of the site [4, 57]. The ϕ dihedral angle formed between the N(His)-Cu-N(His) and S(Met)-Cu-S(Cys) planes are related to the Cu-S(Cys) interaction (Figure 4C): as ϕ decreases, the pseudo σ interaction increases, rendering a more perturbed site; in contrast, classic blue copper sites have a larger ϕ angle and a decreased pseudo σ interaction. The NiR site displays the maximum pseudo σ interaction, its severe tetragonal distortion leads to a strong absorption at 450 nm and an intense green color (Figure 4A).

A comparison of a series of naturally occurring T1 Cu centers (Pc, CBP, NiR) can provide further insights into the fine mechanisms that control the spectral features of perturbed sites. Although these proteins share the same ligand set [S(Cys), 2 N(His), and S(Met)] (Figure 4B), their color ranges from blue in Pc to green in NiR, reflecting a change in the nature of the Cu(II)–S(Cys) bond: from a significant π donor character in the classic Cu site in Pc to a greater contribution of a pseudo σ donor interaction in the green site of NiR [6, 7, 34]. Indeed, the S(Cys) $\pi \rightarrow$ Cu $3d_{x^2-y^2}$ charge transfer (CT) signal intensity decreases as that of S(Cys) pseudo $\sigma \rightarrow$ Cu $3d_{x^2-y^2}$ CT increases in this series, going from Pc via CBP to NiR (Figure 4A); reflecting changes in their geometric structure. Specifically, the Cu(II)–S(Met) bond contracts by 0.3 Å and the Cu(II)–S(Cys) bond is elongated by 0.1 Å; while the S(Met)-Cu-S(Cys) plane rotates into the N(His)-Cu-N(His) plane, triggering a more tetragonal structure (Figure 4B and C). Overall, a longer Cu–S(Cys) distance and a shorter Cu–S(Met) bond favor the pseudo σ interaction with the Cys ligand, leading to a more perturbed site. Thus, although the protein structure imposes certain structural properties of T1 Cu sites, the strength of the Cu–S(Met) bond determines their distortion, fine-tuning their geometric, electronic, and redox behavior (Section 3.2) [6, 7, 34].

Another good example of the role of the axial interaction in the spectroscopic properties of T1 Cu sites is illustrated by the thermodynamic equilibrium that the enthalpically favored green copper site in NiR displays, as it converts into an entropically favored blue site with an increase in temperature [58–61]. The change from green to blue is associated with a decreased intensity of the pseudo σ CT band, an increased intensity of the π CT transition, and the weakening of the axial thioether ligand (Cu–S(Met) ~ 2.5 Å) at elevated temperatures (Figure 3 and 5A). Consistently, the mutation of the axial Met for a non-coordinating Thr residue in NiR yields a NiR variant with blue Cu features, confirming the role of the axial ligand in the green/blue transition (Figure 5B) [60].

On the other hand, a series of mutations of the axial ligand in cucumber stellacyanin further illustrates the role of this interaction in determining the electronic structure properties of a T1 Cu site. In the wild-type site the axial S(Met) interaction has been replaced by a stronger O(Gln) ligand [62], while two mutations have been studied: Gln99Met and Gln99Leu; providing a O(Gln) to S(Met) to a non-coordinating axial ligand (Leu) series that covers the range of native axial interactions in blue copper sites [63]. The absorption spectra of this set of proteins display a high-intensity π S to Cu LMCT and a weak pseudo σ S to Cu

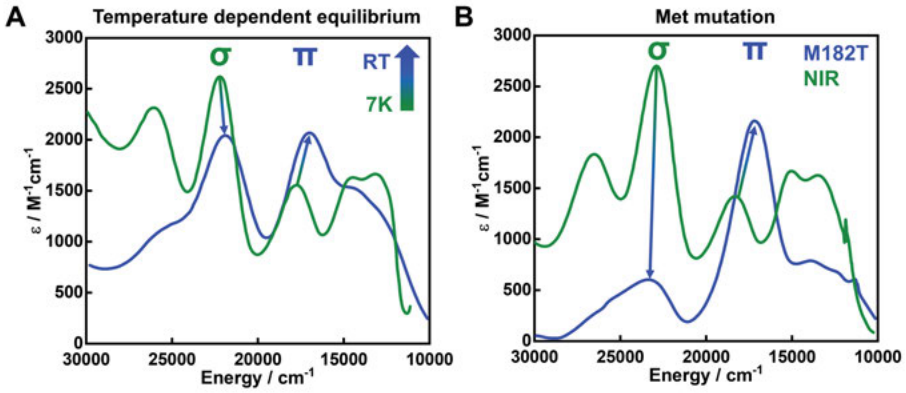


Figure 5. (**A**) Temperature-dependent electronic absorption spectra, showing the thermodynamic equilibrium between green and blue Cu sites in NiR. (**B**) Comparison of electronic absorption spectra of wild-type NiR (green) and Met182Thr NiR (blue). Adapted with permission from [7], copyright 2016, John Wiley and Sons.

LMCT transition, reflecting a similar tetrahedral environment and a blue color for stellacyanins and all its variants. Their MCD spectra show an increase in the LF transitions energies on going from wild-type to Gln99Met to Gln99Leu, indicating that the Cu ion moves into the N(His)-S(Cys)-N(His) plane and the site becomes more trigonal, as the axial ligand is weakened or lost. Consistently, decreased $g_{||}$ values and increased $A_{||}$ values were observed by EPR. Sulfur K-edge XAS shows an increase in the pre-edge intensity, going from wild-type to Gln99Met to Gln99Leu stellacyanin, indicating a more covalent Cu(II)–S(Cys) bond.

Overall, this series of mutated stellacyanins illustrate the correlation of the weakened (or lost) axial ligand interaction with the stronger and more covalent Cu-S(Cys) interaction. A similar observation can be derived from a comparison of T1 Cu sites in multicopper oxidase (MCOs) proteins, such as fungal laccase and Fet3p, where a non-coordinating residue (Leu or Phe) is located at the axial position [64–67]. These sites have a trigonal structure, display an intense π S to Cu CT transition, a weak pseudo σ to Cu CT band, decreased $g_{||}$ and increased $A_{||}$ values, as compared to the classic tetragonal T1 Cu sites in other MCOs [53, 68–70]. The lack of an axial ligand results in a compensating increase in covalency of the Cu–S(Cys) bond. Consistently, mutation of the non-ligating axial residue (Phe463) to a Met in *Polyporus pinsitus* fungal laccase yields a T1 Cu with the electronic structure and spectral features of a classic blue Cu site [70].

In summary, studies of series of mutated T1 Cu sites in different proteins demonstrate that the axial ligand modulates the geometric and electronic structure of a blue copper center. As the axial ligand interaction is weakened (or lost), the metal ion moves into the N(His)-S(Cys)-N(His) plane, the geometry

becomes more trigonal, and the Cu(II)–S(Cys) bond becomes shorter and more covalent; indicating that the strength of the axial ligand inversely affects the strength of the equatorial Cu(II)–S(Cys) bond. The implications of these changes in electronic structure onto the redox properties of the T1 Cu center will be further discussed in Section 3.2.

Other perturbed T1 Cu centers have been generated by mutations of the Cys residue. The double mutation Cys112Asp/Met121Glu in Az results in a novel Cu site that no longer fits the classification of T1 Cu due to the loss of the copper-thiolate interaction. This site has been named "type zero Cu", it has a small parallel hyperfine splitting and displays rapid electron transfer rates characteristic of T1 copper centers [71–74]. The T1 Cu in Az has also been modified by incorporation of selenocysteine (SeC) and selenomethionine (SeM) to generate the Cys112SeC and Met121SeM azurin variants [75]. The spectral features of Met121SeM Az were essentially identical to those of wild-type Az, while the Cys112SeC variant displayed a red-shifted band at 677 nm and an increased parallel hyperfine splitting value ($A_{\|} = 99 \times 10^{-4}\,cm^{-1}$), as compared to wild-type Az ($A_{\|} = 67 \times 10^{-4}\,cm^{-1}$). The Cu(II)–Se(SeC) bond length is only 0.14 Å longer than the Cu–S(Cys) distance, yet it displays only minor changes upon reduction of the metal center [76]. The fact that the structural features of the T1 Cu center suffers little change upon redox cycling, even in the Cys112SeC Az variant, underscores the important role of the protein scaffold in maintaining the 'entatic' state, as it will be discussed in Section 3.1.

Finally, other perturbed T1 Cu centers worth mentioning are those with an open coordination position, namely: the red Cu center in nitrosocyanin (Nc), the purple site in the nitrifying archaeon *Nitrosopumilus maritimes* (Nmar1307) and the His117Gly variant of azurin. Nitrosocyanin is a mononuclear copper protein that shares a high sequence homology and a similar overall fold with Pc, Az, and Rc. However, Nc exhibits an unprecedented trimer of single domain and its T1 Cu displays a square pyramid geometry with a S(Cys) ligand, two N(His), and a water molecule; while a carboxylate oxygen from Glu is on the pyramid apex. The site in Nc lacks the long Cu–S(Met) and short Cu–S(Cys) bonds characteristic of most blue copper centers, and it exhibits a bright red color due to a strong absorption band at 390 nm ($\varepsilon_{390} \sim 7{,}000\,M^{-1}\,cm^{-1}$) [77, 78]. The spectroscopic parameters of this site are more typical of a type 2 Cu center than of a T1 Cu site, exhibiting a large Cu hyperfine splitting ($g_{\|} = 2.25$ and $A_{\|} = 144 \times 10^{-4}\,cm^{-1}$) and a low reduction potential ($E_0 = +85$ mV); its open coordination site suggests a putative catalytic role, rather than an ET function for Nc.

Another T1 Cu site with an open binding position is the purple mononuclear Cu center found in Nmar1307. The site displays a tetragonal distortion with a weak Cu(II)–S(Cys) bond and it can bind a water molecule as an exogenous ligand. The EPR spectrum indicates that this site can adopt two different coordination geometries in nearly equal population: one axial species with a $A_{\|} = 98 \times 10^{-4}\,cm^{-1}$, and a more rhombic one with a smaller $A_{\|}$ ($69 \times 10^{-4}\,cm^{-1}$), corresponding to two different water binding modes. The deep purple color of the protein arises from two strong absorption bands with similar intensity, at 413

and 558 nm [79]. The open coordination position in the purple Cu center in Nmar1307 may also bind other exogenous ligands, such as nitrous oxide, for catalytic purposes. Indeed, having an open coordination position may improve catalytic function, as shown for the His117Gly Az variant, where the loop containing the Cu ligating residues gains flexibility and allows for the coordination of exogenous ligands [80–82]. The His117Gly variant displays a lower reactivity, as compared to wild-type Az, yet the introduction of an exogenous imidazole ligand restores the reactivity above the wild-type level [83]. While the His117Gly Az variant illustrates the role of having an open coordination position in the reactivity of the T1 Cu center, the naturally occurring perturbed T1 Cu centers of Nc and Nmar1307 with open coordination sites may fulfill catalytic functions, going beyond the classic role of the T1 site in electron transfer.

3. FINE-TUNING THE REDOX POTENTIAL OF A BLUE COPPER SITE

T1 Cu centers fulfill an important function in biology as ET sites. Their reduction potentials are fine-tuned to fit closely between those of their respective electron donor and acceptors. For example, in the photosynthetic transport chain cytochrome *f*-Pc-$P700^+$, the E_0 of the blue Cu site in Pc (+370 mV) falls between that of its electron donor cytochrome *f* (+340 mV) and its electron acceptor $P700^+$ (+490 mV). The thermodynamic properties of T1 Cu sites are finely modulated to match their function in biological ET chains, as even a small shift in their reduction potential could have severe consequences. Thus, a wide range of reduction potentials is observed for T1 Cu sites, ranging from +184 mV relative to standard hydrogen electrode (SHE) in stellacyanins [84] to even > +1000 mV for one of the T1 Cu centers in human ceruloplasmin [85]. The latter belongs to the family of MCOs, which are multi-domain proteins with four or six Cu centers, at least one of them being a T1 Cu site. The highest reduction potentials for T1 Cu centers are found in MCOs, as they are optimized to oxidize a wide range of substrates [78, 86–88]. The different reduction potentials for blue copper proteins are depicted in Figure 6A. A family of BCPs that illustrates well the wide variety in spectral features and reduction potentials that T1 Cu sites can have, are the auracyanins, a family of four proteins (A-D) from the filamentous anoxygenic phototroph *Chloroflexus aurantiacus*. The E_0 for their T1 Cu sites ranges from +83 mV (auracyanin D) to +423 mV (auracyanin C), while their spectral features vary from those of a classical blue copper site (auracyanin B) to those with highly perturbed spectra (auracyanins C and D) [89].

The reduction potential for the Cu(II)/Cu(I) pair depends on several factors, including steric hindrance, distortion of the coordination sphere, the hydrophobicity of the environment, and the type of ligands around the site. Sulfur donors and nitrogen π-acceptors generally raise the potential, while anionic

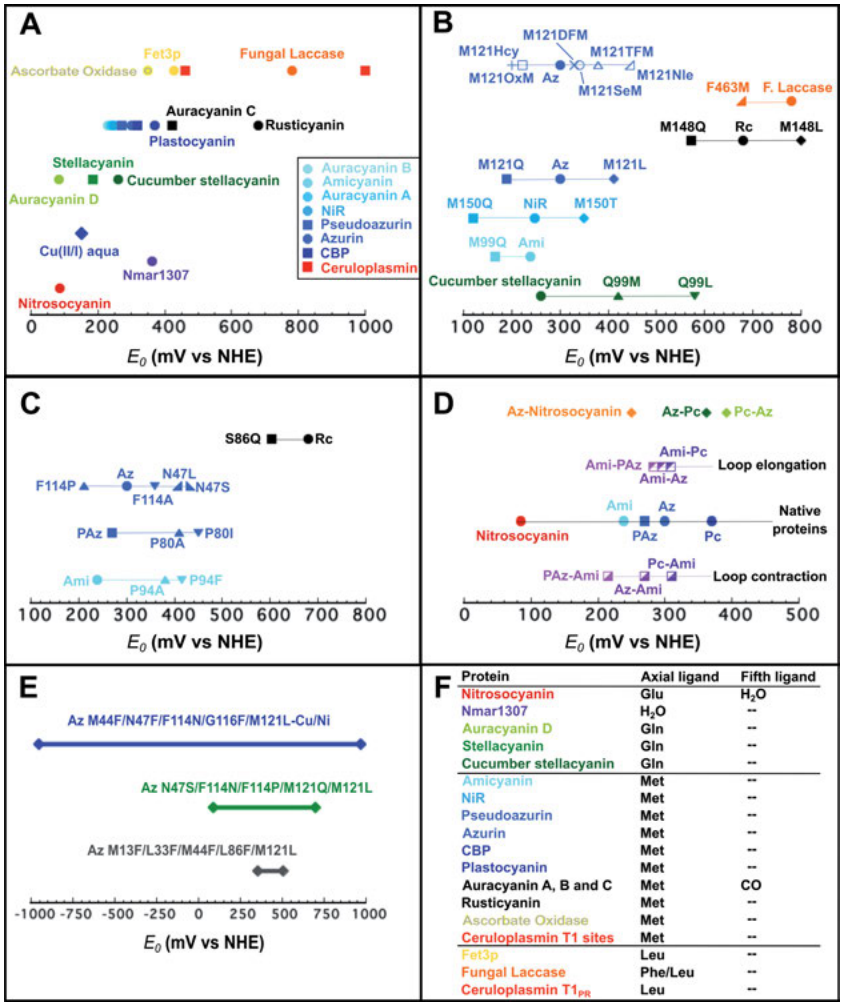


Figure 6. (**A**) Wide range of reduction potentials of native blue copper sites; the redox potential of aqueous Cu(II/I) is shown for comparison. (**B**) Variation of E_0 values of T1 Cu sites as a result of mutating the axial ligand in cucumber stellacyanin, amicyanin, NiR, Az, Rc, and fungal laccase. In the series non-ligating axial ligand to Met to Gln, an increase in donor strength of the axial ligand correlates with a decrease of the Cu–S(Cys) bond strength, resulting in a decreased E_0. (**C**) Variation of E_0 produced by mutations in the outer coordination sphere of the T1 Cu centers in Ami, PAz, Az, and Rc. Elimination or formation of hydrogen bonds result in significant changes in the reduction potential. (**D**) E_0 values of T1 Cu sites in new proteins produced by loop-elongation or loop-contraction mutagenesis. Chimeras display reduction potential values that are closer to that of the protein loop that was introduced. (**E**) A wide range of reduction potentials is achieved in designed blue Cu proteins using the Az scaffold, and fine-tuning E_0 with several mutations. (**F**) Axial ligation of T1 Cu sites in native blue proteins; $T1_{PR}$ denotes the permanently reduced T1 Cu site in human ceruloplasmin. E_0 values compiled here were taken from [22, 60, 63, 75, 84, 85, 87, 89, 91, 93–94, 96, 102, 103, 105 112–119, 123–127, 130–132, 147–153, 155].

ligands lower it [90]. The trigonal core of three donors [2 N(His), S(Cys)] in the classic blue copper proteins, such as Az and Pc, is more amenable for Cu(I) and thus, it leads to a higher E_0 (250–370 mV) [91], as compared to the reduction potential of the Cu(II)/Cu(I) pair in water (~ +150 mV). Fine-tuning the reduction potential of T1 Cu sites is attributed to several factors, including the folding of the protein around the site [47, 92], hydrogen bonding networks involving the second coordination sphere [93], and the degree of interaction of the axial ligand [94–97]. The following sections describe each of these factors in more detail.

3.1. The 'Entatic' State

In general, it is assumed that the intermediates of (fast/efficient) ET reactions involving metal complexes require a geometry which is between the favored geometry of the two valence states. However, Cu(I) and Cu(II) display different geometry and coordination number preferences. While Cu(I) prefers lower numbers of ligands in geometries that include linear, trigonal, and tetrahedral, Cu(II) prefers higher coordination numbers with tetragonal or square planar geometries for four coordination sites, and squared based pyramidal or octahedral for five and six coordinate centers, respectively. T1 Cu sites are generally four coordinate in a distorted tetrahedral geometry, or trigonal three coordinate sites when there is no coordinating residue at the axial position. In all cases, they have three strong ligands arranged in a pseudo-trigonal manner: one Cys thiolate and two His imidazoles, while weak axial interactions typically involve a Met thioether or Gln amide. This unusual geometry of blue copper centers is reflected in their unique spectral features, as discussed above, and it is highly conserved upon redox cycling. Indeed, crystal structures of BCPs have demonstrated that the ligand positions are essentially unchanged upon reduction of the Cu(II) form to Cu(I) (Figure 7), and are practically identical to those in the apoprotein [98]. While the strong set of ligands [2 N(His), S(Cys)] arranged in a pseudo-trigonal manner may suffice the preferences of a Cu(I) ion, the Cu(II) ion is constrained by a rigid geometry that is distorted from its preferred tetragonal configuration, resulting in the 'entatic' state. Crystal structures of apo and metal-substituted BCPs have also supported the notion of the 'entatic' state: structures of Az [49, 99], Pc [14], PAz [12], and amicyanin (Ami) [13] show little difference in the position of the ligating residues (0.1–0.3 Å) between their apo and Cu-bound forms. The structural and electronic features of the T1 Cu site, as generated by the 'entatic' state, seem to be supported by the βαβ fold in the BCPs [23, 28].

The 'entatic' state occurs in a protein when a group, metal or non-metal, is forced into an unusual, energized geometric or electronic state [47]. In many proteins involved in biological ET, the redox-active metal ion displays a distorted geometry approximating that of the transition state in which it is involved. This results in a design in which the redox-active site itself would achieve a condition energetically favorable to catalysis. The T1 Cu sites are a prime example of a

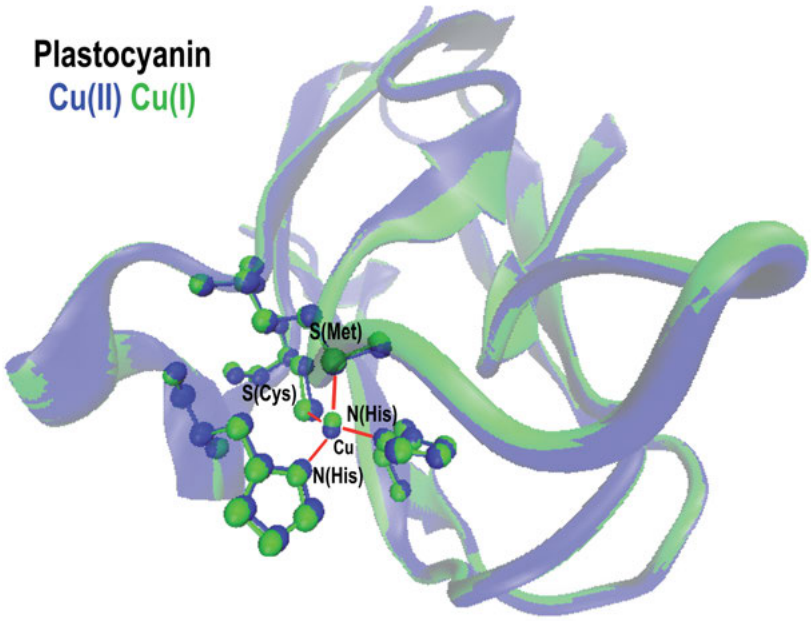


Figure 7. Superposition of the crystal structures of oxidized and reduced plastocyanin. A comparison of the coordination sphere of the T1 Cu center shows that the site suffers minimal changes upon oxidation. Figure generated from PDB ID 1BXU (blue) and 1BXV (green) (oxidized and reduced Pc, respectively) [166] coordinates using Visual Molecular Dynamics (VMD).

metal ion in an 'entatic' state that modulates ET, both through the constraint of its redox potential and of its relaxation energy upon redox state change. Indeed, the metal–ligand bonds in T1 Cu centers, as determined by X-ray crystallography and extended X-ray absorption fine structure spectroscopy, are lengthened by 0.1 Å or less between the two oxidation states. These small changes in bond length are crucial to assure the low reorganization energy of T1 Cu sites, and, thus, a fast ET for its function (Section 4). The small conformational change between the Cu(II) and Cu(I) states in the T1 Cu center can be explained in terms of the electronic structure of the T1 Cu(II) site. The unusually distorted tetrahedral geometry imposed by the short S(Cys)–Cu bond causes the loss of degeneracy for the $d_{x^2-y^2}$ and d_{xy} orbitals, eliminating the Jahn-Teller distorting force that is common to Cu(II) ions in a "normal" T_d ligand field [56]. The absence of the Jahn-Teller distortion in the oxidized state means that there is little reorganization of the ligand sphere upon redox change. The relative lower energy of the $3d_{x^2-y^2}$ orbital results in a significantly more positive potential than would be the case for the same ligand set in a tetragonal environment. Unconstrained Cu complexes with a 4N-ligand set have an $E_0 \sim -200$ mV, while a constrained distorted tetrahedral complex with tetrakis-imidazole coordination has

an $E_0 \sim +350$ mV [100], suggesting that the distorted tetrahedral geometry contributes $\sim +500$ mV to the reduction potential.

3.2. Role of the Axial Ligand

While a 2 N(His), S(Cys) ligand set is a unique feature of all T1 Cu centers, the axial ligand is not conserved [101]. In most BCPs, the axial position is occupied by a Met residue acting as a weak ligand with its thioether group at 2.6–3.15 Å from the metal ion [87]. However, some other blue proteins, like stellacyanin, have a Gln residue in this position at 2.2 Å, while others have a non-coordinating residue, such as Leu or Phe, as is the case for fungal laccases and one of the T1 sites in ceruloplasmin (Figure 6F) [85, 102]. A stronger bond with the axial ligand correlates with larger tetragonal distortions and a weaker Cu(II)–S(Cys) bond, as illustrated for the green sites (Section 2.3) [95, 103–105]. The T1 Cu center in Az represents a special case, having a fifth oxygen ligand at 3.0 Å, provided by a Gly carbonyl residue at the opposite position to the axial Met [106]. Overall, axial ligation in T1 Cu sites plays a very important role in determining the properties of each site [101, 103, 107, 108]. For example, the presence of an axial ligand protects the site against exogenous ligand coordination [109], it provides stability towards temperature and pH changes [110], and most importantly, it contributes to the fine-tuning of the reduction potential of these sites [95, 109].

The large range of reduction potentials found for BCPs is a result of evolutionary pressure to match the different biological functions of each protein. In general, T1 Cu sites with strong oxygen-based axial ligands, such as Gln, display lower values of E_0 ($+190$ to $+320$ mV), due to a stabilization of the Cu(II) state. This phenomenon is found in stellacyanin, where the Cu center has a Gln axial ligand, representing the strongest axial coordination and the lowest reduction potential (184 mV) among all T1 Cu sites [84, 111]. T1 Cu centers with a Met axial ligand show a wide range of reduction potentials, between 200 and 680 mV (Figure 6A). The strength of the axial ligand interaction is one of the main factors tuning the value of E_0: as the thioether-metal bond strength decreases, the Cu–S(Cys) bond interaction is stronger, resulting in increased stabilization of Cu(I) over the oxidized state, and thus, an increased reduction potential. The modulation of the axial ligand bond strength provides a first coordination sphere mechanism for the protein to tune the reduction potential by up to 200 mV [56, 101]. On the other hand, some T1 Cu sites have a fifth carbonyl axial ligand from a Gly residue that, being an oxygen-based ligand, stabilizes Cu(II) over the Cu(I) state. This is the case of the blue center in Az, located at the low end of Met-containing T1 Cu centers with a $E_0 \sim +300$ mV. As compared to Az, the carbonyl O-Cu interaction in Pc is much weaker, resulting in a somewhat higher potential ($+370$ mV) (Figure 6A). Consistently, a mutation in Az (Phe114Asn) that generates a hydrogen bond between the carbonyl oxygen of Gly and Asn114, also disrupts the interaction between the Cu ion and its fifth ligand, causing an increase of 122 mV in the reduction potential [112]. Finally, several blue Cu centers have non-coordinating hydrophobic residues in the axial position (Leu in fungal

laccase), where the complete lack of axial ligand coordination leads to a three-coordinate Cu site with a strong Cu-S(Cys) interaction, stabilizing the Cu(I) state. These effects, coupled with the hydrophobicity of the residue, results in the T1 Cu centers with the highest reduction potentials (+427 to +1000 mV) (Figure 6A) [113]. Figure 6B summarizes the role of the axial ligand in fine-tuning the reduction potential of T1 Cu sites: an increase in donor strength of the axial ligand is related to lower E_0 values, demonstrating that the chemical nature of the axial ligand has a great impact on the reduction potentials in the T1 Cu sites.

The fine mechanism by which the strength of the axial ligand fine-tunes the reduction potential of T1 Cu sites is also demonstrated by a comparison of Pc, CBP, and NiR (Section 2.3). These sites share the same ligand set [S(Cys), 2 N(His), S(Met)], yet their redox potentials vary from 370 mV in Pc, 317 mV in CBP to 247 mV in NiR (Figure 6A). As described in Section 2.3, the color of these proteins ranges from blue in Pc to green in NiR, reflecting changes in their geometric and electronic structure (Figure 4). Basically, the axial Met ligation becomes stronger from the blue Pc site to the green NiR center, with the concomitant decrease of the Cu–S(Cys) bond strength and destabilization of the Cu(I) reduced state, resulting in a decrease in reduction potential along this series [6, 7, 34].

Site-directed mutagenesis has been extensively used in native BCPs to evaluate the role of the axial ligand in their electronic structure and redox properties. Met121 in Az has been mutated to all possible 19 amino acids, finding that its replacement by hydrophobic residues without a donor atom (Ala, Val, Leu, Ile) stabilizes Cu(I) over Cu(II) and increases the E_0 by 60–140 mV [114, 115]. Conversely, Met substitutions by Glu [114, 115], Gln [105] or Cys [116] residues lower the reduction potential by 100–200 mV (Figure 6B). Interestingly, His, Lys and Asp mutations in the axial position do not generate a change in E_0 [114, 115]. Moreover, a series of iso-structural non-natural amino acids – including oxomethionine (OxM), difluoromethionine (DFM), trifluoromethionine (TFM), selenomethionine (SeM), and norleucine (Nle) – were introduced at the position of Met121 in Az [75]. These mutated T1 Cu sites exhibited little change in their UV-Vis electronic absorption and EPR spectra, while their reduction potentials span in a 225 mV range, correlating linearly with the hydrophobicity of the axial ligand (Figure 6B). Met121 in Az has also been replaced by homocysteine (Hcy), containing an extra methylene group. This substitution decreases the distance between the axial ligand and the Cu ion, resulting in a longer Cu–S(Cys) bond length (2.22 Å), a large parallel hyperfine splitting ($A_{||} = 180 \times 10^{-4}$ cm^{-1}), and electronic absorption features that are very similar to those of the red Cu center in nitrosocyanin [116]. The reduction potential of this T1 Cu variant is ~100 mV lower than that of wild-type Az, a trend that is consistent with having a stronger interaction with the axial ligand and a weaker Cu–S(Cys) bond (Figure 6B).

Another series of mutations that illustrates the role of the axial ligand is that of cucumber stellacyanin, where the axial Gln has been replaced by Met and Leu [63]. The redox potential of this site (260 mV) [37] increases to 420 [37] and 580 mV [63], upon the Gln → Met and Gln → Leu mutations, respectively. This

trend is consistent with that found in naturally occurring T1 Cu centers, as discussed above. Mainly, as the strength of the axial ligand decreases (in the series going from Gln to Met to non-coordinating Leu), the reduction potential increases. Consistently, the Met to Gln substitution in the blue Cu center of NiR causes a decrease in E_0 of 127 mV [117], while in Rc this mutation also decreases E_0 by 108 mV [94]. Moreover, mutation of a Met to a non-coordinating residue in the axial position increases the reduction potential by ~100 mV, as observed for the Met → Leu mutation in Rc, which leads to a very high E_0 value, ~ +800 mV at pH 3.2 [94], similar to those found in sites naturally having a Leu axial ligand, such as fungal laccase and one of the T1 Cu sites in ceruloplasmin [94]. Consistently, the Phe → Met mutation in fungal laccase results in a decrease of the reduction potential by ~100 mV (Figure 6B) [118].

In summary, the protein fold in BCPs provides different degrees of hydrophobicity and axial ligands of variable strength. As illustrated in the examples discussed here, the axial ligand modulates the geometric and electronic structure of the blue copper site to fine-tune the reduction potential of the T1 Cu center. Overall, as the donor strength of the axial ligand decreases, the strength of the Cu–S(Cys) bond increases with subsequent stabilization of the Cu(I) state and an increase in reduction potential. However, the wide range of reduction potentials observed within families of blue Cu proteins with the same axial ligation suggests that, in spite of the great impact of the axial ligand, there are other properties of the protein matrix that play important roles in tuning the reduction potentials of T1 Cu sites, as discussed below.

3.3. Inner versus Outer Coordination Sphere Effects

As discussed above, the reduction potential of T1 Cu centers is highly dependent on the metal coordination geometry and the donor atoms around the Cu ion. However, the inner coordination sphere cannot fully account for the wide range (~ 600 mV) of reduction potentials observed for blue copper proteins with a given fixed ligand set. The axial ligand only accounts for tuning the redox potential in a range of ~ 220 mV [96]. Thus, the nature of the chemical environment around the ligating residues, namely the outer coordination sphere, should contribute to the specific reduction potential of each protein. While the protein provides the 'entatic' state to maintain the copper ion in a geometry that is intermediate between the preferences of Cu(II) and Cu(I) states, there is a cost in energy to sustain such a constrained geometry. This is overcome by outer sphere interactions between the Cu ligands and surrounding residues [61]. In fact, the 'entatic' state in blue Cu proteins is induced mainly by an extensive network of hydrogen bonds [46]. The protein scaffold is not a passive entity, and variations in its hydrogen bonding network play an important role in fine tuning the electronic and physical properties of T1 Cu centers, such as the reduction potential [119].

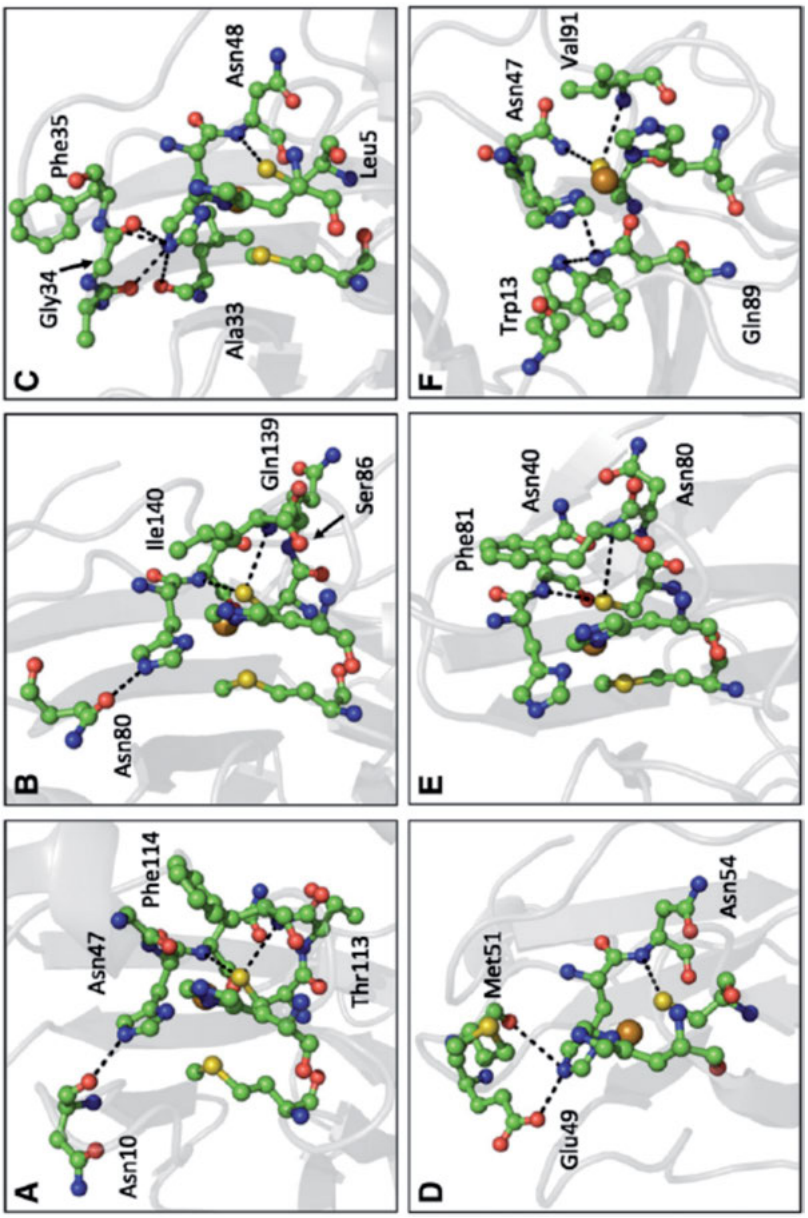


Figure 8. Outer-sphere coordination in representative blue copper proteins in their oxidized forms (**A**) *P. aeruginosa* azurin; (**B**) *T. ferrooxidans* rusticyanin; (**C**) *P. nigra* plastocyanin; (**D**) *P. denitrificans* amicyanin; (**E**) *C. sativus* cucumber basic protein; (**F**) *C. sativus* stellacyanin. Reproduced with permission from [167], copyright 2012, Elsevier.

The inner sphere ligands of T1 Cu sites display several hydrogen bonds, often conserved in several BCPs (Figure 8). For example, in Az (Figure 8A) one of the His ligands binds Cu through its N_δ, whereas N_ε has a hydrogen bond with the carbonyl oxygen of Phe15. This hydrogen bond is conserved in other blue Cu proteins, but other residues provide the proton acceptor atom: Glu49 in Ami (Figure 8D), Asn80 in Rc (Figure 8B), and a water molecule in phytocyanins. Also, in all blue copper proteins, the sulfur atom of the coordinating Cys forms one or two hydrogen bonds with the amide N-H groups from adjacent residues. These H bonds modulate the electron density of cysteine S, which is crucial for the highly covalent nature of the Cu−S(Cys) bond. In Az, Cys112 forms two hydrogen bonds with the backbone amide groups from Asn47 and Phe114 [29, 120]. Additionally, Asn47 forms another hydrogen bond with the oxygen from the Thr113 side chain, strengthening even more the two ligand loops (Figure 8A). A similar scenario is found in Rc, where the coordinating Cys138 forms hydrogen bonds with amide groups from Ser86 (equivalent to Asn47 in Az) and Ile140; while the carbonyl oxygen of Ser86 hydrogen-bonds with the amide from Gln139, strengthening the two ligand loops (Figure 8B) [93]. In contrast, Pc, PAz and Ami (Figure 8C, D) have only one hydrogen bond around the coordinating Cys residue, since the corresponding Phe residue (Phe114 in azurin) is a Pro in these proteins [48, 121, 122].

Site-directed mutagenesis studies have revealed important information about how the hydrogen bonding network around the blue Cu center tunes its reduction potential and other properties. Unlike mutations of the metal ligating residues, mutations in the outer coordination sphere are less likely to change the spectroscopic features of the T1 Cu site, yet allow to probe the role of hydrogen bonding networks modulating/influencing its redox potential. For example, the Ser86Asn mutation in Rc resulted in a decrease of the reduction potential of the T1 Cu center, and this was attributed to the strengthening of the H-bonding interactions between two ligand-containing loops (Figure 6C) [93]. Conversely, mutations of the corresponding Asn by Ser [112] or Leu [123] in Az, that eliminate the hydrogen bond between Asn47 and Thr113, result in an increase of E_0 by 130 and 110 mV, respectively. Furthermore, the additional hydrogen bond between Cys112 and Phe114 in Az was probed with the Phe114Pro mutation and it resulted in a decreased reduction potential of the blue Cu center. It is proposed that abolishing the hydrogen bond to the thiolate of Cys112 provides more conformational flexibility of the Cys residue (Figure 6C) [124]. Consistently, in PAz and Ami, where a Pro residue is in place of the corresponding Phe114 in Az, mutating the Pro resulted in an increase of the reduction potentials of these sites [119, 125, 126]. Additionally, placing a residue that can form H bonds (Ala and Ile) at Pro80, which is adjacent to the His81 ligand in PAz, alters the hydrogen-bonding network in the vicinity of the T1 center and results in a significant increase of the reduction potential: 139 mV in the Pro80Ala mutant [125] and 180 mV in the Pro80Ile mutant (Figure 6C) [127].

Overall, these studies demonstrate that the hydrogen-bonding network in the second coordination sphere fine-tunes the reduction potential of blue copper proteins. In fact, an additive effect of mutations on the outer coordination sphere

has been observed for Az, where the reduction potential of the T1 Cu center could be altered from +90 to +640 mV, a range that is much wider than that reported for native T1 copper proteins (Figure 6E) [112].

Beyond the hydrogen-bonding network, other outer-sphere factors can influence the redox potential. The hydrophobic nature of residues in the second coordination sphere of the metal center is also believed to play an important role in tuning the reduction potential of the T1 Cu site [91, 119, 128]. Hydrophobic encapsulation increases reduction potentials by exclusion of water and it encloses the positively charged metal center in a low dielectric protein medium. For example, the Cu site of Rc is surrounded by hydrophobic residues, namely Phe54, 51, 76, 83 and 111 and Ile140, that stabilize the Cu(I) state. The higher reduction potential of Rc, in relation to others blue copper proteins, has been ascribed to the effect of these hydrophobic residues [128, 129]. Additionally, the longer N-terminal region of Rc protects the hydrophobic patch of the protein, contributing to its high stability and elevated reduction potential at low pH (+680 mV at pH 3.2) [130, 131]. The role of a hydrophobic environment is further confirmed by the effect of incorporating four Phe residues in the positions Met13, Leu33, Met44, and Leu86, which are part of the secondary coordination sphere of the T1 Cu in Az. These mutations were inspired by the highly hydrophobic environment of the blue Cu in Rc; they yielded little change in the electronic absorption and EPR spectra of the site, but a modest increase in its reduction potential that correlated with the number of Phe incorporated and amounted to up to 30 mV (Figure 6E) [132].

Clearly, several factors provided by the protein scaffold lead to a fine-tuning of the reduction potential of T1 Cu centers. This knowledge allows us to control the E_0 in new biosynthetic blue copper centers, taking advantage of a variety of redox reactions that can be used for numerous applications, from alternative energy generation to synthesis of intermediates or final products for industrial applications.

3.4. Biosynthetic Blue Copper Centers

The knowledge of how metalloproteins work can help us to design new biosynthetic molecules that reproduce the structures and functions of native proteins. Biosynthetic metalloproteins can reveal hidden structural features that may be missing from studies of native metalloproteins, but also, the new metalloenzymes can be used for biotechnological applications. A full understanding of the factors that govern metal functions in proteins makes it possible to design new biosynthetic molecules with improved properties, such as higher stability and greater efficiency, or even new functions not found so far in nature. However, designing an inner and outer coordination sphere to obtain a new functional metalloprotein is a challenge. The design of small models containing the basic components of a native metalloprotein has been an approach that helps to identify the minimum features needed for function. Also, the rationally-designed systems are

often simpler than the native systems, allowing de-convolution of their properties and a deeper understanding of them.

Two complexes that model the key geometric and electronic features of the blue copper site have been developed by Kitajima, Fujisawa, and Moro-oka (Complex 1) [133, 134] and Holland and Tolman (Complex 2) [135]; in both cases a triphenylmethylthiolate is used to model the Cys ligand. Complex 1 is based on a tridentate tris(pyrazolyl)hydroborate $[HB(pz')_3]^{-1}$ ligand, while Complex 2 uses the bidentate β-diketiminate ligand. Complex 1 has a trigonally distorted tetrahedral structure, with a short and strong Cu(II)–S(thiolate) bond at 2.1 Å, and an axial N(pz′) ligand at 2.2 Å. The structure of Complex 1 is very similar to that of the copper center in Pc, having a long axial thioether (2.8 Å). In contrast, Complex 2 has a three-coordinate, trigonal planar Cu center, with a structure very similar to the one of three-coordinate blue copper sites in fungal laccases [113]. The availability of these models allowed a deeper understanding into the electronic structure of T1 Cu sites, particularly, with respect to the covalency of the Cu(II)–S(thiolate) bond, the nature of the equatorial nitrogen ligands and the role of the axial ligand.

Initial attempts to model the copper site in BCPs were done by replacement of Zn by Cu in Zn-binding proteins, such as horse liver alcohol dehydrogenase [136–139] or insulin [140–142]. Horse liver alcohol dehydrogenase contains a catalytic Zn(II) ion that is coordinated in a distorted tetrahedral geometry by 2S(Cys) 2N(His) O(H_2O). Replacement of the catalytic Zn(II) ion with Cu(II) results in a protein with an intense blue color due to a strong absorption maximum at 623 nm and a rhombic Cu(II) EPR spectrum [136, 137]. The properties of this Cu(II)-substituted protein suggested that the unique spectroscopic properties of the T1 Cu center can be replicated by Cys and His residues in a distorted tetrahedral environment. Another close mimic of the T1 Cu center was developed using insulin. The Zn(II)-insulin R_6 hexamer contains two identical zinc sites, each coordinated to three N(His) and one chloride or phenolate ion in a distorted tetrahedral geometry. Upon addition of thiolate ligands, such as pentafluorobenzenethiolate or tetrafluorobenzenethiolate, the Cu(II)-insulin R_6 hexamer exhibits a strong absorption around 630 nm and an axial EPR spectrum that is similar to those of blue Cu proteins [140, 142]. Interestingly, addition of benzenethiolate or 4-methylbenzenethiolate resulted in derivatives with rhombic EPR spectra that resemble those of green Cu centers.

A red Cu center has been engineered into the T2 Cu site of Cu-Zn superoxide dismutase (CuZnSOD) [143, 144]. CuZnSOD is a dimeric enzyme with two identical subunits, each of which contains a T2 Cu and a Zn site. The Cu(II) site is coordinated to four His residues and a water molecule in a distorted square pyramidal geometry, while the Zn(II) ion is ligated by three His residues and one Asp residue in a distorted tetrahedral geometry. The Cu(II) and Zn(II) ions are bridged by a His ligand. When one of the ligating His residues in the Cu site (His46 or His120) is substituted by a Cys residue, the new site displays EPR signals that are typical of T2 Cu centers, and new strong absorption maxima at 379 and 406 nm for His46Cys and His120Cys, respectively. On the other hand, the replacement of His80, which is a ligand for Zn(II), by a Cys residue, and

replacing the Zn(II) ion by Cu(II) in the His80Cys variant, leads to a CuCuSOD with a green T1 Cu site showing electronic absorption, MCD and resonance Raman spectra that are similar to those of the T1 Cu site in NiR [104], yet its EPR spectrum is similar to that of stellacyanin [143]. The strong Cu(II)–S(Cys) bonding interaction is the most fundamental structural component of the T1 Cu site and is one that is extremely difficult to mimic with synthetic systems. However, the constraints of the binding sites in metalloproteins can stabilize metal geometries that are not usually observed in small Cu(II) complexes. Thus, these proteins have provided useful model systems through which the spectroscopic features of the blue copper center may be systematically probed.

Another example of a designed protein with a blue Cu center is one where the protein scaffold is provided by four-stranded α-helical coiled-coils. In this case, the Cu(II) is coordinated by two N(His) and one S(Cys), adopting a trigonal planar geometry in the hydrophobic core of the scaffold. This *de novo* designed protein displayed a blue color, electronic absorption and EPR features that resemble those of a typical T1 Cu site, while the metal center showed a reduction potential of 328 mV [145].

In order to take advantage of the structures made by nature, site-directed mutagenesis has been extensively used to generate new blue Cu proteins with different combinations of the loops containing the metal ligands. The blue Cu center is placed between loop regions, involving three ligands on a loop linking β-strands 7 and 8 and the fourth coordinating residue on the loop linking β-strands 3 and 4 [146]. Although the blue copper proteins have a conserved inner coordination sphere, the ligand-containing loops from different proteins show variation in length and sequence. Loop-directed mutagenesis can be used to graft active sites onto the rigid β-barrel structure of blue Cu proteins that provides an ideal scaffold for protein engineering studies. The loops in PAz, Pc, and Az have been replaced by the shorter loop from Ami, resulting in chimeras named PAz-Ami, Pc-Ami, and Az-Ami variants, respectively [147, 148]. In all cases these mutations conserve the active site structure with limited effects on the spectroscopic properties of the T1 Cu center. However, loop contraction always results in a decrease of the reduction potential, ranging from ∼ 30 mV in Az-Ami to ∼60 mV in Pc-Ami and PAz-Ami at about neutral pH values (Figure 6D) [148]. Ami has the lowest E_0 value with respect to PAz, Pc, and Az; thus, the introduction of its C-terminal active site loop in PAz-Ami, Pc-Ami, and Az-Ami variants results in a reduction potential that is closer to that of the Ami protein. Conversely, when the longer active site loops of PAz, Pc, and Az were introduced into the Ami protein scaffold, resulting into chimeras named Ami-PAz, Ami-Pc, and Ami-Az [149–151], the new proteins displayed an increase in reduction potential, resulting in a E_0 value that is closer to that of the protein loop that was introduced into the Ami scaffold (Figure 6D). Thus, loop-contraction generally provides an active site environment preferable for the Cu(II) state, whereas loop elongation leads to an active site which favors the Cu(I) state.

The chimeras discussed above suggest that the chemical environment provided by each loop plays a determining role in the reduction potential of the blue Cu site. For example, in the Az-Pc variant, where the Pc loop is introduced into the

Az scaffold, the loop adopts a conformation identical to that in Pc and its E_0 increases to a value that matches the reduction potential of Pc [152]. Moreover, the Az-Nc mutant, where the loop of the red Cu site in nitrosocyanin is introduced into the Az scaffold, displays a red color and its electronic absorption and EPR spectra closely resemble those of Nc, despite the fact that the axial ligand is a His instead of a Glu. The reduction potential of the Az-Nc chimeric protein lies between the reduction potentials of Az and Nc (~ 50 mV below that of Az), probably reflecting the effect of having different axial ligands [153]. Thus, inserting the Nc loop into the Az scaffold is enough to provide the chimeric T1 Cu center with the red Cu spectroscopic features of Nc, while reducing its reduction potential (Figure 6D). Other mutations in the loop of Az include its replacement by a range of sequences containing Ala, Gly, and Val residues, in order to assess the importance of amino acid composition and length of this region. Variations in the nature of residues, while maintaining the loop length, preserve the structural features of the loop, suggesting that loop structure is dictated by length and not sequence [154].

Finally, the knowledge gained about the properties that fine-tune the reduction potential of the T1 Cu site within the same protein scaffold, has allowed the design of new proteins that cover a wide range of E_0 values. For example, changes in hydrophobicity and hydrogen-bonding networks around the T1 Cu site in Az were caused by combinations of mutations that include Phe114Pro, Phe114Asn, Met121Gln, Met121Leu, and Ans47Ser, yielding blue sites where the reduction potential could be tuned over a 700 mV range, with a gradual variation of approximately 50 mV between the individual mutants [112]. Another study used the Az scaffold and different combinations of five mutated residues and two metal ions (Cu and Ni) to tune the reduction potential from $+970 \pm 20$ mV to -954 ± 50 mV. The mutations included: Asn47Ser and Phe114Asn, that tune the hydrogen-bonding networks near the Cu center; Met121Leu, that adds a hydrophobic residue in the axial position; and Met44Phe and Gly116Phe, that increase the hydrophobic nature of the surrounding environment (Figure 6E) [155]. The ability to design blue Cu centers covering a wide range of E_0 values clearly documents the understanding that has been achieved about the structural features that tune the reduction potentials in metalloproteins.

The reduction potential is a critical parameter in determining the efficiency of most biological and chemical reactions. Extensive experimental and theoretical work has been devoted to the understanding of the molecular factors responsible for fine-tuning the reduction potential of metal active sites. This knowledge may enable scientists and engineers to design new redox agents for chemical, biochemical, biophysical, and biotechnological applications.

4. TYPE 1 COPPER REACTIVITY AND ELECTRON TRANSFER PATHWAYS

Blue Cu centers act as ET sites in a wide variety of electron transport chains. As discussed above, while the electronic and geometric structure of T1 Cu cen-

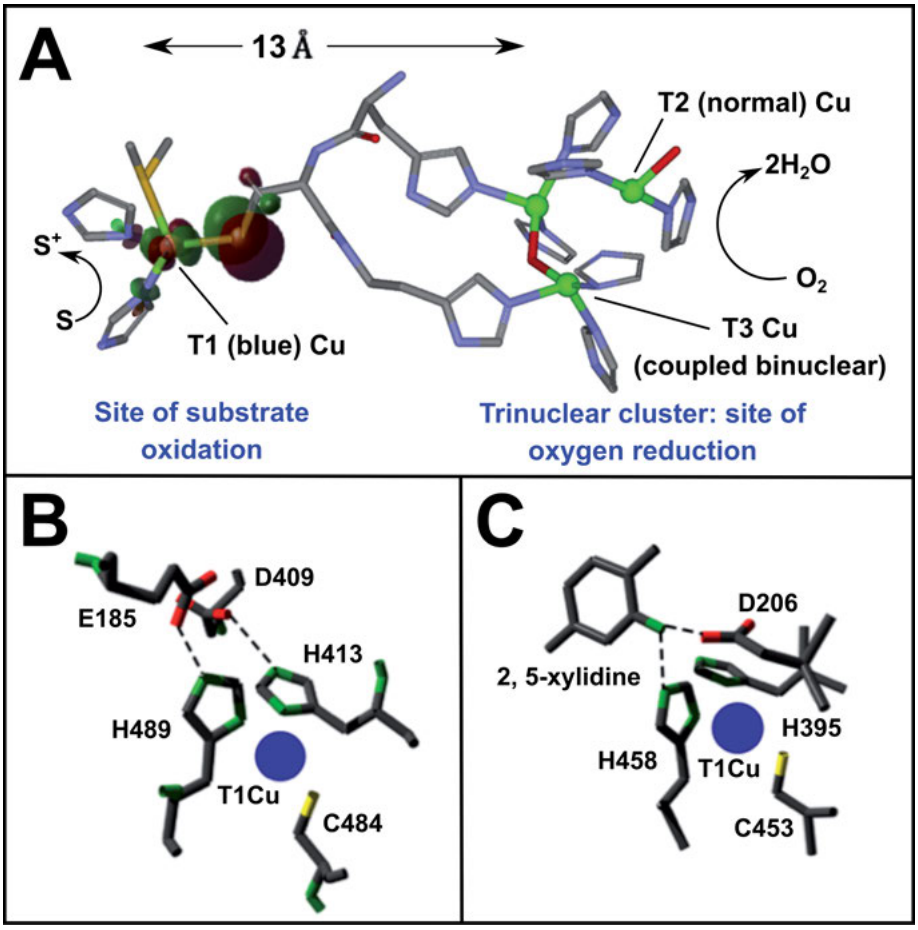


Figure 9. (**A**) Active site of the multicopper oxidases showing the T1, T2, and T3 Cu sites. Electron transfer pathways from the substrate to the T1 Cu sites in the ferroxidase Fet3p (**B**) and the fungal laccase TvLc (**C**). In Fet3p (B), the His ligands to the T1 Cu (His489 and His413) are H-bonded to E185 and D409, while in TvLc (C) the 2,5-xylidine H-bonds to D206 and H458, a ligand to the T1 Cu site. Adapted with permission from [88], copyright 2007, American Chemical Society.

ters are conserved across different biological systems, the reduction potentials are tuned by the protein scaffold to be consistent with the physiologically relevant electron donors and acceptors. However, in many cases the physiological redox partners of BCPs have not been identified, representing a challenge for future research in the biology of these proteins. Examples of BCPs where their redox partners are known are Pc and Ami. Pc accepts electrons from cytochrome *f* and transfers them to $P700^+$ in photosystem I [15, 16, 19, 20], while Ami accepts electrons from methylamine dehydrogenase to transfer them to a cytochrome *c* [156, 157]. In both cases, protein-protein complexes are formed to

achieve ET [42]. In contrast, some blue Cu centers have their electron acceptors within the same protein, as is the case in a family of enzymes named multicopper oxidases (MCO), where the T1 Cu accepts electrons from a wide variety of small substrates, but it always transfers them to the same acceptor: a trinuclear Cu center located 13 Å away in the same protein, where O_2 is reduced to water (Figure 9A) [4, 53, 158]. The reactivity of T1 Cu sites in MCOs is quite diverse, as these enzymes are present in several organisms and are important for a variety of physiological functions, including: lignin formation in plants, lignin degradation in fungi, iron metabolism in yeast and mammals, copper homeostasis in bacteria, and manganese oxidation in bacterial spores. Thus, the electron donors for the T1 Cu center in MCOs vary from organic substrates to metal ions [88]. The diverse reactivity displayed by BCPs brings up several important questions: How can a T1 Cu center sustain electron transfer from/to such a wide range of electron transfer partners? What are the factors that determine the reactivity of blue Cu proteins? And what makes these sites so efficient for electron transfer? Marcus theory for ET can help us to answer these questions. In the following sections, the contributions of the electronic structure and geometric properties of T1 Cu centers, as well as the important properties provided by the protein, that tune these sites for biological ET are discussed.

4.1. Anisotropic Covalency and Electron Transfer Pathways

The active site of MCOs provides an excellent example of how the anisotropic covalency of the T1 Cu center contributes to rapid ET. In these enzymes, the Cys ligand at the blue Cu center is flanked by two His residues in the sequence, forming a highly conserved triad HisCysHis, where the His imidazoles are ligands to the type 3 site in the trinuclear cluster (Figure 9A). Rapid ($>10^{-3}$ s^{-1}) intramolecular ET from the T1 Cu center over 13 Å to the trinuclear cluster in MCOs has been analyzed in terms of the rate of non-adiabatic ET (Equation 1) [7]:

$$k_{ET} = (4\pi/h^2\lambda kT)^{1/2}(H_{AB})^2\exp[-(\Delta G^0+\lambda)^2/4\lambda kT] \quad (1)$$

where H_{AB} is the electronic coupling matrix element, ΔG^0 is the driving force, and λ is the reorganization energy [159]. The driving force ΔG^0 is determined by the difference in reduction potentials of the T1 Cu site and the trinuclear center, which are tuned for the electron flow to go in that direction. In Section 3, the factors involved in the fine-tuning of the reduction potential of blue Cu sites were discussed. The reorganization energy λ refers to the energetic cost of the geometric changes that the site suffers upon redox cycling; in Section 3.1, the role of the protein in imposing an 'entatic' state that minimizes this term for the T1 site was discussed. While ΔG^0 and λ are important for ET, it was found in the MCOs that the main contribution comes from the H_{AB} term, which refers to the electron transfer pathways that involve the connectivity between the Cys ligand to the T1 Cu center and the two His ligands to the trinuclear site. The H_{AB} term depends linearly on the covalency of the Cu–S(Cys) bond. The high

covalency of the Cu–S(Cys) bond and its relative orientation in the direction of the two Cys-His pathways provide a superexchange pathway for rapid ET (Figure 9A) [7]. Thus, the anisotropic covalency of the T1 Cu center primes this site to be ready to pass on the electron effectively towards the trinuclear Cu cluster. It should be mentioned that this is not the only case where anisotropic covalency contributes to the efficiency of an electron transfer Cu center. In the case of the binuclear Cu_A site, it has been shown that the anisotropic covalency provided by a Cu–S(Cys) bond contributes to both, intermolecular and intramolecular ET in cytochrome *c* oxidase [160].

4.2. Tuning the Type 1 Copper Reactivity for Biological Function

Let us now consider intermolecular electron transfer from an electron donor or substrate that has to form a complex with the blue Cu protein; this would be a protein-protein complex for Pc or Ami, or a small substrate binding to the protein in MCOs. The semiclassical expression for the rate of *inter*molecular ET, k_{ET}, is given by Equation (2) [159]:

$$k_{ET} = SK_A(4\pi^3/h^2\lambda kT)^{1/2}(H_{AB})^2\exp[-(\Delta G^0+\lambda)^2/4\lambda kT] \quad (2)$$

k_{ET} is now not only dependent on the electronic coupling matrix element (H_{AB}), the driving force (ΔG^0), and the reorganization energy (λ); but also on the steric term (S), and the equilibrium constant for formation of the interaction complex (K_A). Together, the S and K_A terms relate to the formation of a substrate-protein or protein-protein complex and the binding affinity associated to this interaction. Also, in this case the H_{AB} term is given by the electron transfer pathways that are provided by the interface between the two molecules upon formation of the interaction complex. For the case of protein-protein complexes, quantitative analysis of the rate of intermolecular ET is limited by the availability of structural information, and this is needed to assess potential electron transfer pathways. Examples of structurally characterized protein-protein complexes include: Pc with cytochrome *f* [161], Az with aromatic amine dehydrogenase (AADH) [162], and a ternary complex of Ami with its two redox partners, the electron donor methylamine dehydrogenase (MADH) and the electron acceptor cytochrome c_{551i} [18]. A comparison of the protein complexes of Az and Ami with their respective electron donor partners illustrates how, while having electron transfer pathways with different efficiencies, the protein scaffold compensates with other terms of Equation (2), to make electron transfer rates optimal in both systems [162]. Namely, while the MADH-Ami complex provides an electron transfer pathway that is ~ 100 times faster than that of the AADH-Az complex, the Az blue Cu site compensates this by having a lower reorganization energy and a larger driving force for ET, as compared to the T1 Cu site in Ami.

MCOs also help illustrate the contributions of the protein environment around the T1 Cu center to intermolecular electron transfer from the substrate. MCOs

can oxidize a wide variety of substrates that range from organic molecules to metal ions. As discussed in Section 3, the nature of the axial ligand and the outer coordination sphere contribute to fine-tuning the reduction potential of the blue Cu center. Not surprisingly, the reduction potentials of T1 Cu sites in MCOs range from 300–400 mV for sites with a Met axial ligand, as in ascorbate oxidase, tree laccase and the two redox-active T1 sites in human ceruloplasmin, to high potentials (from 550 to >1000 mV) in fungal laccases and the non-redox active T1 site in ceruloplasmin, which lack an axial ligand [88].

An interesting case is the T1 Cu site in the yeast ferroxidase Fet3p, having the lowest reduction potential (427 mV) among T1 sites with no axial ligand. Strikingly, with such a low reduction potential, the driving force (ΔG^0) to oxidize Fe(II) to Fe(III) would be very small, since the E_0 of the Fe(II)/Fe(III) redox couple at pH 6.5 is 420 mV. However, Fet3p has an Fe(II) binding site that tunes down the reduction potential of the substrate to make the ferroxidase reaction possible; it is estimated that the protein-bound Fe(II) must have an $E_0 \leq 190$ mV [163]. The Fe(II) binding site in Fet3p is provided by carboxylate residues in the vicinity of the T1 Cu site, namely Glu185 and Asp409, which are connected through hydrogen bonds to the two His ligands of the blue Cu center (Figure 9B) [164]. In this way, by having a binding site for its preferred substrate, Fet3p optimizes three terms of Equation (2): (i) it provides a high affinity binding site for Fe(II), optimizing the SK_A term; (ii) it tunes down the reduction potential of the substrate to provide a large ΔG^0; (iii) and the ligands for Fe(II) provide efficient electron transfer pathways from the substrate to the T1 Cu center, optimizing the H_{AB} term [88, 163]. The multicopper ferroxidases, such as yeast Fet3p and human ceruloplasmin, are an example of how the protein environment around the T1 Cu site tunes their reactivity and optimizes it to achieve high specificity towards Fe(II) ions, their physiologically relevant substrate. Providing a binding site for a metal ion that maximizes each contributing term to intermolecular ET (Eq. 2) from the metal ion substrate to the T1 Cu site may be a general strategy in multicopper metallo-oxidases.

Finally, a comparison of Fet3p with a fungal laccase, *Trametes versicolor* laccase (TvLc), provides further insights into how the protein fold around the T1 Cu site provides the basis for substrate specificity [88]. The physiological substrates of fungal laccases include lignin and pigments, and *in vitro* these enzymes are very efficient in the oxidation of organic compounds, such as phenols and hydroquinones. In contrast, Fet3p displays a very rapid reduction of the T1 Cu center by Fe(II), while being quite slow when confronted with organic substrates. The overall protein folds of TvLc and Fet3p are quite similar, and in both proteins one of the His ligands to T1Cu is solvent-exposed. The crystal structure of the complex of TvLc with 2,5-xylidine reveals the formation of hydrogen bonds between the substrate and His458, a ligand for the blue Cu site; and with Asp206 in the vicinity of the metal center (Figure 9C) [165]. This connectivity provides efficient electron transfer pathways from the organic substrate to the T1 Cu center, optimizing the H_{AB} term in Equation (2). Moreover, the presence of hydrophobic residues in the vicinity of the T1 Cu site favors the binding of aromatic substrates, optimizing the SK_A term for this type of substrates. In con-

trast, Fet3p lacks such hydrophobic residues, having the carboxylate residues that provide a binding site for Fe(II) instead, as discussed above. Thus, while Fet3p and fungal laccases are homologous and display almost identical protein folds, the chemical environment around the T1 site in each of them optimizes the interaction and intermolecular ET with their respective physiological substrates [88].

5. CONCLUDING REMARKS

Blue Cu proteins have evolved to fulfill an important function in biology as electron transfer sites that interact with a wide variety of redox partners. The cupredoxin protein scaffold plays a major role in tuning the distinctive structural and reactivity properties of T1 Cu sites for the physiological function of each blue Cu protein. The distorted tetrahedral geometry is imposed by the protein structure (entatic state) and the unusually high covalency of the Cu–S(Cys) bond provides the blue Cu site with unique features that make it optimal for ET function: small geometric changes upon redox cycling and anisotropic covalency to direct and make ET more efficient. While the nature of the axial ligand in the first coordination sphere of the T1 Cu site has a direct impact on its spectroscopic features and reduction potential, other factors in the outer coordination sphere, such as hydrogen-bonding networks and the hydrophobicity of the residues around the T1 site, also play a role in tuning its reduction potential and assuring a large driving force for ET.

Finally, the protein also provides efficient electron transfer pathways upon formation of complexes between the blue Cu protein and its redox partners. Future research efforts towards the identification of physiologically relevant electron donors and acceptors of blue Cu proteins and the structural characterization of interaction complexes shall provide further insights on how these proteins are optimized for biological electron transfer.

ACKNOWLEDGMENTS

Liliana Quintanar thanks Prof. Edward I. Solomon (Stanford University) for introducing her to the interesting research area of bio-inorganic spectroscopy of copper metalloenzymes, and for his important scientific influence, which is clearly evidenced in this chapter. Nils Schuth acknowledges a postdoctoral fellowship from CONACYT (National Council of Technology in Mexico) grant PN2076.

ABBREVIATIONS AND DEFINITIONS

AADH	aromatic amine dehydrogenase
Ami	amicyanin

Az	azurin
BCPs	blue copper proteins
CBP	cucumber basic protein
CD	circular dichroism
CT	charge Transfer
CuZnSOD	copper-zinc superoxide dismutase
DFM	difluoromethionine
DFT	density functional theory
ΔG^0	driving force
E_0	reduction potential
EPR	electron paramagnetic resonance
ET	electron transfer
H_{AB}	electronic coupling matrix element
HCy	homocysteine
HOMO	highest occupied molecular orbital
λ	reorganization energy
K_A	equilibrium constant for formation of the interaction complex
LF	ligand field
LMCT	ligand to metal charge transfer
MADH	methylamine dehydrogenase
MCD	magnetic circular dichroism
MCOs	multicopper oxidases
Nc	nitrosocyanin
NiR	nitrite reductase
Nle	norleucine
OxM	oxomethionine
PAz	pseudoazurin
Pc	plastocyanin
Rc	rusticyanin
S	steric term
Sec	selenocysteine
SeM	selenomethionine
SHE	standard hydrogen electrode
TFM	trifluoromethionine
T1	type 1
XAS	X-ray absorption spectroscopy

REFERENCES

1. C. L. Dupont, A. Butcher, R. E. Valas, P. E. Bourne, G. Caetano-Anolles, *Proc. Natl. Acad. Sci. USA* **2010**, *107*, 10567–10572.
2. C. L. Dupont, G. Grass, C. Rensing, *Metallomics* **2011**, *3*, 1109–1118.
3. R. A. Holwerda, S. Wherland, H. B. Gray, *Annu. Rev. Biophys. Bioeng.* **1976**, *5*, 363–396.
4. E. I. Solomon, D. E. Heppner, E. M. Johnston, J. W. Ginsbach, J. Cirera, M. Qayyum, M. T. Kieber-Emmons, C. H. Kjaergaard, R. G. Hadt, L. Tian, *Chem. Rev.* **2014**, *114*, 3659–3853.

5. E. I. Solomon, *Inorg. Chem.* **2006**, *45*, 8012–8025.
6. E. I. Solomon, R. G. Hadt, *Coord. Chem. Rev.* **2011**, *255*, 774–789.
7. E. I. Solomon, R. G. Hadt, B. E. R. Snyder, *Isr. J. Chem.* **2016**, *56*, 649–659.
8. S. Katoh, *Nature* **1960**, *186*, 533–534.
9. I. W. Sutherland, J. F. Wilkinson, *Microbiology* **1963**, *30*, 105–112.
10. W. E. Blumberg, J. Peisach, *Biochim. Biophys. Acta* **1966**, *126*, 269–273.
11. A. M. Nersissian, E. L. Shipp, in *Advances in Protein Chemistry: Copper-Containing Proteins,* Vol. 60, Eds. J. Selverstone Valentine, E. Butler Gralla, Academic Press, San Diego, CA, **2002**, pp. 271–340.
12. K. Petratos, M. Papadovasilaki, Z. Dauter, *FEBS Lett.* **1995**, *368*, 432–434.
13. R. Durley, L. Chen, L. W. Lim, F. S. Mathews, V. L. Davidson, *Protein Sci.* **1993**, *2*, 739–752.
14. T. P. J. Garrett, D. J. Clingeleffer, J. M. Guss, S. J. Rogers, H. C. Freeman, *J. Biol. Chem.* **1984**, *259*, 2822–2825.
15. M. Hervas, J. A. Navarro, A. Dıaz, M. A. De la Rosa, *Biochemistry* **1996**, *35*, 2693–2698.
16. P. M. Wood, D. S. Bendall, *Biochim. Biophys. Acta* **1975**, *387*, 115–128.
17. W. Lockau, *Arch. Microbiol.* **1981**, *128*, 336–340.
18. L. Chen, R. C. Durley, F. S. Mathews, V. L. Davidson, *Science* **1994**, *264*, 86–90.
19. W. Haehnel, T. Jansen, K. Gause, R. B. Klosgen, B. Stahl, D. Michl, B. Huvermann, M. Karas, R. G. Herrmann, *EMBO J.* **1994**, *13*, 1028–1038.
20. P. B. Crowley, M. Ubbink, *Acc. Chem. Res.* **2003**, *36*, 723–730.
21. L. G. Rydén, L. T. Hunt, *J. Mol. Evol.* **1993**, *36*, 41–66.
22. A. M. Nersissian, C. Immoos, M. G. Hill, P. J. Hart, G. Williams, R. G. Herrmann, J. S. Valentine, *Protein Sci.* **1998**, *7*, 1915–1929.
23. F. De Rienzo, R. R. Gabdoulline, M. C. Menziani, R. C. Wade, *Protein Sci.* **2000**, *9*, 1439–1454.
24. L. D. Kanbi, S. Antonyuk, M. A. Hough, J. F. Hall, F. E. Dodd, S. S. Hasnain, *J. Mol. Biol.* **2002**, *320*, 263–275.
25. M. E. P. Murphy, P. F. Lindley, E. T. Adman, *Protein Sci.* **1997**, *6*, 761–770.
26. T. Yamanaka, Y. Fukumori, *FEMS Microbiol. Rev.* **1995**, *17*, 401–413.
27. I. Harvey, Q. Hao, E. M. H. Duke, W. J. Ingledew, S. S. Hasnain, *Acta Crystallogr. D* **1998**, *54*, 629–635.
28. A. G. Sykes, *Adv. Inorg. Chem.* **1991**, *36*, 377–408.
29. E. T. Adman, *Adv. Protein Chem.* **1991**, *42*, 145–197.
30. L. F. Green, J. F. McCarthy, C. G. King, *J. Biol. Chem.* **1939**, *128*, 447–453.
31. S. Katoh, A. Takamiya, *Nature* **1961**, *189*, 665–666.
32. R. P. Ambler, L. H. Brown, *Biochem. J.* **1967**, *104*, 784–825.
33. F. E. Dodd, Z. H. L. Abraham, R. R. Eady, S. S. Hasnain, *Acta Crystallogr. D* **2000**, *56*, 690–696.
34. M. Murata, G. S. Begg, F. Lambrou, B. Leslie, R. J. Simpson, H. C. Freeman, F. J. Morgan, *Proc. Natl. Acad. Sci. USA* **1982**, *79*, 6434–6437.
35. K. A. Markossian, V. T. Aikazyan, R. M. Nalbandyan, *Biochim. Biophys. Acta* **1974**, *359*, 47–54.
36. V. T. Aikazyan, R. M. Nalbandyan, *FEBS Lett.* **1975**, *55*, 272–274.
37. A. M. Nersissian, C. Immoos, M. G. Hill, P. J. Hart, G. Williams, R. G. Herrmann, J. S. Valentine, *Protein Sci.* **1998**, *7*, 1915–1929.
38. L. Xu, X. J. Wang, T. Wang, L. B. Li, *Biol. Plant.* **2017**, *61*, 445–452.
39. R. J. Cho, M. Mindrinos, D. R. Richards, R. J. Sapolsky, M. Anderson, E. Drenkard, J. Dewdney, T. L. Reuber, M. Stammers, N. Federspiel, A. Theologis, W.-H. Yang, E. Hubbell, M. Au, E. Y. Chung, D. Lashkari, B. Lemieux, C. Dean, R. J. Lipshutz, F. M. Ausubel, R. W. Davis, P. J. Oefner, *Nature Genet.* **1999**, *23*, 203.

40. H. C. Freeman, J. M. Guss, in *Handbook of Metalloproteins*, Ed. A. Messerschmidt, John Wiley & Sons, Ltd., Hoboken, New Jersey, **2006**, pp. 1–17.
41. M. Hervas, J. A. Navarro, A. Diaz, H. Bottin, M. A. De la Rosa, *Biochemistry* **1995**, *34*, 11321–11326.
42. R. Hulsker, M. V. Baranova, G. S. Bullerjahn, M. Ubbink, *J. Am. Chem. Soc.* **2008**, *130*, 1985–1991.
43. P. M. Wood, *Biochim. Biophys. Acta* **1974**, *357*, 370–379.
44. M. D. Esposti, M. Mentel, W. Martin, F. L. Sousa, *Front. Microbiol.* **2019**, *10*, 499.
45. E. I. Solomon, J. W. Hare, H. B. Gray, *Proc. Natl. Acad. Sci. USA* **1976**, *73*, 1389–1393.
46. B. G. Malmström, *Eur. J. Biochem.* **1994**, *223*, 711–718.
47. R. J. P. Williams, *Eur. J. Biochem.* **1995**, *234*, 363–381.
48. J. M. Guss, H. C. Freeman, *J. Mol. Biol.* **1983**, *169*, 521–563.
49. W. E. B. Shepard, R. L. Kingston, B. F. Anderson, E. N. Baker, *Acta Crystallogr.* **1993**, *D49*, 331–343.
50. R. A. Scott, J. E. Hahn, S. Doniach, H. C. Freeman, K. O. Hodgson, *J. Am. Chem. Soc.* **1982**, *104*, 5364–5369.
51. S. E. Shadle, J. E. Penner-Hahn, H. J. Schugar, B. Hedman, K. O. Hodgson, E. I. Solomon, *J. Am. Chem. Soc.* **1993**, *115*, 767–776.
52. K. W. Penfield, A. A. Gewirth, E. I. Solomon, *J. Am. Chem. Soc.* **1985**, *107*, 4519–4529.
53. E. I. Solomon, R. K. Szilagyi, S. D. George, L. Basumallick, *Chem. Rev.* **2004**, *104*, 419–458.
54. J. L. Hughey-IV, T. G. Fawcett, S. M. Rudich, R. A. Lalancette, J. A. Potenza, H. J. Schugar, *J. Am. Chem. Soc.* **1979**, *101*, 2617–2623.
55. S. J. George, M. D. Lowery, E. I. Solomon, S. P. Cramer, *J. Am. Chem. Soc.* **1993**, *115*, 2968–2969.
56. E. I. Solomon, K. W. Penfield, A. A. Gewirth, M. D. Lowery, S. E. Shadle, J. A. Guckert, L. B. LaCroix, *Inorg. Chim. Acta* **1996**, *243*, 67–68.
57. R. K. Szilagyi, E. I. Solomon, *Curr. Opin. Chem. Biol.* **2002**, *6*, 250–258.
58. K. Olesen, A. Veselov, Y. Zhao, Y. Wang, B. Danner, C. P. Scholes, J. P. Shapleigh, *Biochemistry* **1998**, *37*, 6086–6094.
59. A. Veselov, K. Olesen, A. Sienkiewicz, J. P. Shapleigh, C. P. Scholes, *Biochemistry* **1998**, *37*, 6095–6105.
60. L. Basumallick, R. K. Szilagyi, Y. Zhao, J. P. Shapleigh, C. P. Scholes, E. I. Solomon, *J. Am. Chem. Soc.* **2003**, *125*, 14784–14792.
61. S. Ghosh, X. Xie, A. Dey, Y. Sun, C. P. Scholes, E. I. Solomon, *Proc. Natl. Acad. Sci. USA* **2009**, *106*, 4969–4974.
62. P. J. Hart, A. M. Nersissian, R. G. Herrmann, R. M. Nalbandyan, J. S. Valentine, D. Eisenberg, *Protein Sci.* **1996**, *5*, 2175–2183.
63. S. D. George, L. Basumallick, R. K. Szilagyi, D. W. Randall, M. G. Hill, A. M. Nersissian, J. S. Valentine, B. Hedman, K. O. Hodgson, E. I. Solomon, *J. Am. Chem. Soc.* **2003**, *125*, 11314–11328.
64. D. M. Dooley, J. Rawlings, J. H. Dawson, P. J. Stephens, L.-E. Andreasson, B. G. Malmstrom, H. B. Gray, *J. Am. Chem. Soc.* **1979**, *101*, 5038–5046.
65. U. A. Germann, G. Muller, P. E. Hunziker, K. Lerch, *J. Biol. Chem.* **1988**, *263*, 885–896.
66. A. Messerschmidt, R. Huber, *Eur. J. Biochem.* **1990**, *187*, 341–352.
67. E. I. Solomon, U. M. Sundaram, T. E. Machonkin, *Chem. Rev.* **1996**, *96*, 2563–2605.
68. A. A. Gewirth, E. I. Solomon, *J. Am. Chem. Soc.* **1988**, *110*, 3811–3819.
69. L. B. LaCroix, S. E. Shadle, Y. Wang, B. A. Averill, B. Hedman, K. O. Hodgson, E. I. Solomon, *J. Am. Chem. Soc.* **1996**, *118*, 7755–7768.

70. A. E. Palmer, D. W. Randall, F. Xu, E. I. Solomon, *J. Am. Chem. Soc.* **1999**, *121*, 7138–7149.
71. K. M. Lancaster, S. D. George, K. Yokoyama, J. H. Richards, H. B. Gray, *Nature Chem.* **2009**, *1*, 711–715.
72. K. M. Lancaster, K. Yokoyama, J. H. Richards, J. R. Winkler, H. B. Gray, *Inorg. Chem.* **2009**, *48*, 1278–1280.
73. K. M. Lancaster, S. Sproules, J. H. Palmer, J. H. Richards, H. B. Gray, *J. Am. Chem. Soc.* **2010**, *132*, 14590–14595.
74. K. M. Lancaster, O. Farver, S. Wherland, E. J. Crane, J. H. Richards, I. Pecht, H. B. Gray, *J. Am. Chem. Soc.* **2011**, *133*, 4865–4873.
75. S. M. Berry, M. Ralle, D. W. Low, N. J. Blackburn, Y. Lu, *J. Am. Chem. Soc.* **2003**, *125*, 8760–8768.
76. M. Ralle, S. M. Berry, M. J. Nilges, M. D. Gieselman, W. A. v. d. Donk, Y. Lu, N. J. Blackburn, *J. Am. Chem. Soc.* **2004**, *126*, 7244–7256.
77. R. L. Lieberman, D. M. Arciero, A. B. Hooper, A. C. Rosenzweig, *Biochemistry* **2001**, *40*, 5674–5681.
78. D. M. Arciero, B. S. Pierce, M. P. Hendrich, A. B. Hooper, *Biochemistry* **2002**, *41*, 1703–1709.
79. P. Hosseinzadeh, S. Tian, N. M. Marshall, J. Hemp, T. Mullen, M. J. Nilges, Y.-G. Gao, H. Robinson, D. A. Stahl, R. B. Gennis, Y. Lu, *J. Am. Chem. Soc.* **2016**, *138*, 6324–6327.
80. T. d. Blaauwen, C. W. G. Hoitink, G. W. Canters, J. Han, T. M. Loehr, J. Sanders-Loehr, *Biochemistry* **1993**, *32*, 12455–12464.
81. L. J. C. Jeuken, P. v. Vliet, M. P. Verbeet, R. Camba, J. P. McEvoy, F. A. Armstrong, G. W. Canters, *J. Am. Chem. Soc.* **2000**, *122*, 12186–12194.
82. L. J. C. Jeuken, M. Ubbink, J. H. Bitter, P. v. Vliet, W. Meyer-Klaucke, G. W. Canters, *J. Mol. Biol.* **2000**, *299*, 737–755.
83. S. Alagaratnam, N. J. Meeuwenoord, J. A. Navarro, M. Hervas, M. A. De la Rosa, M. Hoffmann, O. Einsle, M. Ubbink, G. W. Canters, *FEBS J.* **2011**, *278*, 1506–1521.
84. N. Sailasuta, F. C. Anson, H. B. Gray, *J. Am. Chem. Soc.* **1979**, *101*, 455–458.
85. T. E. Machonkin, H. H. Zhang, B. Hedman, K. O. Hodgson, E. I. Solomon, *Biochemistry,* **1998**, *37*, 9570–9578.
86. H. Claus, *Micron,* **2004**, *35*, 93–96.
87. J. Liu, S. Chakraborty, P. Hosseinzadeh, Y. Yu, S. Tian, I. Petrik, A. Bhagi, Y. Lu, *Chem. Rev.* **2014**, *114*, 4366–4469.
88. L. Quintanar, C. Stoj, A. B. Taylor, P. J. Hart, D. J. Kosman, E. I. Solomon, *Acc. Chem. Res.* **2007**, *40*, 445–452.
89. J. D. King, C. L. McIntosh, C. M. Halsey, B. M. Lada, D. M. Niedzwiedzki, J. W. Cooley, R. E. Blankenship, *Biochemistry* **2013**, *52*, 8267–8275.
90. B. R. James, R. J. P. Williams, *J. Chem. Soc.* **1961**, 2007–2019.
91. H. B. Gray, B. G. Malmström, R. J. P. Williams, *J. Biol. Inorg. Chem.* **2000**, *5*, 551–559.
92. B. L. Vallee, R. J. P. Williams, *Proc. Natl. Acad. Sci. USA* **1968** *59*, 498–505.
93. J. F. Hall, L. D. Kanbi, I. Harvey, L. M. Murphy, S. S. Hasnain, *Biochemistry* **1998**, *37*, 11451–11458.
94. J. F. Hall, L. D. Kanbi, R. W. Strange, S. S. Hasnain, *Biochemistry* **1999**, *38*, 12675–12680.
95. M. H. M. Olsson, U. Ryde, *J. Biol. Inorg. Chem.* **1999**, *4*, 654–663.
96. D. K. Garner, M. D. Vaughan, H. J. Hwang, M. G. Savelieff, S. M. Berry, J. F. Honek, Y. Lu, *J. Am. Chem. Soc.* **2006**, *128*, 15608–15617.
97. G. N. Ledesma, D. H. Murgida, H. K. Ly, H. Wackerbarth, J. Ulstrup, A. J. Costa-Filho, A. J. Vila, *J. Am. Chem. Soc.* **2007**, *129*, 11884–11885.

98. W. E. B. Shepard, B. F. Anderson, D. A. Lewandoski, G. E. Norris, E. N. Baker, *J. Am. Chem. Soc.* **1990**, *112*, 7817–7819.
99. H. Nar, A. Messerschmidt, R. Huber, M. v. d. Kamp, G. W. Canters, *FEBS Lett.* **1992**, *306*, 119–124.
100. S. Knapp, T. P. Keenan, X. Zhang, R. Fikar, J. A. Potenza, H. J. Schugar, *J. Am. Chem. Soc.* **1990**, *112*, 3452–3464.
101. D. W. Randall, D. R. Gamelin, L. B. LaCroix, E. I. Solomon, *J. Biol. Inorg. Chem.* **2000**, *5*, 16–29.
102. F. Xu, W. Shin, S. H. Brown, J. A. Wahleithner, U. M. Sundaram, E. I. Solomon, *Biochim. Biophys. Acta* **1996**, *1292*, 303–311.
103. E. I. Solomon, M. J. Baldwin, M. D. Lowery, *Chem. Rev.* **1992**, *92*, 521–542.
104. Y. Lu, L. B. LaCroix, M. D. Lowery, E. I. Solomon, C. J. Bender, J. Peisach, J. A. Roe, E. B. Gralla, J. S. Valentine, *J. Am. Chem. Soc.* **1993**, *115*, 5907–5918.
105. R. E. M. Diederix, G. W. Canters, C. Dennison, *Biochemistry* **2000**, *39*, 9551–9560.
106. *Comprehensive Coordination Chemistry II: From Biology to Nanotechnology*, Vol. 8, Eds. L. Que, Jr., W. B. Tolman, J. McCleverty, T. J. Meyer, Elsevier: Oxford, UK, 2004, pp 91–122.
107. G. W. Canters, G. Gilardi, *FEBS Lett.* **1993**, *325*, 39–48.
108. N. Bonander, B. G. Karlsson, T. Vanngård, *Biochemistry* **1996**, *35*, 2429–2436.
109. O. Farver, L. K. Skov, T. Pascher, B. G. Karlsson, M. Nordling, L. G. Lundberg, T. Vänngård, I. Pecht, *Biochemistry* **1993**, *32*, 7317–7322.
110. J. Salgado, S. J. Kroes, A. Berg, J. M. Moratal, G. W. Canters, *J. Biol. Chem.* **1998**, *273*, 177–185.
111. H. Li, S. P. Webb, J. Ivanic, J. H. Jensen, *J. Am. Chem. Soc.* **2004**, *126*, 8010–8019.
112. N. M. Marshall, D. K. Garner, T. D. Wilson, Y.-G. Gao, H. Robinson, M. J. Nilges, Y. Lu, *Nature* **2009**, *462*, 113–116.
113. V. Ducros, A. M. Brzozowski, K. S. Wilson, S. H. Brown, P. Ostergaard, P. Schneider, D. S. Yaver, A. H. Pedersen, G. J. Davies, *Nature Struct. Biol.* **1998**, *5*, 310–316.
114. B. G. Karlsson, M. Nordling, T. Pascher, L.-C. Tsai, L. Sjolin, L. G. Lundberg, *Protein Eng.* **1991**, *4*, 343–349.
115. T. Pascher, B. G. Karlsson, M. Nordling, B. G. Malmström, T. Vanngård, *Eur. J. Biochem.* **1993**, *212*, 289–296.
116. K. M. Clark, Y. Yu, N. M. Marshall, N. A. Sieracki, M. J. Nilges, N. J. Blackburn, W. A. v. d. Donk, Y. Lu, *J. Am. Chem. Soc.* **2010**, *132*, 10093–10101.
117. K. Kataoka, K. Yamaguchi, S. Sakai, K. Takagi, S. Suzuki, *Biochem. Biophys. Res. Commun.* **2003**, *303*, 519–524.
118. F. Xu, A. E. Palmer, D. S. Yaver, R. M. Berka, G. A. Gambetta, S. H. Brown, E. I. Solomon, *J. Biol. Chem.* **1999**, *274*, 12372–12375
119. M. C. Machczynski, H. B. Gray, J. H. Richards, *J. Inorg. Biochem.* **2002**, *88*, 375–380.
120. H. Nar, A. Messerschmidt, R. Huber, *J. Mol. Biol.* **1991**, *218*, 427–447.
121. L. M. Cunane, Z.-W. Chen, R. C. E. Durley, F. S. Mathews, *Acta Crystallogr.* **1996**, *D52*, 676–686.
122. T. Inoue, N. Nishio, S. Suzuki, K. Kataoka, T. Kohzuma, Y. Kai, *J. Biol. Chem.* **1999**, *274*, 17845–17852.
123. C. W. G. Hoitink, G. W. Canters, *J. Biol. Chem.* **1992**, *267*, 13836–13842.
124. S. Yanagisawa, M. J. Banfield, C. Dennison, *Biochemistry* **2006**, *45*, 8812–8822.
125. M. Nishiyama, J. Suzuki, T. Ohnuki, H. C. Chang, S. Horinouchi, S. Turley, E. T. Adman, T. Beppu, *Protein Eng.* **1992**, *5*, 177–184.
126. C. A. P. Libeu, M. Kukimoto, M. Nishiyama, S. Horinouchi, E. T. Adman, *Biochemistry* **1997**, *36*, 13160–13179.
127. D. W. Low, M. G. Hill, *J. Am. Chem. Soc.* **2000**, *122*, 11039–11040.

128. A. Donaire, B. Jimenez, J.-M. Moratal, J. F. Hall, S. S. Hasnain, *Biochemistry* **2001**, *40*, 837–846.
129. M. H. M. Olsson, G. Hong, A. Warshel, *J. Am. Chem. Soc.* **2003**, *125*, 5025–5039.
130. J. G. Grossmann, J. F. Hall, L. D. Kanbi, S. S. Hasnain, *Biochemistry* **2002**, *41*, 3613–3619.
131. B. Jimenez, M. Piccioli, J.-M. Moratal, A. Donaire, *Biochemistry* **2003**, *42*, 10396–10405.
132. S. M. Berry, M. H. Baker, N. J. Reardon, *J. Inorg. Biochem.* **2010**, *104*, 1071–1078.
133. N. Kitajima, K. Fujisawa, Y. Moro-oka, *J. Am. Chem. Soc.* **1990**, *112*, 3210–3212.
134. N. Kitajima, K. Fujisawa, M. Tanaka, Y. Moro-oka, *J. Am. Chem. Soc.* **1992**, *114*, 9232–9233.
135. P. L. Holland, W. B. Tolman, *J. Am. Chem. Soc.* **1999**, *121*, 7270–7271.
136. W. Maret, H. Dietrich, H.-H. Ruf, M. Zeppezauer, *J. Inorg. Biochem.* **1980**, *12*, 241–252
137. W. Maret, M. Zeppezauer, A. Desideri, L. Morpurgo, G. Rotilio, *FEBS Lett.* **1981**, *136*, 72–74.
138. W. Maret, A. K. Shiemke, W. D. Wheeler, T. M. Loehr, J. Sanders-Loehr, *J. Am. Chem. Soc.* **1986**, *108*, 6351–6359.
139. W. Maret, H. Kozlowski, *Biochim. Biophys. Acta* **1987**, *912*, 329–337.
140. M. L. Brader, M. F. Dunn, *J. Am. Chem. Soc.* **1990**, *112*, 4585–4587.
141. M. L. Brader, D. Borchardt, M. F. Dunn, *J. Am. Chem. Soc.* **1992**, *114*, 4480–4486.
142. M. L. Brader, D. Borchardt, M. F. Dunn, *Biochemistry* **1992**, *31*, 4691–4696.
143. Y. Lu, E. B. Gralla, James A. Roe, J. S. Valentine, *J. Am. Chem. Soc.* **1992**, *114*, 3560–3562.
144. Y. Lu, J. A. Roe, C. J. Bender, J. Peisach, L. Banci, I. Bertini, E. B. Gralla, J. S. Valentine, *Inorg. Chem.* **1996**, *35*, 1692–1700.
145. D. Shiga, D. Nakane, T. Inomata, Y. Funahashi, H. Masuda, A. Kikuchi, M. Oda, M. Noda, S. Uchiyama, K. Fukui, K. Kanaori, K. Tajima, Y. Takano, H. Nakamura, T. Tanaka, *J. Am. Chem. Soc.* **2010**, *132*, 18191–18198.
146. G. Battistuzzi, M. Borsari, L. Loschi, F. Righi, M. Sola, *J. Am. Chem. Soc.* **1999**, *121*, 501–506.
147. S. Yanagisawa, C. Dennison, *J. Am. Chem. Soc.* **2003**, *125*, 4974–4975.
148. S. Yanagisawa, C. Dennison, *J. Am. Chem. Soc.* **2004**, *126*, 15711–11571.
149. C. Dennison, E. Vijgenboom, W. R. Hagen, G. W. Canters, *J. Am. Chem. Soc.* **1996**, *118*, 7406–7407.
150. C. Buning, G. W. Canters, P. Comba, C. Dennison, L. Jeuken, M. Melter, J. Sanders-Loehr, *J. Am. Chem. Soc.* **2000**, *122*, 204–211.
151. R. Remenyi, L. J. C. Jeuken, P. Comba, G. W. Canters, *J. Biol. Inorg. Chem.* **2001**, *6*, 23–26.
152. C. Li, M. J. Banfield, C. Dennison, *J. Am. Chem. Soc.* **2007**, *129*, 709–718.
153. S. M. Berry, E. L. Bladholm, E. J. Mostad, A. R. Schenewerk, *J. Biol. Inorg. Chem.* **2011**, *16*, 473–480.
154. K. Sato, C. Li, I. Salard, A. J. Thompson, M. J. Banfield, C. Dennison, *Proc. Natl. Acad. Sci. USA* **2009**, *106*, 5616–5621
155. P. Hosseinzadeh, N. M. Marshall, K. N. Chacón, Y. Yu, M. J. Nilges, S. Y. New, S. A. Tashkov, N. J. Blackburn, Y. Lu, *Proc. Natl. Acad. Sci. USA* **2016**, *113*, 262–267.
156. J. Tobari, Y. Harada, *Biochem. Biophys. Res. Commun.* **1981**, *101*, 502–508.
157. V. L. Davidson, L. H. Jones, *J. Biol. Chem.* **1995**, *270*, 23941–23943.
158. D. J. Kosman, *J. Biol. Inorg. Chem.* **2010**, *15*, 15–28.
159. R. A. Marcus, N. Sutin, *Biochim. Biophys. Acta* **1985**, *811*, 265–322.
160. S. D. George, M. Metz, R. K. Szilagyi, H. Wang, S. P. Cramer, Y. Lu, W. B. Tolman, B. Hedman, K. O. Hodgson, E. I. Solomon, *J. Am. Chem. Soc.* **2001**, *123*, 5757–5767.

161. M. Ubbink, M. Ejdebäck, B. G. Karlsson, D. S. Bendall, *Structure* **1998**, *6*, 323–335.
162. N. Sukumar, Z. Chen, D. Ferrari, A. Merli, G. L. Rossi, H. D. Bellamy, A. Chistoserdov, V. L. Davidson, F. S. Mathews, *Biochemistry* **2006**, *45*, 13500–13510.
163. L. Quintanar, M. Gebhard, T.-P. Wang, D. J. Kosman, E. I. Solomon, *J. Am. Chem. Soc.* **2004**, *126*, 6579–6589.
164. A. B. Taylor, C. S. Stoj, L. Ziegler, D. J. Kosman, P. J. Hart, *Proc. Natl. Acad. Sci. USA* **2005**, *102*, 15459–15464.
165. T. Bertrand, C. Jolivalt, P. Briozzo, E. Caminade, N. Joly, C. Madzak, C. Mougin, *Biochemistry* **2002**, *41*, 7325–7333.
166. T. Inoue, H. Sagawara, S. Hamanaka, H. Tsukuil, E. Suzuki, T. Kohzumal, Y. Kai, *Biochemistry* **1999**, *38*, 6063–6069.
167. J. J. Warren, K. M. Lancaster, J. H. Richards, H. B. Gray, *J. Inorg. Biochem.* **2012**, *115*, 119–126.

4

Purple Mixed-Valent Copper A

Marcos N. Morgada, [1] ***Daniel H. Murgida,*** [2] ***and Alejandro J. Vila*** [1]

[1]Instituto de Biología Molecular y Celular de Rosario (IBR), Área Biofísica, Departamento de Química Biológica, Facultad de Ciencias Bioquímicas y Farmacéuticas, Universidad Nacional de Rosario-CONICET, Rosario, Argentina
<morgada@ibr-conicet.gov.ar>
<vila@ibr-conicet.gov.ar>

[2]Instituto de Química Física de los Materiales, Medio Ambiente y Energía (INQUIMAE), Departamento de Química Inorgánica, Analítica y Química Física, Facultad de Ciencias Exactas y Naturales, Universidad de Buenos Aires – CONICET, Buenos Aires, Argentina.
<dhmurgida@qi.fcen.uba.ar>

Metal Ions in Life Sciences, Volume 20 Guest Editors: Martha E. Sosa Torres and Peter M. H. Kroneck
Series Editors: Astrid Sigel, Eva Freisinger, and Roland K. O. Sigel

https://doi.org/10.1515/9783110589757-004

Abstract: Cu_A is a binuclear copper center acting as an electron transfer hub in terminal oxidases such as cytochrome *c* oxidase and nitrous oxide reductase. Its unique electronic structure is intimately linked to its function and has puzzled the community of biological inorganic chemistry for decades. Here we review the insights provided by different spectroscopic techniques of Cu_A centers, and the different experimental approaches to tackle its study, that encompass the synthesis of model compounds as well as protein engineering efforts. The contribution of the electronic structure to the thermodynamic and kinetic of electron transfer is extensively discussed. We also describe the proposed mechanism of Cu_A assembly in different organisms. The recent discovery of a novel Cu_A site opens new perspectives to this field.

Keywords: Cu_A · cytochrome *c* oxidase · electronic structure · long range electron transfer · metallochaperones · nitrous oxide reductase

1. INTRODUCTION: A REDOX HUB COMMON TO DIOXYGEN AND NITROUS OXIDE REDUCTASE

Cu_A is a binuclear copper center acting as the electron entry port in terminal oxidases, i.e., enzymes transferring electrons to a final acceptor of a respiratory chain [1, 2]. The most notorious enzymes from this family are the oxygen reductases, better known as cytochrome *c* oxidase (abbreviated as COX or Complex IV) [3–5], a central enzyme in the respiration of most aerobic cells, and N_2O reductase (abbreviated as N_2OR or NosZ) [6–10], the terminal enzyme in the anaerobic respiration in denitrifying bacteria, which transforms nitrous oxide into molecular nitrogen. Indeed, the discovery of the structural and electronic features of the Cu_A center could be pursued by almost simultaneous spectroscopic, mutational, and structural studies in cytochrome *c* oxidases and N_2O reductases. This metal site is also present in the Cu_A-dependent nitric oxide reductase (Cu_A NOR) from *Bacillus azotoformans*, a denitrifying gram-positive bacterium. In contrast to most NORs, this enzyme lacks a heme *c* and instead, utilizes the Cu_A center as electron acceptor from a soluble cytochrome [11, 12]. Several crystal structures are available for COX and N_2OR from different organisms under variate conditions, but the presence of a Cu_A site in NOR is only based in spectroscopic methods at the moment [12].

Cytochrome *c* oxidase is an integral membrane protein composed of several subunits, that can range from the three essential, preserved ones (COX I,

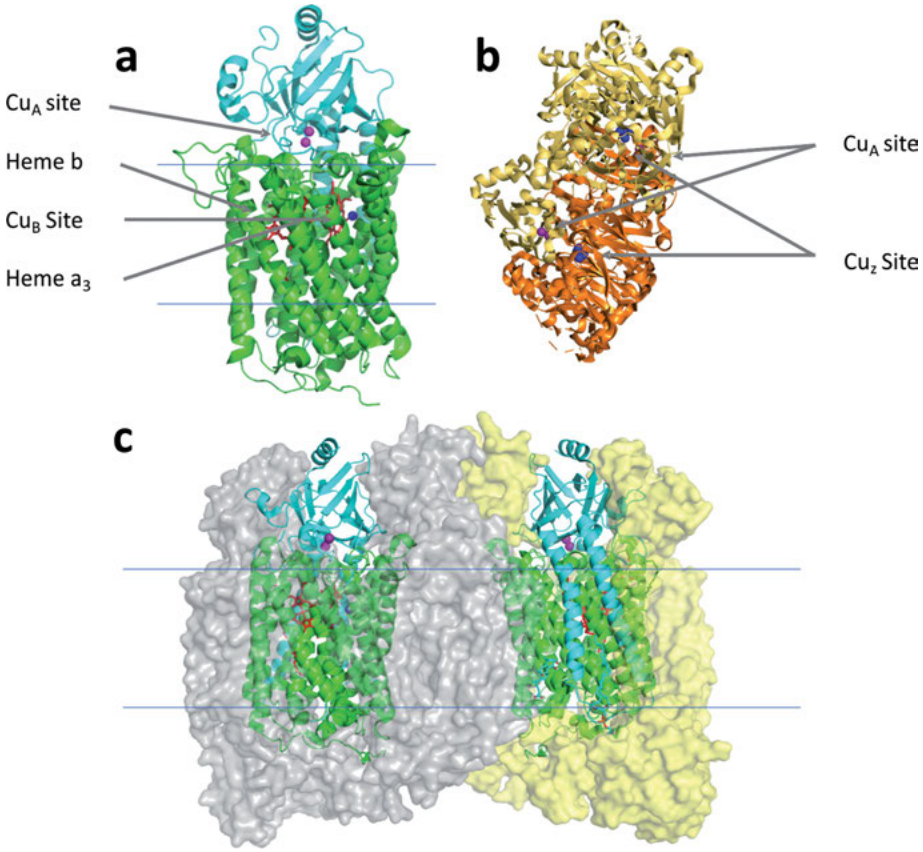


Figure 1. Crystal structures of representative Cu_A-containing enzymes. (**a**) *Thermus thermophilus* ba_3 cytochrome *c* oxidase (PDB 3S8F) [19]. The Cu_A center (purple spheres) is located in subunit COX II (in cyan). The oxidase also contains two heme moieties (in red), one of them tightly associated to the Cu_B site (blue sphere) forming the catalytic oxygen reduction site. (**b**) *Pseudomonas stutzeri* N_2O reductase (PDB 3SBR) [20] is a dimeric protein with one Cu_A site (purple) in each monomer and two multicopper Cu_z site (blue) in the protein-protein interface. (**c**) Bovine cytochrome *c* oxidase (PDB 1V54 [21]. The complex is a dimer with each monomer composed of 13 subunits. All the metal cofactors are located in subunits COX I (green) and COX II (cyan) with the other subunits stabilizing the complex (grey and yellow). The horizontal thin lines in (a) and (c) indicate the membrane.

COX II, and COX III) in bacteria to 14 subunits in mammals (Figure 1a, c) [13, 14]. The essential metal cofactors (Cu_A, Cu_B, and two heme centers) are bound to the three conserved subunits that form the catalytic core of the respiratory enzyme. The Cu_A site accepts electrons from cytochrome *c*, that are then shuttled to the heme *a*, and finally to the heme a_3-Cu_B binuclear center where O_2 reduction takes place [5, 15]. In some bacteria, the heme *a* is replaced by a heme *b* moiety [16–18].

N_2O reductase is a homodimeric, periplasmic bacterial protein containing six copper atoms per monomer arranged into two sites (Figure 1b): Cu_A, which is the electron acceptor, and Cu_Z, the tetranuclear catalytic cluster where N_2O is bound and reduced [7, 22, 23]. Electron transfer from Cu_A and Cu_Z takes place between centers in different subunits located at 10 Å [20, 24] (see Chapter 5 of this Volume).

In all these enzymes, the Cu_A center is bound to a small subunit adopting the so-called cupredoxin fold. This center cycles between the fully reduced, Cu^{1+}-Cu^{1+} state, and the mixed-valent $Cu^{1.5+}$-$Cu^{1.5+}$ state by means of a one-electron reaction. The Cu^{2+}-Cu^{2+} state, instead, is not accessible under physiological conditions (indeed, the electrochemically-driven oxidation to this state is irreversible) [25]. Cu_A is one of the most efficient electron transfer hubs in Nature, a fact that is due to its unique geometry and electronic structure that are tailored to optimize long range electron transfer to and from this center [2, 26, 27].

An exception is the membrane-anchored, periplasmic protein PmoD from the methane-oxidizing bacterium *Methylosinus trichosporium* OB3b [28, 29], in which a symmetric Cu_A site is localized at the interface of a homodimer (see below). In contrast to the Cu_A-containing oxidases, this protein does not harbor additional metal cofactors that may serve as catalytic redox sites. No function has been assigned yet to this protein, despite the fact that it has been shown to be critical for the copper-dependent growth of this organism [28].

The identity, metal content and structure of the Cu_A site have been matter of intense debate, dating back to the 1930s, when Elvehjem reported that copper was essential for the respiratory activity in yeast cells [30]. Later, Keilin and Hartree provided evidence showing that copper was associated to cytochrome *c* oxidase [31]. These early discoveries, crucial for the understanding of the mechanism of an essential enzyme for aerobic life, triggered an intense debate in the community of bioenergetics, biochemistry, biophysics, and biological inorganic chemistry that lasted for almost six decades [1, 2, 26]. The strong absorption features of the heme cofactors in the oxidases and the presence of an additional copper site (Cu_B) precluded the clean observation of the spectroscopic features from Cu_A [31]. The fact that COX is an integral membrane protein has further hampered the study of this center. Therefore, two main strategies have been exploited to circumvent the interference of the other cofactors: (1) the expression of soluble domains from the oxidases containing only the Cu_A center, that has not always been successful [32–36] and (2) the engineering of Cu_A centers in a soluble protein scaffold [37–40]. These studies were also supplemented by different attempts to synthesize model compounds, that have also proven challenging [41–44].

During this period, spectroscopic techniques played a key role in the elucidation of the structure of the Cu_A site, prior to the availability of the first crystal structures for Cu_A-containing proteins in the mid-1990s [45–47]. Excellent reviews from Helmut Beinert [1] and (more recently) from Peter Kroneck [2] provide first-hand historic reports of the discovery and structure elucidation of this metal site. This aspect will not be extensively covered here. Instead, we will

cover the variety of Cu_A centers and their spectroscopic features, focusing on most recent progress regarding the alternative ground states, a thermodynamic description of Cu_A redox features, and the biological assembly of these centers in proteins.

2. THE ARCHITECTURE OF NATIVE Cu_A AND MODEL SYSTEMS

2.1. Native Cu_A Sites

The Cu_A site is characterized by a rhombic Cu_2S_2 diamond core in which the two copper ions are bridged by two Cys thiolate sulfurs (Figure 2) [7, 13, 46, 48]. Each copper ion is coordinated to an extra His ligand, and there are two weak axial ligands (one on each copper ion) in a *trans* arrangement: the thioether moiety from a Met residue, and the oxygen atom from a peptide bond of variate nature [33, 46, 49]. These differences in the axial ligation make the site slightly asymmetric, despite the charge delocalization in the oxidized state being such that the two copper ions are equivalent. These centers are bound to small domains displaying a cupredoxin fold, the same originally characterized for the Type 1 (or blue) copper centers [50] (see Chapter 3 of this Volume). This fold consists of seven strands in a Greek-key β-barrel arrangement, with the copper ligands located in two loops, in such a way that the metal site is located close to the protein surface (Figure 2b). Five out of the six metal ligands are arranged in a $CysX_3CysX_3HisX_2Met$ motif within the ligand loop (PmoD being an exception) [51]. The remaining His ligand is located upstream in the primary sequence in a loop close to the metal site (Figure 2b). The His ligands are not equivalent, while the His residue from the ligand loop is partially solvent-exposed, the other is totally buried, resulting in one copper ion located closer to the protein surface (Figure 2).

The presence of S-, N-, and O-donating atoms resembles the chemistry of the mononuclear Type 1 (blue) copper centers, and this combination of hard and soft donor atoms stabilizes both, the oxidized mixed-valent $Cu^{1.5+}$-$Cu^{1.5+}$ and the reduced Cu^{1+}-Cu^{1+} state [54, 55]. The Cu–S(Cys) bond distances are between 2.2 and 2.4 Å, while the Cu–N(His) distances are between 2.0 and 2.2 Å. Instead, the distance to the axial Met residue depends on the protein source. In native Cu_A centers, the Cu-S(Met) distances range from 2.34 to 2.5 Å, thus are significantly shorter than those reported for Type 1 Cu centers. A short Cu–Cu distance of 2.4 Å was initially suggested by EXAFS experiments [56] and later was confirmed to be in this range by crystallographic data [46]. Biosynthetic Cu_A sites engineered in cupredoxin scaffolds are characterized by longer Cu–S(Met) bonds (2.9–3.1 Å in Cu_A azurin and 3.0 Å in CyoA), that also correspond to shorter Cu–Cu bonds (2.35–2.44 Å). The Cu_A site is the first example of a metal site with a direct inter-metallic bond within a protein. These unusual features are driven by the constrained nature of this rhombic structure that minimizes structural changes upon one-electron reduction, resulting in a low reorganization

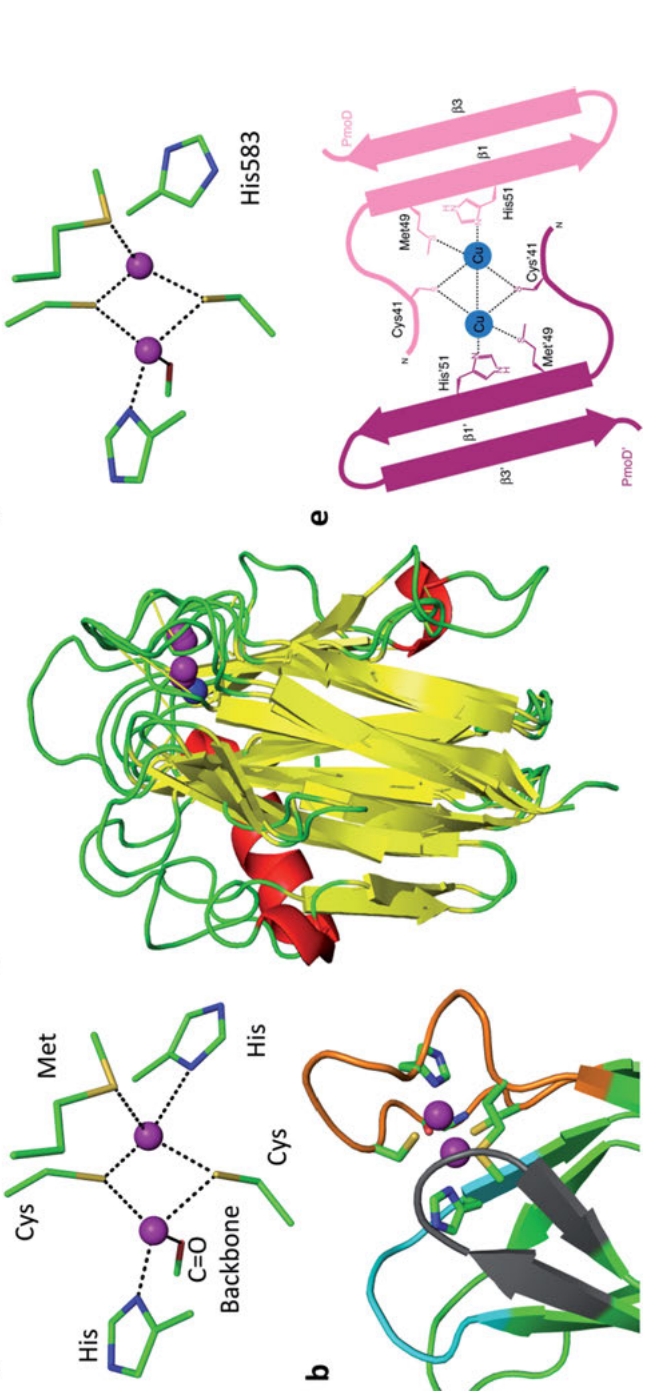


Figure 2. Structural features of Cu_A sites. (**a**) Typical geometry of a Cu_A center, showing first coordination sphere conserved ligands [20]. (**b**) Location of the CuA site within the cupredoxin fold. The loops surrounding the metal site can be identified according to the presence of metal site ligands: ligand loop (orange), His loop (cyan) and the loop indicated as the redox partner interacting loop (gray) [52]. (**c**) Structural alignment of cupredoxin proteins, *Thermus thermophilus* Cu_A (PDB 2CUA) [52], *Pseudomonas stutzeri* nitrous oxide reductase Cu_A harboring domain (from PDB 3SBR) [20] and *Populus nigra* plastocyanin (PDB 1PLC) [53]. The ß barrel motif is well preserved (yellow) connected by loops (green). Helical parts of the structures (red) are more divergent among the different cupredoxin solved structures. (**d**) Five ligands Cu_A site in nitrous oxide reductase (from PDB 3SBR) [20]. Detachment of the buried His583 is favored by substrate binding to the protein active site [20]. (**e**) Model of the novel Cu_A site in PmoD (reprinted from [28]).

energy that facilitates long range electron transfer [15]. Recent high resolution structures of the bovine cytochrome *c* oxidase in the fully oxidized and fully reduced states have revealed that the geometry of the Cu_A center [57] is largely unchanged upon redox change, and that transmission of the redox change in the protein structure takes place by subtle structural changes. Overall, the metal ligand set and the geometry in native Cu_A sites in cytochrome *c* oxidases has been shown to be preserved upon different perturbations.

A quite different situation has been reported for *Pseudomonas stutzeri* N_2O reductase, in which crystal structures have disclosed the detachment of the upstream His ligand from the copper ion bound to the axial Met (Figure 2d) [20]. This results in a metal site rearrangement, with an increased asymmetry in the Cu_A center, with a shorter Cu–S(Met) bond, a longer Cu–Cu distance (2.6–2.7 Å) and a longer Cu–S(Cys) distance for one of the Cys ligands. Surprisingly, these structural changes do not alter the characteristic spectroscopic features of this site (see below) [2]. This perturbation is further stimulated upon N_2O binding to the catalytic Cu_Z cluster [20]. The conformational change has been related to the possibility of gated electron transfer, but this has not been demonstrated yet. A possible mimic of the structure of this site is the His120Ala mutant in Cu_A-azurin, that despite losing a His ligand (in this case, the one located in the ligand-containing loop) retains the mixed-valence character due to the high rigidity of the core [58, 59].

While the ligand set is fully conserved in Cu_A centers, their second coordination sphere is more variable. The identity of the non-ligated amino acids of the ligand loop varies among different organisms, probably in order to optimize the interaction with the corresponding redox partner or to tune the electronic structure [33, 46, 49, 60, 61]. On the other loop, next to the buried His ligand, a residue of variable nature forms a hydrogen bond with the S atom of the first Cys within the ligand loop (Cys 149 in *Thermus* Cu_A) [20, 62]. NMR data have revealed that this hydrogen bond is stronger for the oxidized cluster than for the reduced cluster [63]. This oxidation-dependent hydrogen bond strength, can represent a general mechanism for the regulation of the electron transfer reactivity of the Cu_A center, that may account for the minor differences suggested by the crystal structures of the whole bovine oxidase [57]. The modification of the hydrogen bond network can be a way to manifest the redox state of the cluster to distant parts of the protein and to regulate the interaction with redox partners [64, 65].

The rigid structure of the Cu_A site, displaying an identical geometry both for the reduced and the oxidized form, follows the principle of the "rack" or "entatic" state [55, 66, 67], that postulates that a rigid protein scaffold forces the metal ions to adopt similar geometries upon electron transfer, thus decreasing the reorganization energy. This theory was supported by the finding of crystal structures of cupredoxins proteins in the apo (demetallated) form displaying the same conformation as in the holo (metal-containing) form, suggesting the existence of preorganized metal binding sites [68]. This finding is quite in contrast to our previous knowledge about the functioning of a metallochaperone [69], which typically involves a protein-protein interaction that is not compatible with the

existence of a preorganized site in the apoprotein [70]. NMR studies have shown that apocupredoxins (both for Type 1 and for Cu_A sites) are mobile, dynamic structures that enable protein-protein recognition events with the purported metallochaperones, and that this dynamic is quenched by metal binding that ultimately contributes to the required rigidity [70, 71].

2.2. Protein Model Systems

The spectroscopic study of Cu_A sites in whole oxidases was largely hampered by the presence of other metal centers both in COX and in N_2OR. There have been several strategies to obtain soluble proteins containing only the Cu_A center in the cupredoxin scaffold.

N_2O reductases can be expressed devoid of the catalytic Cu_Z site in a mutant strain of *P. stutzeri* MK402, deficient in the biosynthesis of this cofactor. This N_2OR variant was essential to pursue the quintessential EPR studies by which Kroneck, Antholine, Zumft, and coworkers attributed the S = 1/2 paramagnet to a Type III mixed-valence binuclear site [72–74]. N_2OR variants containing only the Cu_A site can also be achieved by addition of free copper ions to the purified apo-protein, since the Cu_Z site required inorganic sulfur atoms cannot be assembled *in vitro* [75].

In the case of the cytochrome *c* oxidases, two approaches have been pursued: (1) engineering a Cu_A-binding site into a non-Cu_A cupredoxin scaffold [37, 76] and (2) expression of the soluble COX II fragments [32, 33, 35].

The first strategy relies on the observation that most of the copper ligands of the Cu_A site are contained in a loop with a $CysX_3CysX_3HisX_2Met$ sequence in the β-barrel of the cupredoxin fold (Figure 2b, c). A similar loop (with different length and sequence) is present in the Type 1 copper proteins and in related cupredoxin domains devoid of metal ligands, such as the one in the *E. coli* cytochrome *o* quinol oxidase [37]. Indeed, the latter was the first successful Cu_A-construct achieved by Saraste and coworkers, who engineered copper binding sites in the soluble CyoA domain of this oxidase that reproduced the spectroscopic features of Type 1 and Cu_A sites [47]. Indeed, later mutagenesis studies validated the identity of the proposed copper ligands [54] and the crystal structure of this construct provided one of the first glimpses into the three-dimensional structure of a Cu_A site [38, 47].

This approach has been successful for engineering Cu_A centers in Type 1 proteins such as azurin, by the Lu group [76, 77] and amicyanin, initially by Canters et al. and later by Davidson et al. [39, 40]. Due to the robust nature of the azurin scaffold, Cu_A-azurin has been a unique model for spectroscopic and redox studies of this center [58, 59, 77–81]. Overall, these constructs point to a close evolutionary relationship among these domains [32, 80].

The second approach involves expressing the water-soluble Cu_A-containing domain of subunit 2 from cytochrome *c* oxidase (COX II) in a soluble form. This can be achieved by truncating the N-terminal helices anchoring this subunit to the membrane (Figure 1a, c). This strategy was successful for many bacterial

oxidases, such as the ones from *P. denitrificans* [32], *B. subtilis* [34], *T. thermophilus* [33], *P. versutus* [82], *Synechocystis* PCC 6803 [35], and from the archaeal oxidase from *S. acidocaldarius* [83]. All these centers have provided excellent working systems for the reproduction of the spectroscopic and functional features of Cu_A sites. The *T. thermophilus* COX II subunit has proven the most exploited construct, based on the high stability of this protein, for which there also exists a high-resolution crystal structure [52].

The expression and purification of COX II subunits from eukaryotic oxidases has not been successful so far. This can be attributed to the different degrees of complexities of prokaryotic and eukaryotic enzymes. In bacteria, this subunit is located in the periplasmic space, exposed to the solvent, while in eukaryotes it is embedded within the inner mitochondrial space interacting with other subunits, generally of hydrophobic nature (Figure 1). This renders unstable constructs which have been obtained by simple truncation of the N-terminal anchor. Instead, successful mimics of eukaryotic COX II subunits have been achieved by engineering the loops of the human and plant proteins into the stable, soluble scaffold of *T. thermophilus* Cu_A (*Tt*Cu_A hereafter) [71, 84].

2.3. Synthetic Model Compounds

Another common approach to the study of metal sites in proteins is the chemical synthesis of model compounds to mimic the structure of the metal sites within proteins. The synthesis of copper-thiolate compounds is challenging due to the high reactivity of thiol groups with copper ions, and this has been a problem for mimics of Type 1 copper proteins. The presence of the bridged dithiolate structure in Cu_A has made this endeavor even more challenging, since this results mostly in the formation of a disulfide bond. Among many efforts [44], there are only two successful examples of Cu_A mimics [41, 42].

Houser, Young, and Tolman reported in 1996 the synthesis of a delocalized mixed valence bis(μ-thiolato) dicopper complex, $[Cu_2(iPrdacoS)_2]$ (iPrdacoS = 1-isopropyl-5-ethylthiolato 1,5-diazacyclooctane), that features a seven-line EPR spectrum indicative of a fully delocalized mixed valence center (Figure 3a) [41]. The absorption spectrum shows somehow perturbed features in the visible range, and an intense band at ca. 1500 nm not present in native Cu_A centers, that has been attributed to the presence of a distinct ground state in this compound (see below) [85, 86]. The Cu–Cu distance in this complex (2.92 Å) is considerably larger than in Cu_A centers in proteins (2.4–2.5 Å), either native or engineered [41, 56].

In 2011, Gennari et al. reported a new bis(μ-thiolato)dicopper complex based on a 2,2′-(2,2′-bipyridine-6,6′-diyl)bis(1,1-diphenylethanethiolate) ligand that reproduced a planar rhombic Cu_2S_2 core and displayed the characteristic EPR spectra of Cu_A sites (Figure 3b) [42]. This is the first Cu_A model that can be reversibly oxidized or reduced between the native $Cu^{1.5+}$-$Cu^{1.5+}$ and Cu^{1+}-Cu^{1+} states. The Cu–Cu distance of the oxidized state (2.93 Å) was similar to the one in the Tolman model, in agreement with the similar absorption spectra. However, the reduced state showed an intermetallic distance of 2.64 Å, closer to those of Cu_A

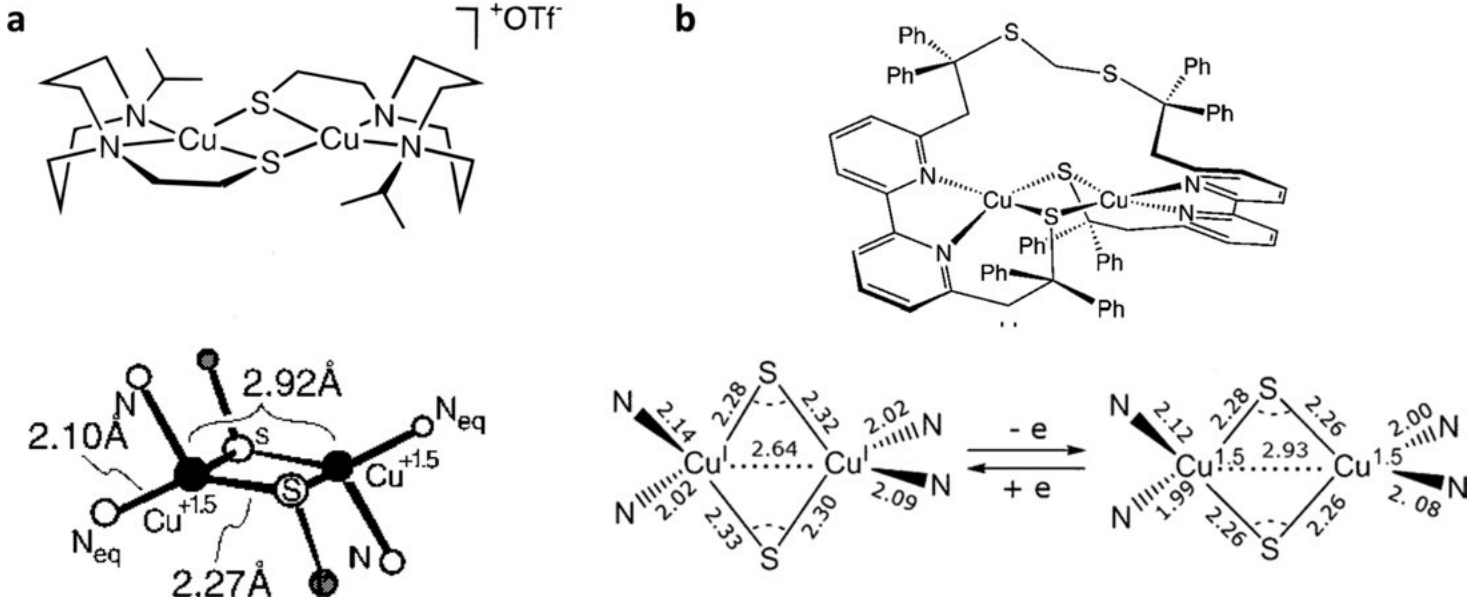


Figure 3. Model compounds of Cu_A sites. (**a**) Structure of the model synthesized by Houser, Young, and Tolman [41]. Figure reproduced from [86] with permission of the American Chemical Society, copyright 1998. (**b**) Structure of the model compounds synthesized by Gennari et al. [42]. The relative geometries of the metal centers in the oxidized and reduced states are depicted for the latter model, while the former one could only be characterized in the oxidized form. Reproduced from [42] with permission of Wiley, copyright 2011.

sites in proteins. The change in distance is consistent with the high reorganization energy reported for this compound (1.93 eV) [87], much higher to those expected for an efficient ET center in nature (see below) [88]. These results suggest that, despite the electronic structure of these complexes, biological metal centers may be reproduced by model chemistry nevertheless, the protein framework induces a rigidity in these sites that is hardly reproduced in small molecules.

3. BIOGENESIS AND ASSEMBLY

Copper uptake by metalloproteins in the cell requires the action of specific proteins and mechanisms, since copper ions are essential, but they also can be toxic under conditions of metal dyshomeostasis [89, 90]. Thus, copper levels are strictly regulated within cells [91, 92]. This regulation involves specific metal pumps, compartmentalization in different organelles, metal storage proteins, and metallochaperones. The latter proteins are responsible of delivering the copper ions to the target proteins. This requires trafficking of the metal ion to the cellular compartment where metal site assembly takes place. This event defines the cellular localization of the proteins involved in the final steps of the copper delivery pathway [69].

COX II (containing the Cu_A center) is located in the periplasm of gram-negative bacteria and in the extracellular space of gram-positive organisms. Instead, in eukaryotes subunit II is within the mitochondrial intermembrane space (IMS). In N_2O reductases, the Cu_A site is also located in the periplasmic space [93, 94]. In all these cases, Cu_A assembly has been shown to occur in the cellular compartment where the corresponding subunits are located [71, 93, 95, 96]. Thus, the identification of possible metallochaperones can be pursued by identifying proteins essential for the COX or N_2OR function with copper binding domains within the same subcellular compartment. This led to the discovery of the family of Sco (synthesis of cytochrome *c* oxidase) proteins, which are involved in Cu_A biogenesis in COX and N_2OR's [97–100].

Assembly of the Cu_A site requires not only copper transfer, but also the concerted binding to the thiolate moieties from the bridging Cys ligands (Figure 2a). Since Cu_A binding domains are located in oxidizing environments (such as the periplasm or the IMS), the redox homeostasis of these thiol groups is linked to the metal-binding event, since the Cys residues must be maintained in the reduced form to bind the copper ions. Thus, Cu_A assembly requires not only a metallochaperone, but also a thiol-disulfide oxide/reductase for the final assembly step. In this chapter we will describe only this last step in the Cu_A biogenesis, referring to other literature for an integral description of copper trafficking in cells [101–103].

Sco proteins are small membrane-anchored proteins with a conserved thioredoxin-like soluble domain located in the periplasmic space or in the IMS [98], with the insertion of additional secondary structure features (Figure 4a). They possess two Cys residues in a $CysX_3Cys$ motif that can be oxidized to form a disulfide bridge (Figure 4c). In contrast to thioredoxins, Sco proteins have a

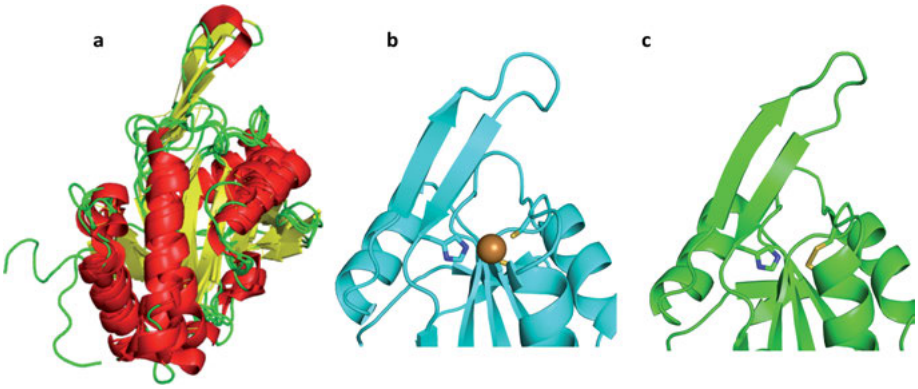


Figure 4. Structural features of Sco proteins. (**a**) Scaffold conservation of Sco protein from different organisms *Bacillus subtilis* (PDB 1XZ0), human (PDB 1WP0) [104], *Thermus thermophilus* (PDB 2K6V)[96], and *Arabidopsis thaliana* (PDB 6N5U) [84] (**b**) Metal binding site of Cu(I)-bound Sco from *Arabidopsis thaliana* (PDB 6N5U) [84] and (**c**) oxidized human Sco1 protein (PDB 2GGT) [104].

protruding loop with a His residue that, together with the two Cys, can bind Cu(I) or Cu(II) ions (Figure 4b). Because of these structural features, Sco proteins can act as metallochaperones by donating the copper ions to the target proteins, or act as a thiol-disulfide oxidoreductases to maintain the copper-binding cysteines in the proper oxidation state to accept the copper ions.

3.1. Biogenesis of the Cu_A Site in Cytochrome *c* Oxidase

The assembly of cytochrome *c* oxidases is a complex, stepwise process that involves the synthesis of the different subunits, their translocation and insertion into the membrane, and (finally) the assembly of the different metal cofactors [95]. In eukaryotes, the different subunits can be encoded either in the nucleus or in the mitochondria. Subunit COX II, of interest here, is mitochondrially encoded.

Proteins from the Sco family have been identified in all organisms with a Cu_A-dependent oxidase (Figure 5) [97]. Knock-out experiments in yeast [105] (and later in bacteria) [106] pointed to Sco as an essential protein for the assembly of the copper cofactors in cytochrome *c* oxidase [107]. Sco homologues were found in higher organisms together with other assembly proteins. Some eukaryotic organisms have been shown to express more than one Sco variant, as it is the case for yeast and mammalian cells. In humans, specific mutations in Sco1 and Sco2 have been linked to lethal pathologies related to loss of oxidase function due to an impairing of the Cu_A assembly [98].

Cu_A assembly in mammals depends on three key proteins: Cox17, Sco1, and Sco2 [108]. Cox17 is a soluble protein located in the IMS that belongs to a family of mitochondrial proteins with twin $CysX_9Cys$ motifs that forms two disulfide bonds and utilizes two Cys in the flexible N-terminal stretch to bind Cu(I) [107,

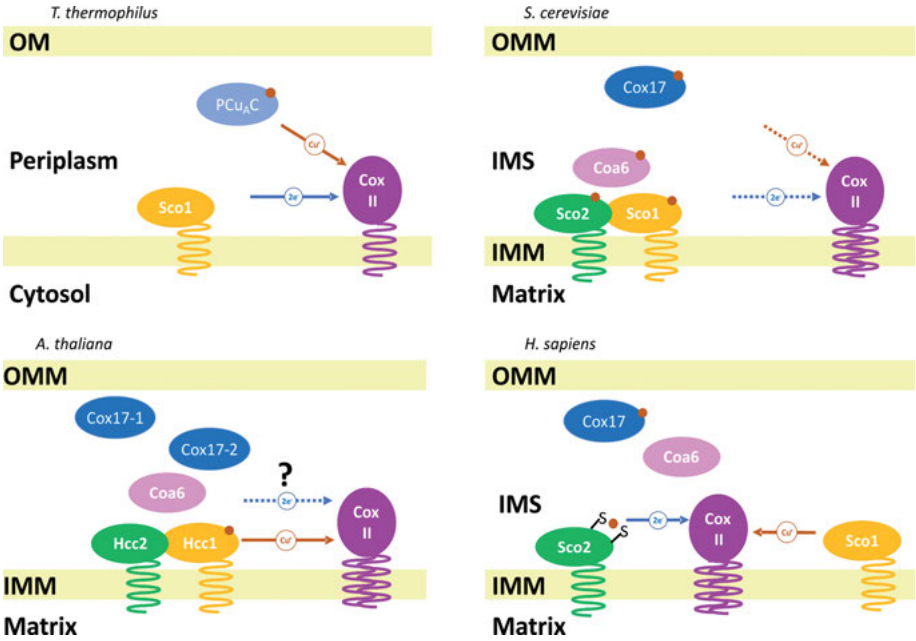


Figure 5. Proposed mechanisms of Cu_A assembly in different organisms. The electron and copper donors are indicated with arrows. Dotted arrows indicate suggested functions. OM: Outer membrane in bacteria, IM: inner membrane in bacteria, OMM: outer mitochondrial membrane, IMM: inner mitochondrial membrane, IMS: mitochondrial intermembrane space.

108]. This protein is conserved in other eukaryotes and has been shown to transfer copper to Sco1 and Sco2, but with different mechanisms. Cox17 is able to transfer one equivalent of Cu(I) to Sco1 in an event coupled with a redox reaction in which the disulfide bond in apo Sco1 is reduced, thus promoting metal transfer. Instead, Cox17 directly transfers Cu(I) to reduced Sco2 [109].

Sco1 and Sco2 are paralogous proteins that are both able to bind copper [110]. In yeast, only Sco1 is essential, while both are essential for Cu_A assembly in the human oxidase [93, 107, 110]. The details of the chemistry involving Sco1 and Sco2 in yeast have not been elucidated yet. In the case of the human mitochondria, Sco2 has been shown to act upstream to Sco1 and both proteins play complementary roles [111].

Since all Sco proteins are homologous, genetic experiments can only help in discerning the essential from the non-essential proteins in the assembly process. Note that such experiments do not clearly define the biochemical roles of each factor. Thus, biochemical studies *in vitro* are required to finally annotate each protein function. However, the isolation of subunits COX II in soluble, pure form from eukaryotic oxidases has proven difficult due to the hydrophobic interactions on the surface of these subunits that render them unstable. As described above, a chimeric protein with the scaffold of soluble *Tt*Cu_A and the three loops of the human protein could recreate a model of human COX II (h*Tt*3L for

human *T. thermophilus*) [71]. Biochemical and biophysical studies using this model have demonstrated that Sco2 acts (in the copper-bound form) as a thiol-disulfide oxidoreductase, while Sco1 is the copper chaperone (Figure 5) [71]. This process takes place by specific protein-protein recognition involving the loops of the Sco proteins and those from the COX II domain that provide the metal-binding region.

A similar strategy has been recently applied to elucidate Cu_A assembly in the plant oxidase from *Arabidopsis thaliana* [84]. In this case, a Sco homologue known as Hcc1 has been shown to act as the metallochaperone. Instead, Hcc2 is a Sco-like protein devoid of the active Cys residues, remaining at the moment without any functional annotation (Figure 5).

It is likely that many other proteins are involved in this complex process. Factors already identified involved in these processes are Cox17 and Coa6 [112, 113]. Both genes affect the assembly of the Cu_A site indirectly. In yeast, Coa6 seems to have an overlapping function with Sco2, but none of them is individually essential for the assembly of the metal site [114].

In prokaryotes, the stability of the COX II domain of the ba_3 oxidase of *T. thermophilus* [33] early enabled biochemical experiments that provided insights into the mechanistic details of the assembly. The first reported *in vitro* Cu_A assembly process was the formation of the Cu_A site in this protein [96]. In this particular case, a Sco protein from *T. thermophilus*, despite being able to bind copper ions, cannot load the copper ions into the oxidase subunit but, instead, acts as a thiol oxidoreductase of the cystine bond on COX II [96]. A soluble periplasmic protein, PCu_AC, encoded in the same operon as COX II, binds Cu(I) and acts as the metallochaperone transferring two Cu(I) ions sequentially to COX II forming a functional Cu_A site (Figure 5).

PCu_AC is not conserved in many bacteria, and some homologues have been identified in gram-positive bacteria. It is proposed that in some bacterial species lacking a homologue of this metallochaperone, the Sco proteins act as metal donors for the Cu_A site.

3.2. Biogenesis of the Cu_A Site in Nitrous Oxide Reductase

N_2O reductase is a bacterial protein located in the periplasmic space [100]. N_2OR is translocated across the membrane by the *tat* system as a folded protein, and the assembly of the metal centers takes place within the periplasm [100]. Two genes possibly involved in Cu_A metalation have been located in the same operon coding for NosZ (N_2OR) [115]. NosA is an integral membrane protein inserted in the outer membrane, characterized as a copper transporter that is important for the assembly of both Cu_A and CuZ. Another gene product is NosL, a lipidated protein anchored to the inner leaflet of the outer membrane. Since NosL binds one Cu(I) equivalent with high affinity, it was originally described as a potential metallochaperone for the Cu_A site [116]. Knockout experiments, however, later showed that NosL is not involved in the assembly of the

metal center, and in *P. denitrificans* it was shown to be essential for Cu_Z assembly under copper-limited conditions [117].

A member of the Sco family, ScoP was also postulated as a possible Cu_A chaperone for N_2OR [99]. Nevertheless, deletion of the ScoP gene in different bacteria was not detrimental for assembly of the Cu_A site. Based on this evidence, further experiments are required to identify copper proteins, or accessory proteins, with relevance to the formation of the Cu_A site in N_2OR [100]. It is not unlikely that conserved proteins across the genomes of different related bacteria could be playing different, or complementary roles, showing a diversity as such reported for the assembly of the Cu_A site in COX.

4. ELECTRONIC STRUCTURE

4.1. Spectroscopic and Theoretical Description of the Native Site

The Cu_A site in its reduced form, with two d^{10} Cu^{1+} ions, is diamagnetic and colorless. Instead, the oxidized mixed-valent center can be readily recognized by its intense purple color due to its unique electronic structure, intimately related to its efficiency as electron transfer hub [2, 26, 27]. The oxidized form of Cu_A corresponds to a S = 1/2 paramagnet, with unique EPR properties (see below). These two spectroscopic features can be readily identified in proteins containing only the Cu_A chromophore, but in the early times, before the discovery of N_2OR, these features were hard to assign. Early work from Thomson and coworkers [118] identified the absorption band in the near infrared region with a maximum at 830 nm by low temperature magnetic circular dichroism as arising from a S = ½ paramagnet with *g* values close to 2. They also linked these features to the absorption maxima at 480 and 530 nm [118]. Later, comparative studies of the copper centers in COX and N_2OR established that there was a common metal center in both proteins [119].

The purple color of Cu_A is due to two intense S(Cys)-Cu ligand → metal charge transfer (LMCT) absorption maxima at 480 and 530 nm (Figure 6) [76], assigned based on resonance Raman (RR) experiments [120, 121]. The band in the near infrared region is also characteristic of Cu_A sites, and can be located between 750 and 800 nm. This absorption feature corresponds to a $\psi \rightarrow \psi^*$ intervalence transition resulting from a direct copper-copper interaction [76, 121, 122]. The position of the two LMCT bands is well conserved among different Cu_A centers, but their relative intensity can vary (Figure 6). Moreover, a decreased intensity in these bands correlates with an increase in the intensity of an absorption at approx. 360 nm (Figure 6), that has been attributed either to a N(His) → Cu LMCT band or to a S(Cys) → Cu LMCT band from the π_u state (see below). Instead, the position and the intensity of the near infrared band are variable in different Cu_A sites. Solomon and coworkers have shown that this band reflects the electronic coupling between the two copper ions (H_{AB}), and therefore can be taken as an indicator of the strength of the Cu–Cu bond [86,

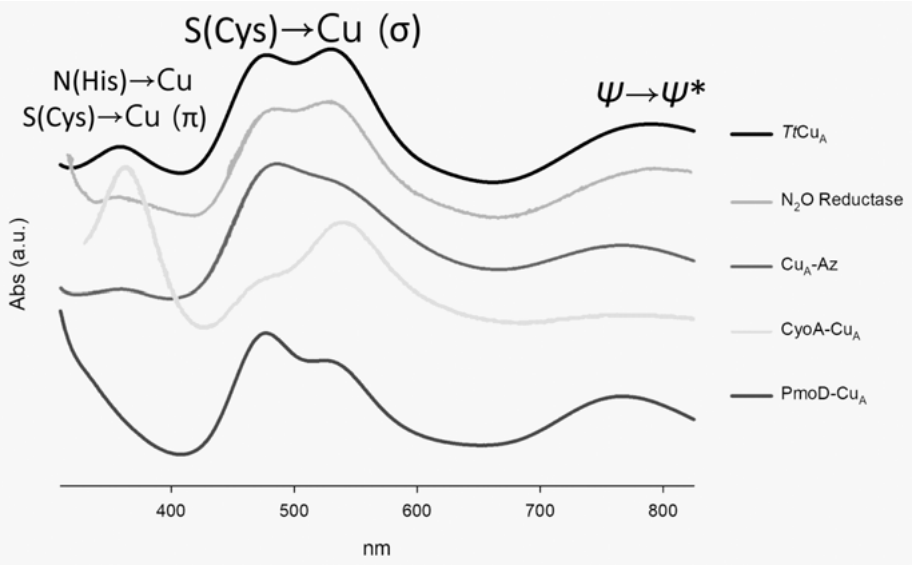


Figure 6. UV-vis spectra of Cu_A centers in native and model proteins. The Cu_A site from *Thermus thermophilus* soluble COX II domain from the ba3 COX (*Tt*Cu_A). The Cu_A center of *Pseudomonas stutzeri* nitrous oxide reductase. Adapted from [20] with permission from Springer Nature, copyright 2001. The engineered CuA centers of Cu_A-Az and CyoA-Cu_A are adapted from [124] and [37], by permission of Springer Nature and Wiley; copyright 2010 and 1992, respectively. The PmoD Cu_A center is adapted from [29] by permission from the American Chemical Society; copyright 2019.

122]. Indeed, recent data on Cu_A variants have shown a correlation between H_{AB} and the Cu–Cu distance as determined by EXAFS [123].

Naturally, the existence of a Cu–Cu metal bond in Cu_A has been a matter of intense discussion [56]. EXAFS experiments performed by Blackburn and associates provided the first evidence of a short Cu-Cu distance (2.5 Å) in a bacterial COX II before the availability of a crystal structure [56]. This result was later confirmed by data obtained for bovine heart COX [125].

The unique EPR spectrum of this center originates from a S = 1/2 paramagnet with similar hyperfine couplings to two copper nuclei (I = 3/2), that results in a seven-line splitting at the $g_{||}$ region [74, 126]. This spectroscopic feature indicates a fully delocalized mixed-valent species of class III, according to the Robin and Day classification [2]. Compared to centers in blue T1 Cu proteins, Cu_A centers show even smaller $A_{||}$ on the basis of EPR simulations [72], reflecting the greater covalent interaction and delocalization of the unpaired electron between the copper ions and the bridging Cys residues. The initial EPR evidence on the Cu_A-only N_2OR variant allowed Kroneck, Zumft, Antholine, and coworkers to confirm the binuclear structure of this center before the availability of any crystal structure and to extend these results to the Cu_A center in COX's [72, 73, 119, 125].

EPR spectra of N_2OR selectively enriched with ^{63}Cu or ^{65}Cu performed by Neese, Kroneck, and coworkers later confirmed the mixed-valence delocalized

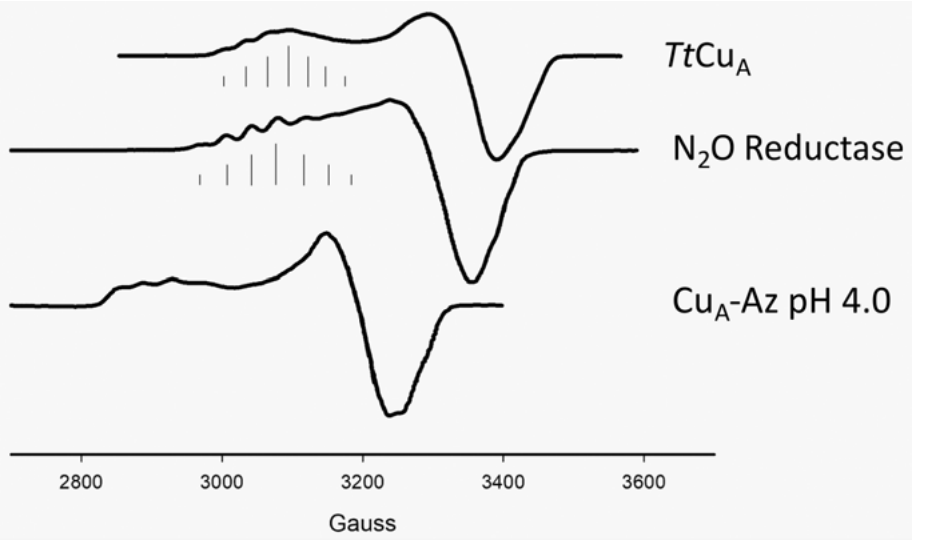


Figure 7. X-band EPR of *Tt*Cu$_A$ (top) adapted from [126] with permission from Elsevier; copyright 1996. Nitrous oxide reductase (middle) showing the typical seven-line hyperfine structure in the $g_{||}$ zone typical of a mixed-valence site, adapted from [20] with permission from Springer Nature, copyright 2001. Cu$_A$-azurin at pH 4 (bottom) lacks this spectral feature due to charge localization on one Cu atom. Adapted from [130] with permission of the National Academy of Sciences, USA, copyright 2004.

structure [127]. Quantum mechanical calculations with the experimentally derived g-tensor provided the first molecular orbital description of Cu$_A$, enabling the assignment of the absorption bands and the identification of a σ_u^* ground state and a low-lying π_u excited state [48, 123]. These states refer to the antibonding σ interaction between the $d_{x^2-y^2}$ copper orbitals, and the bonding π interaction between the d_{xy} copper orbitals, which were originally described as $^2B_{3u}$ and $^2B_{2u}$, respectively, based on symmetry considerations. The Solomon group later integrated these pictures with RR and XAS spectroscopies, providing a detailed description of the electronic structure of this complex metal center [85, 86, 126, 127].

Indeed, ligand K-edge and metal L-edge X-ray absorption spectra reported by Solomon and coworkers revealed that the involvement of the sulfur 3p orbitals in the RAMO (redox active molecular orbital) is 46 %, while the 3d orbitals of the copper ion participate with 44 % [128, 129] This highly delocalized nature has been suggested to favor long range electron transfer by activation of specific superexchange pathways and low reorganization energy upon redox changes.

NMR spectroscopy has provided a complementary description of the Cu$_A$ centers by being able to map the electron spin density in all nuclei from the copper ligands. This has confirmed the high electron spin delocalization into the Cys ligands, and a much less delocalization onto the His ligands (Figure 8) [49, 131–135]. The existence of electron spin density onto axial ligands was assessed in the case of mutants at the axial position [136], but recent experiments on selec-

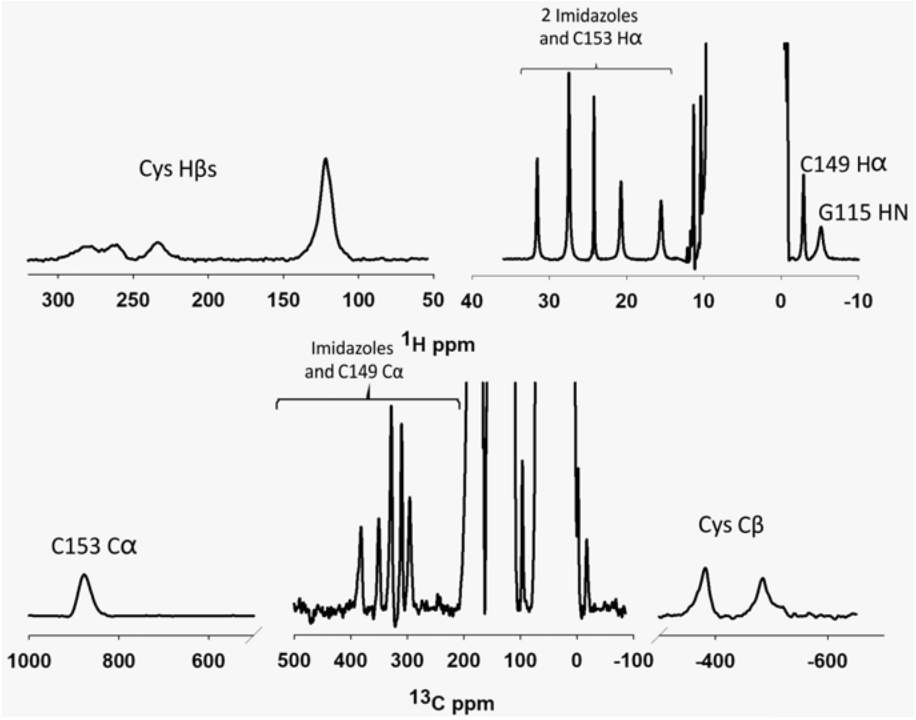


Figure 8. ^{1}H (top) and ^{13}C (bottom) NMR spectra of the *Tt*Cu_A protein in the oxidized state [135]. Hyperfine-shifted signals of the Cu_A ligand residues can be assigned through reported spectra acquisition schemes and analysis. Adapted from [135] with permission from the American Chemical Society; copyright 2009.

tively labelled Met residues confirmed the presence of electron delocalization onto the axial ligand [137].

4.2. The Two-State Issue in Cu_A

There is a consensus in defining the σ_u^* level as the ground state for Cu_A. However, there has been a discrepancy regarding the energy gap between this level and a low-lying π_u excited state [85, 127, 135]. The unusual g value of the EPR spectrum can be accounted for by the spin-orbit coupling between these two levels with a splitting of ca. 5000 cm^{-1}. However, NMR spectra show an anomalous temperature dependence (signals deviating from the Curie magnetization law) [49, 134, 136] that can only be accounted for by a very fast thermal equilibrium between two states following a Boltzmann distribution. NMR data also suggest that the low-lying excited state is the π_u level, with an energy gap of 600 cm^{-1}, i.e., one order of magnitude lower than that found by EPR spectroscopy [135]. This apparent contradiction has been reconciled by a model proposed by Gorelsky, Solomon, and coworkers that describes the electronic structure of Cu_A as a double-well

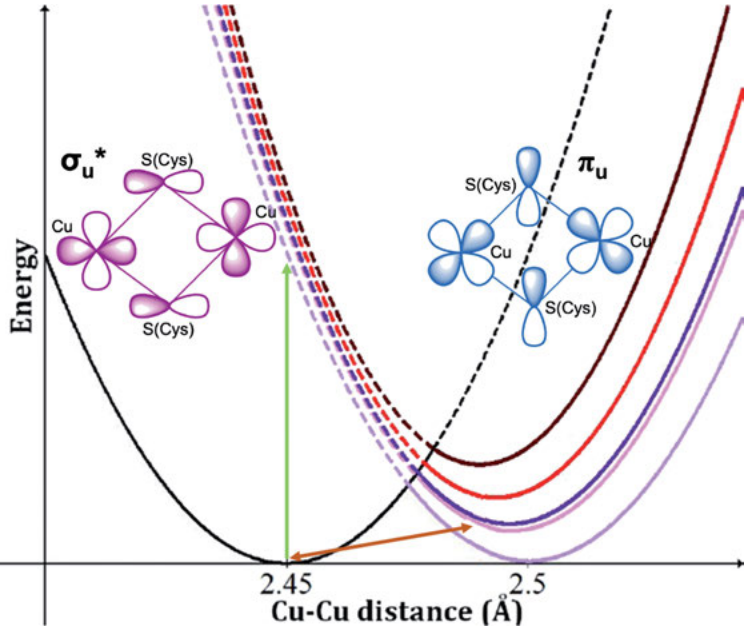


Figure 9. Energy profile scheme of the electronic structure of the Cu_A proteins, inspired by the model proposed by Solomon and coworkers [85]. The green arrow indicates the Frank-Condon energy transition observed by EPR and the orange arrow indicates the fast equilibria between the two alternative ground-state observed by NMR [123]. Adapted with permission from the American Chemical Society, copyright 2019.

potential with minima corresponding to the σ_u^* and π_u states (Figure 8) [85] that correspond to Cu–Cu distances differing by 0.5 Å. This assumption relies on data obtained from biomimetic models (Figure 3), with a copper–copper distance of 2.9 Å, as representative of a pure π_u ground state system [85, 86]. Both the Tolman and Duboc model [41, 42] lack the direct Cu–Cu bonding contribution to the absorption spectrum, resulting in a shift of the intervalence maximum to 1500 nm. Based on this, the NMR-detected equilibrium is supposed to sample a Cu–Cu bond oscillating between 2.5 Å and 2.9 Å, corresponding to the minima of the σ_u^* and π_u states, respectively. However, such an oscillation would not be consistent with the rigidity of this metal center. Recent experiments on Cu_A variants helped to refine this diagram, showing that in a protein framework the σ_u^* and π_u levels can be considered as alternative ground states with minimal changes in the Cu–Cu distance, since the H_{AB} value is preserved in all the studied mutants within the same energy range [49, 123, 135, 136, 138]. Indeed, different theoretical calculations revealed that minimal changes in the Cu_A geometry can induce drastic changes in this energy gap [127, 136, 139]. These results also showed that the band observed around 360–380 nm in the absorption spectrum of Cu_A centers (Figure 6) corresponds to a S(Cys) → Cu LMCT from the π_u state as the features at 480 and 530 nm from the σ_u^*, decrease. These two levels have been proposed to play a role in long range electron transfer (see below).

4.3. First and Second Sphere Modulation of the Electronic Structure

Site-directed mutagenesis was initially employed to validate the identity and essentiality of the different Cu_A ligands, and later mutagenesis experiments aimed to explore how mutations can tune the electronic structure and the redox properties of this site. Both Type 1 and Cu_A centers have unique electronic properties that make them very efficient in long range electron transfer. They both share a high spin delocalization to the Cys ligands due to a large Cu-S covalency, and to the two equatorial His residues that help defining the required site geometry [1, 52, 55, 140]. Both sites also share weak axial ligands completing the first coordination sphere. In the case of Type 1 copper sites, the identity of the axial ligand can differ among proteins [55] while in Cu_A sites two axial ligands are well preserved, a peptide carbonyl moiety and a Met residue side chain [54], with the only exception of PmoD [28].

Mutations on the Cu_A-CyoA protein by Saraste and coworkers helped to identify the essential Cu_A ligands [54]. Mutation on all proposed copper ligands led to loss of the purple color, while changing the nature of the residue that contributes with the carbonyl moiety did not affect the formation a purple Cu_A center [54]. Later, the Ludwig group showed that the Met axial ligand was not strictly required for the formation of a purple Cu_A site in *Paracoccus denitrificans* COX, but a Met to Ile replacement disrupted the fully delocalized mixed valence center [141].

The availability of soluble Cu_A proteins was key for the in depth investigation of this metal center. Yi Lu and coworkers performed an extensive investigation of the roles of the weak axial ligand in Cu_A-azurin (Cu_A-Az). Replacement of the axial Met in this protein induced minimal changes in the reduction potential [142], in contrast with similar mutations in the *Tt*Cu_A fragment (see below) [143]. In both cases, the introduction of stronger ligands in the axial position induced larger degrees of perturbations in the Cu_A site, even if the mixed valence character was preserved. Replacement of Met123 by Asp or Glu in Cu_A-Az increased the intensity of the 380 nm maximum, indicating that the π_u electronic state can be populated at room temperature [142] similar to what was observed later for the replacement of the Met160 ligand on *Tt*Cu_A by Ser or His [136, 143]. These results suggest that the presence of a stronger axial ligand can reduce the energy gap between the two alternative ground states and favor the population of the π_u state.

While the preservation of the first coordination sphere is essential for the modulation of the activity of the proteins with Cu_A sites, the second coordination sphere displays variations among the different proteins. While these residues can be involved in the optimization of the protein interaction with redox partners, it was recently shown that these residues are also involved in the fine tuning of the electronic structure. The engineered protein h*Tt*3L, in which three protein loops were replaced without altering the identity of any ligand, experienced a strong change in the energy gap between the σ_u^* and π_u levels [49]. Later, replacement of the individual loops (*Tt*1L-Ligands and *Tt*1L-Entry with the ligand loop and the loop of interaction with cytochrome *c*, respectively) indicated that each loop contributes individually in this modulation [60]. Overall, these data suggest

that second sphere residues do not only tune the reduction potential, but also the electronic structure of the metal site [49, 60]. The contributions of first and second sphere are independent, as it has been shown by a combination that led to a Cu_A site in which both σ_u^* and π_u states were almost degenerate [123].

5. REDOX THERMODYNAMICS

5.1. pH-Dependence of Reduction Potentials

The redox thermodynamics of Cu_A sites has been investigated for three different types of metalloproteins: (i) solubilized integral COX or N_2OR enzymes, (ii) soluble fragments of subunit II of COX (COX II) containing the native Cu_A site, and (iii) artificial Cu_A sites engineered into the cupredoxin scaffold of small soluble proteins.

Within the first group, spectroelectrochemical titrations of solubilized mitochondrial beef heart COX yield a reduction potential (E°) of $+285$ mV for the Cu_A site of the CO-poisoned enzyme that contains heme *a* in the Fe(III) form. This value decreases about 40 mV upon heme *a* reduction, thus indicating an anticooperative interaction between both metal sites [144, 145]. Similar experiments on the N_2OR from *Pseudomonas nautical* yield $E^\circ = 260$ mV at pH 7.6 [22].

Consistently, electrochemical determinations performed at room temperature, reveal reduction potentials from $+216$ to 297 mV at neutral or slightly alkaline pH for a variety of Cu_A-containing COX II fragments from *Pa. denitrificans* aa_3-COX [78], *T. thermophilus* ba_3-COX, [25, 33, 143, 146, 147], *Sulfolobus acidocaldarius* terminal respiratory supercomplex [83], *Pa. versutus* COX [132], and *Synechocystis* PCC 6803 aa_3-COX [35].

Also in good agreement, artificial Cu_A sites engineered into the cytochrome *bo* quinol oxidase of *Escherichia coli* [54], into *Paracoccus denitrificans* amicyanin [40] and into azurin from *Pseudomonas aeruginosa* [76, 77, 142] display E° values from 239 to 273 mV.

The average E° of the Cu_A site taken over all reported values for the three kinds of Cu_A sites described above is $+256 \pm 80$ mV. Given that the ligand set is preserved in all these protein variants, the 80 mV variation may be ascribed to differences in the protein environment, although, experimental conditions may also play a role.

The most extensively and systematically studied Cu_A proteins are the already discussed model protein Cu_A-Az and the soluble fragment $Tt Cu_A$. A great advantage of these two model proteins compared to others, such as COX II from *P. denitrificans* [148], is their thermal stability over a broad range of pH.

Hwang and Lu studied the pH-dependence of the electronic properties of Cu_A-Az [130]. Variation of the pH from 7 to 4 results in a drastic upshift of E° from $+160$ mV to $+340$ mV. This change is accompanied by small but distinct variations of the UV-vis spectrum and by drastic and reversible changes of the EPR spectra from the seven-line signal typical of fully delocalized (class III) mixed valence Cu_A (pH 7) to a four-line signal, similar to that observed mono-

nuclear T1 Cu centers (pH 4). These results were rationalized in terms of a pH-dependent transition between a delocalized class III state and a trapped valence state, probably class II, with an apparent pK_a value of 4.8. Site directed mutagenesis studies revealed the equatorial ligand His120 as the most likely protonation site [130]. Based on these results, the authors put forward a novel feedback inhibition mechanism for the proton-coupled ET activity of COX. In this mechanism the equatorial His ligand of the native Cu_A center in COX, which is involved in the pathway from Cu_A to heme *a* [15], is regarded as an electron relay that regulates the ET reaction. When this His residue is deprotonated, the Cu_A center is in the fully delocalized form and has a reduction potential low enough to efficiently transfer one electron to heme *a*. The ET-driven proton pumping activity of COX is postulated to result in protonation and subsequent detachment of this crucial His residue, thus inducing a drastic upshift of the reduction potential that would impair subsequent ET steps. The proposed feedback mechanism assumes that under these conditions, the Cu_A site would remain in standby until protons are released into the outer membrane space and the His residue is again deprotonated [27, 130].

In a subsequent combined spectroscopic and theoretical study [59] the authors discovered that the class III character of Cu_A-Az actually remains preserved at low pH and, instead, the four-line hyperfine EPR pattern was ascribed to a ca. 1 % 4s orbital contribution from one of the copper ions to the ground state spin wavefunction due to protonation and detachment of His120. In spite of the different interpretation, the authors suggested that the drastic change of E° with pH, as well as possible variations of the reorganization energy and the electronic coupling upon His protonation, may play a central role in regulating the function of COX [59].

In sharp contrast to the Cu_A-Az construct, the UV-vis absorption and resonance Raman spectra of the native *Tt*Cu_A site are insensitive to pH over a wide range from 2 to 7. ^{1}H-NMR titrations of *Tt*Cu_A, on the other hand, afford three reversible acid-base transitions. The first one has a pK_a value of 3.5 and corresponds to the disappearance of the imidazolic $N^\varepsilon H$ signal from the equatorial His157 ligand (equivalent to His120 in Cu_A-Az) at 23.2 ppm. The remaining resonances of the H157 imidazole ring remain unperturbed, thus indicating that this transition involves neither protonation of H157 nor its detachment from the metal center [149]. The loss of the $N^\varepsilon H$ signal at low pH is assigned to a faster proton exchange with the solvent of the partially exposed H157. A second transition, which is not paralleled by changes of the UV-vis or RR spectra, has pK_a = 4.7 and corresponds to changes in the resonances of His114 imidazole protons and the peptide proton of the second shell ligand Gly115. The third NMR detected transition with pK_a = 11 is ascribed to the collapse of the Cu_2S_2 core.

Notably, while the reduction potential of Cu_A-Az shows a sharp drop of more than 150 mV from pH 7 to 4, this variation is only about 15 mV for the native *Tt*Cu_A [130, 149]. Furthermore, the reduction potential of *Tt*Cu_A shows a low-amplitude transition at pH 3 followed by a small and nearly continuous variation at higher pH (Figure 10) that reflects the existence of several acid-base equilibria with small or no impact on the electronic structure.

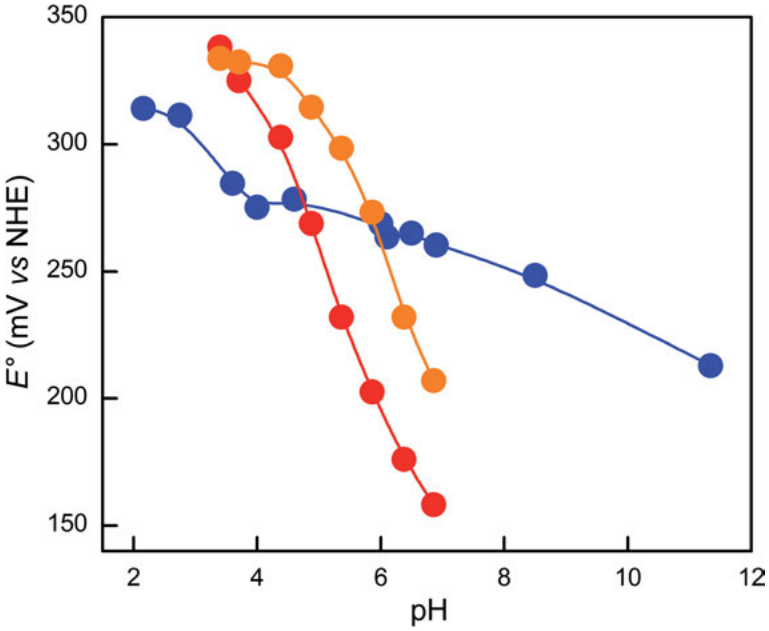


Figure 10. Reduction potentials as a function of pH. Blue: wild-type *Tt*Cu_A. Red: Cu_A-Az. Orange: H120A mutant of Cu_A-Az. The data for *Tt*CuA and Cu_A-Az are taken from [149] and [130], respectively.

Consistent with these results, Abriata and Vila reported a network of residues located around the electron entry point of COX, the Cu_A site in subunit II, that experience collective pH equilibria around neutral pH. This network starts at the occluded side of the Cu_A site and extends to the interface between subunits I and II of COX, where the proton exit is located and through which electrons flow into subunit I [63].

Moreover, eight different variants of *Tt*Cu_A, generated either by single point mutation or by loop engineering, display upshifts of only 8–58 mV upon lowering pH from 7 to 3.5 [60, 150]. The sharp contrast between the pH-dependencies of the spectroscopic and electrochemical features of the native *Tt*Cu_A site and the Cu_A-Az construct challenge the notion of redox properties of the primary electron acceptor of COX being fine-tuned by local pH variations *in vivo*. Most likely the differences between the two proteins are related to the presence of a protonable residue, His35, very close to the metal site of Cu_A-Az.

Interestingly, the enthalpic and entropic contributions to E° measured at acidic pH for the eight *Tt*Cu_A variants afford a good linear relationship with a slope of –2.74 (Figure 11), thus indicating partial enthalpy/entropy compensation. Similar results are obtained at neutral pH, except that in this case the variants Met160His and h*Tt*3L-Met160His fall outside of the linear correlation. These results are consistent with the observation that for Met160His and h*Tt*3L-Met160His, but not for the other *Tt*Cu_A variants, $\Delta E_{\sigma_u^*/\pi_u}$ is strongly pH-depend-

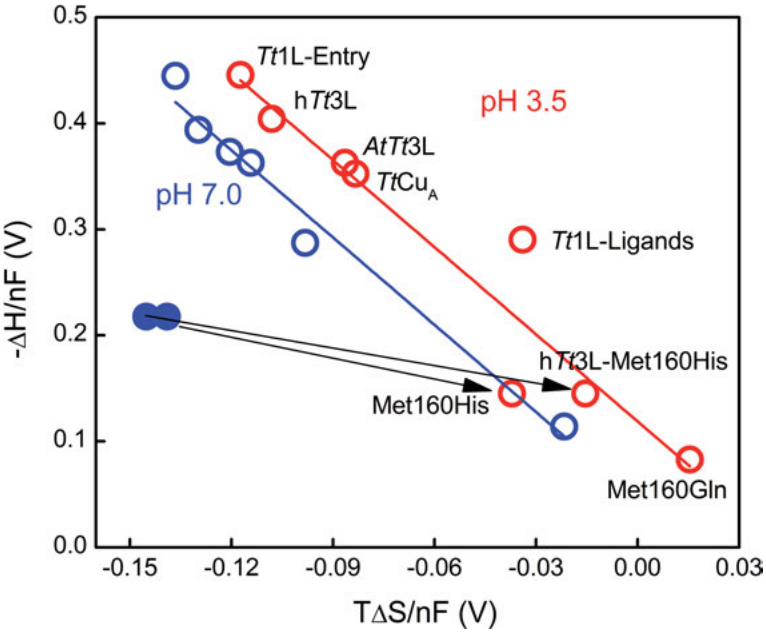


Figure 11. Enthalpic *versus* entropic contributions to the reduction potential for six different variants of $Tt\text{Cu}_\text{A}$ measured electrochemically at pH 3.5 (red) and 7.0 (blue).

ent, such that at pH 3.5 the GS in all cases is predominantly σ_u^* and becomes predominantly π_u at pH 7.0 only for Met160His and h*Tt*3L-Met160His.

5.2. First and Second Sphere Ligand Modulation of the Reduction Potentials

The impact of the first coordination sphere amino acids on the Cu_A electronic and redox properties is a key issue that has been addressed by several groups. Gamelin et al. [86] found out that the thiolate ligands should strongly stabilize the oxidized form to an extent that, in comparison with blue T1 centers, anticipates a much lower reduction potential than the average 256 mV value determined for Cu_A sites in general. This discrepancy suggests that delocalization between the two copper ions rise the reduction potential relative to a trapped mixed-valence dimer.

Mutation of the strong equatorial His ligands has been shown to exert a strong impact on the electronic properties of Cu_A. In the Cu_A-Az, replacement of His46 results in disruption of the metal site [58]. In contrast, mutation of His120 by Asn, Asp, Ala, and Gly is not detrimental to the formation of the binuclear site, although spectroscopic data suggest modulation of the Cu–Cu distance [58] and either transformation into a trapped valence state [150] or a small 4s orbital contribution to the ground-state wave function that results in a significant change

of the EPR spectra while preserving valence delocalization [59]. Reduction potentials for these Cu_A-Az variants are not available, but the equivalent His260Asn exchange in the Cu_A site of the integral COX from *Rhodobacter sphaeroides* results in a 90 mV upshift of $E°$ and in a distorted EPR spectrum, whose origin is not completely clear [151, 152].

The role of the weak axial ligands in modulating the redox potential of Cu_A-Az has been found to be very small. On the one hand, esterification of the backbone carbonyl of Glu114 (weak axial ligand) does not affect $E°$, although it alters the absorption spectrum to some degree [79]. On the other, mutation of Met123 by Leu, Asp, and Glu results in a minor variation of $E°$, between 5 and 16 mV, in spite of the fact that in the Met123Leu mutant, but not for the other two variants, EPR spectra suggest valence localization [142]. Similar mutations in the blue T1 Cu protein azurin result in significantly stronger modulation of $E°$, which was rationalized in terms of the slightly shorter Cu-S(Met) distance in Cu_A-Az that renders this center less sensitive to perturbations on the axial Met ligand compared to the T1 site (Figure 12) [142].

In sharp contrast, mutation of the equivalent Met residues in the native Cu_A centers of the integral COX from *R. sphaeroides* [152] and *Pa. denitrificans* [141] by Leu and Ile result in $E°$ upshifts of 120 and 100 mV, respectively, thus highlighting the relevance of the overall protein scaffold in determining the influence of specific ligands. Systematic mutation of the equivalent Met160 in the soluble *Tt*Cu_A fragment confirms this view, as it results in strong modulation of $E°$ (Figure 12) with preservation of the mixed valence delocalization in all the mutants [142]. As expected, the more conservative mutation of Met160 by a selenomethionine has no significant effect on the electronic properties of *Tt*Cu_A [153].

Modulation of the reduction potential of the Cu_A site is not restricted to the first coordination sphere ligands. Asn47 and Glu114 are two residues in the vicinity of the bridging Cys112 of Cu_A-Az that do not directly interact with the metal ions but participate in H-bonding to Cys112. As shown by New et al. [154], the reduction potentials of the mutants Asn47Ser and Glu114Pro of Cu_A-Az are upshifted and downshifted with respect to the unmodified protein by 30 and 42 mV, respectively. A similar trend but of higher magnitude was found by the same group for equivalent mutations in the T1 blue copper azurin [155]. Notably, the effect of mutating second sphere ligands on the reduction potential of Cu_A-Az more than doubles the effect of mutating the first sphere ligand Met120. Based on the different signs of the shifts in the two mutants, the authors speculate that the modulation cannot be ascribed solely to Coulomb effects and, instead, must also include the effect of hydrophobic and H-bonding interactions. More recent studies based on loop engineering of *Tt*Cu_A provided deeper insight into the regulation of the reduction potential by outer sphere amino acids [49, 60]. In these studies, four different chimeras were generated by replacing either one or the three loops that define the nearby environment of the metal site, but preserving in all cases the set of ligand residues. In the so-called h*Tt*3L and *AtTt*3L chimeras the three loops are replaced by the *Homo sapiens* and *Arabidopsis thaliana* sequences, respectively. The variants *Tt*1L-Entry and *Tt*1L-Ligands have only the entry loop 86–89 and the ligand loop 149–160, re-

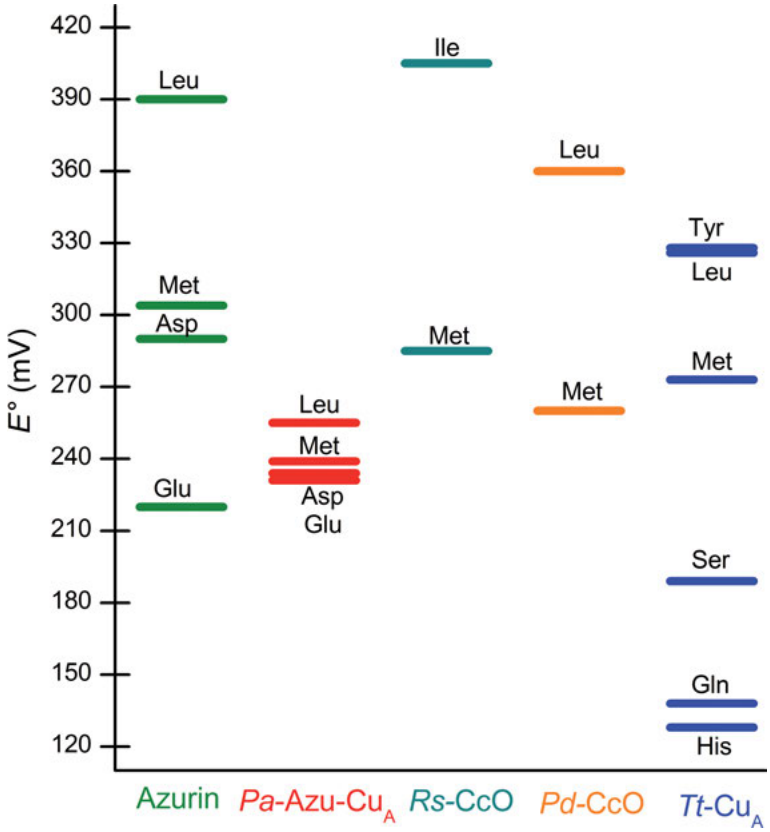


Figure 12. Modulation of the reduction potentials of T1 and Cu_A centers upon replacement of the weak Met ligand by a different amino acid, as indicated in each case. Green: Met120 variants of the T1 Cu protein azurin from *Pseudomonas aeruginosa* [142]. Red: Met120 variants of Cu_A-Az [147]. Cyan: Met263 variants of cytochrome *c* oxidase from *Rhodobacter sphaeroides* [152]. Orange: Met227 variants of cytochrome c oxidase from *Paracoccus denitrificans* [141]. Blue: Met160 variants of the *Tt*Cu_A soluble fragment [143].

spectively, replaced by the homologous *Homo sapiens* sequences. The different outer sphere permutations fine tune the reduction potential by up to ± 35 mV (Figure 13A), which is a significantly softer modulation than that achieved through Met160 replacement (Figure 12). Interestingly, the E° values of the chimeras correlate with the total hydrophobicity index of the metal site environment, calculated as the sum of hydrophobicity indexes of all residues belonging to the three surrounding loops. ΔH° and ΔS° values for reduction are negative for all the investigated variants, thus arguing against a simple Coulomb effect. Moreover, the enthalpy and entropy contributions to E° vary linearly with the total hydrophobicity index, partially compensating each other and, thereby, softening the variation of E° (Figure 13B).

Interestingly, the reduction potential of h*Tt*3L-Met160His, i.e., the Met160His variant of the h*Tt*3L chimera, is 74 mV at pH 7.0. This value is 187mV lower

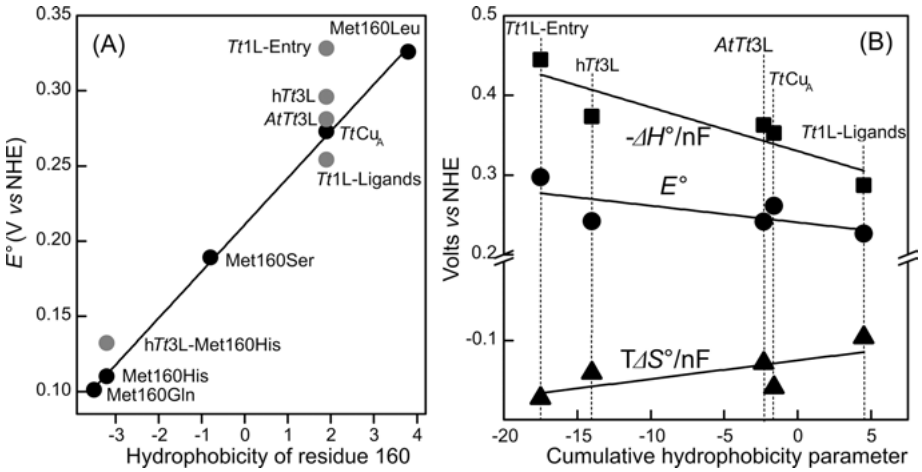


Figure 13. (**A**) Reduction potentials of the different *Tt*Cu_A variants as a function of the hydrophobicity index of Met160. Black circles are Met160 single point mutants, while grey symbols are chimeras with either one or the three loops surrounding the metal site replaced [49, 88, 123, 143, 156] (**B**) Enthalpic and entropic contributions to the reduction potential as a function of the hydrophobicity for *Tt*Cu_A chimeras where either one or the three loops surrounding the metal site have been replaced, while preserving the ligands set [60]. Adapted with permission from the American Chemical Society, copyright 2019.

than for *Tt*Cu_A, 5 mV lower than for Met160His and 168 mV lower than for h*Tt*3L, thus showing some degree of additivity of the effects of first and second sphere mutations on the redox potentials [123, 148].

5.3. The Reduction Potential and the Electronic Partition Function

While the reduction potentials and the enthalpies of reduction of the different point mutants and chimeric forms of *Tt*Cu_A correlate with the hydrophobicity of residue 160 (Figure 13A), the reduction entropy for the same series correlates with the size of the energy gap $\Delta E_{\sigma_u^*/\pi_u}$, which varies not only with the mutations but also with the pH for some of these proteins, (Figure 14) [60, 148]. This correlation has been interpreted in terms of the electronic partition function, which in turn is determined by the degeneracies of the reduced and oxidized forms of the metal site [60, 148]. The oxidized Cu_A is a fully delocalized mixed-valent center with two alternative ground states, each of them being a doublet. Therefore, the degeneracy of the mixed-valent complex may adopt extreme values of 2 and 4, depending on the thermal accessibility of the π_u state given by the ratio $\Delta E_{\sigma_u^*/\pi_u}/k_B T$. The reduced Cu_A, in contrast, is a singlet and has a one-well GS description with σ_u^* HOMO symmetry and, thereby, is a nondegenerate state. Taking these multiplicities into account, the reduction entropy can be expressed in terms of Equation (1):

$$\Delta S = -k_B ln\left(2 + 2e^{-}\Delta E_{\sigma_u^*/\pi_u} / k_B T\right) + k_B ln\left(Q^*_{Red} / Q^*_{Ox}\right) \quad (1)$$

where the first term accounts for the difference in the electronic part of the partition function between the reduced and oxidized forms, and the second terms does the same for the non-electronic components of the partition function.

This relation predicts that the electronic contribution to the reduction entropy will be negative for all *Tt*Cu_A variants, and that this number becomes more negative upon decreasing $\Delta E_{\sigma_u^*/\pi_u}$, as experimentally verified (Figure 14). The impact of the accessibility and multiplicity of the alternative GS on ΔS^o has been found to account for differences in the E^o values between Cu_A variants of up to 100 mV [60, 148].

6. ELECTRON TRANSFER DYNAMICS

6.1. Fine-Tuning of the Reorganization Energy

Pioneering kinetic studies [15, 157] and comparisons with the better known T1 Cu centers [158] during the early 1990s anticipated that the ET reorganization energy (λ) of Cu_A should be a small number between 0.15 and 0.5 eV. Notably,

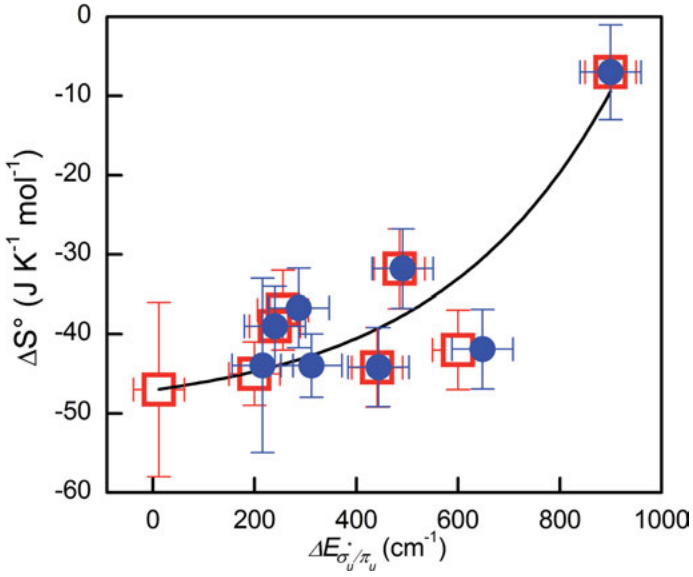


Figure 14. Reduction entropy as a function of σ_u^*/π_u, the energy gap measured at pH 3.5 (blue) and 7.0 (red) for eight different variants of *Tt*Cu_A: *Tt*Cu_A, Met160His, Met160Gln, h*Tt*3L, *AtTt*3L, *Tt*1L-Ligands, *Tt*1L-Entry and h*Tt*3L-Met160His [60, 148]. Adapted with permission from the American Chemical Society, copyright 2019.

these successful predictions were formulated several years before the Cu_A structure and experimental λ determinations became available.

From the point of view of λ, the Cu_A site has been thought of as a dimer of mononuclear T1 Cu sites [26, 86, 158] where one of the equatorial His ligands of each of the monomers is replaced by a Cu–Cu bond. The strong metal-metal electronic coupling in Cu_A has been argued to overcome vibronic coupling, thus leading to a highly covalent delocalized ground state wavefunction. Such delocalization is expected to distribute the structural changes associated to the ET reaction in twice the number of chemical bonds compared to the monomer (or to the localized valence binuclear complex) and, therefore, the inner sphere reorganization energy (λ_{in}) of Cu_A is expected to be about one half of that for T1 Cu sites [26, 86, 158]. The outer sphere reorganization energy (λ_{out}) of Cu_A is also expected to be lower than for T1 Cu sites as a consequence of charge delocalization over the dimer that is expected to diminish solvent rearrangement in the vicinity of the metal site [86, 158].

In very good agreement with these predictions, the experimental determination of Cu_A-Az yields λ = 0.4 eV [159, 160], i.e., about half of the value of 0.82 eV reported for wild-type azurin [156]. Other experimental determinations produced similarly small λ values of 0.3 eV, 0.4 eV, and 0.3–0.4 eV for Cu_A-Az [146], bovine COX [157], and *Tt*Cu_A [60, 136, 149], respectively. The small reorganization energies are consistent with the observation that the structures of the oxidized and reduced Cu_A centers from *T. thermophilus* and *B. subtilis* are virtually identical, as revealed by XAS studies [136, 161].

Moreover, quantum mechanical calculations predict values of 0.22 to 0.32 eV for λ_{in} when the ET reaction proceeds from the σ_u^* GS. This energy, however, is calculated to increase by about a factor of two when the ET reaction takes place from the alternative π_u GS [85, 86, 136]. The difference, which might be attributable to partial localization of the Cu-S(Cys) covalency in two shorter and two longer bonds in the π_u GS, has been an argument to regard the π_u GS as redox-inactive [85, 86].

Recent electrochemical determinations of λ for eight different *Tt*Cu_A variants, however, gave a different picture. In addition to wild-type *Tt*Cu_A, protein film voltammetry experiments were conducted for single variants of the weak axial ligand Met160 and for chimeras in which the loops surrounding the metal site were exchanged by homologous sequences from other organisms but preserving the length of the loops and the set of coordinating residues, as well as for Met160 variants of these chimeras [49, 60, 88, 123, 136]. Interestingly, the λ values of the eight protein variants can be grouped into two well defined clusters (Figure 15). The first cluster includes wild type *Tt*Cu_A, as well as variants Met160Gln and *Tt*1*L*-Ligands. The average λ for this group is $\langle\lambda\rangle = 0.27 \pm 0.05$ eV, and energy gaps vary between 440 and 900 cm^{-1}. The second cluster, integrated by *At*3L, h*Tt*3L, and h*Tt*3L-Met160His, is characterized by $\langle\lambda\rangle = 0.53 \pm 0.05$ eV and smaller $\Delta E_{\sigma_u^*/\pi_u}$ values, from 12 to 256 cm^{-1}. Note that these two average λ values closely resemble those determined by QM calculations for ET from the σ_u^* and π_u alternative GS, respectively [85, 86, 136], thus suggesting that the electrochemical experiments selectively probe one or the other RAMO, depending on the size of $\Delta E_{\sigma_u^*/\pi_u}$.

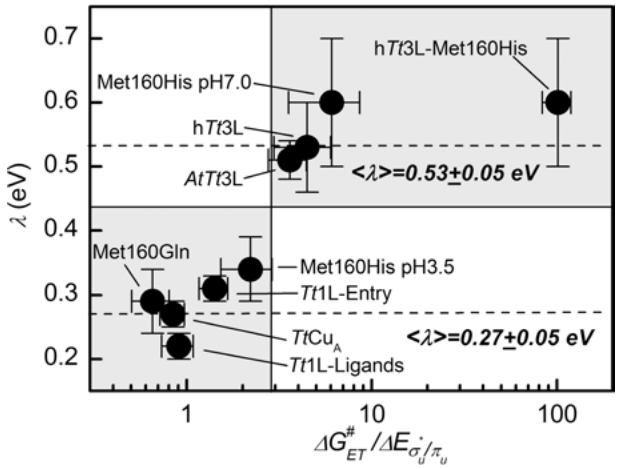


Figure 15. Electron transfer reorganization energies of eight *Tt*Cu_A protein variants as a function of the ratio between the ET activation free energy and σ_u^*/π_u, the energy gap. All values correspond to pH 4.6, except for variant Met160His that is reported at two different pH values and variant h*Tt*3L-Met160His that is reported at neutral pH [60, 88, 162]. Adapted with permission from the American Chemical Society, copyright 2019.

In very good agreement with this interpretation, the Met160His variant falls into one or the other cluster depending on pH (Figure 15). The UV-vis, EPR, RR, and NMR spectra of the Met160His variant clearly show a reversible pH-dependent modulation of $\Delta E_{\sigma_u^*/\pi_u}$ with preservation of the structure and class III MV characteristics. These spectroscopic features correlate with a reversible increase of λ from 0.34 eV at low pH to 0.60 eV at high pH (Figure 16), which is paralleled by an unexpected increase of the ET rate constant, thus suggesting a 7.5-fold increase of the electronic coupling for the heterogeneous ET reaction from the π_u RAMO compared to the σ_u^* GS [88].

The observation that the heterogeneous ET reaction apparently occurs either from the σ_u^* or the π_u alternative GS depending on the thermal accessibility of the higher lying π_u GS can be readily understood provided that the reaction coordinates for the ET reaction and for the $\sigma_u^* \rightarrow \pi_u$ conversion are strongly coupled. A number of experimental observations point out that this might indeed be the case. EXAFS data show that the Cu–Cu distance is slightly different in the reduced and oxidized forms of wild-type *Tt*Cu_A and point mutants, thus suggesting that this stretching may be a crucial component of the ET reaction coordinate [136]. This distance has also been proposed as the main coordinate for the adiabatic $\sigma_u^* \rightarrow \pi_u$ transition, but the difference in bond lengths between the two alternative oxidized states was assumed to be about 0.6 Å, based on experimental and computational studies on model complexes [41, 85, 86, 139]. The assignment of the Cu-Cu distance as a crucial determinant of the σ_u^*/π_u relative populations has been recently confirmed for native Cu_A. In this case, however, the experimental structural and spectroscopic evidence indicates

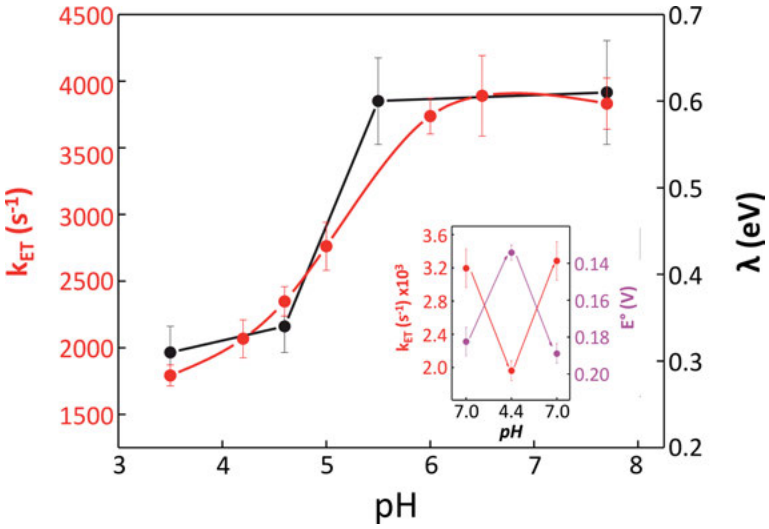


Figure 16. Variation of the heterogeneous electron transfer rate constant and of the reorganization energy of the Met160His mutant of *Tt*Cu_A as a function of pH. The inset shows the reversibility of pH-induced changes of k_{ET} and $E°$. Reprinted from [88] with permission from Wiley, copyright 2015.

that the difference in Cu–Cu bond lengths between the σ_u^* and π_u states is only 0.05 Å [123], i.e., a small variation that is compatible with this stretching being a common crucial reaction coordinate for both ET and $\sigma_u^* \rightarrow \pi_u$ conversion. Thus, under the hypothesis that these two processes are indeed entangled, one can anticipate that for those protein variants with $\Delta E_{\sigma_u^*/\pi_u}$ values larger than the ET activation free energy ($\Delta G_{ET}^{\#}$) the reaction will proceed from the lower-lying σ_u^* RAMO. In contrast, as qualitatively verified in Figure 15, the reaction will proceed through the alternative π_u RAMO, when $\Delta E_{\sigma_u^*/\pi_u} < \Delta G_{ET}^{\#}$.

6.2. Electron Transfer Pathways

The availability in the 1990's of the first crystal structures of COX from bovine heart [14, 46] and *Pa. denitrificans* [45] prompted the investigation of possible pathways for the $Cu_A \rightarrow$ heme *a* intramolecular ET reaction. The first studies [15, 163, 164] identified an efficient pathway that starts at the equatorial ligand His204 (bovine heart COX numbering), continues across one hydrogen bond to Arg438 and finally trough another H-bond to the heme *a* propionate. A slight variation of this pathway involves Arg439 instead of Arg438 (Figure 17) [165]. Latter investigations identified alternative pathways that connect either of the bridging ligands, Cys200 and Cys196, with the heme *a* propionate (Figure 17) [86, 129, 165]. Even though the Cys pathway involves about 10 more bonds than the His204 route, it was proposed to be similarly efficient [86] or even more efficient

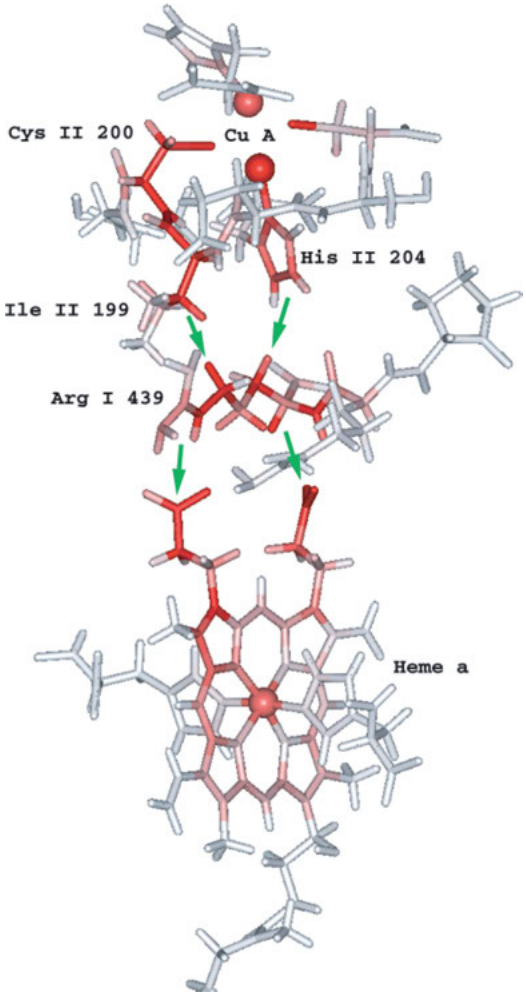


Figure 17. Intramolecular $Cu_A \rightarrow$ heme *a* electron transfer pathways proposed for bovine heart cytochrome *c* oxidase. Reprinted from [165] with permission from the American Chemical Society, copyright 2000.

[165]. This inference is based on the high covalency of the Cu–S(Cys) bond that was found to be about 13 % per S(Cys) but only 1–3 % for the Cu-N(His) interactions [129]. One should note, however, that these magnitudes refer only to the mixed-valent and not to the reduced form of Cu_A. Later, calculations of effective couplings combined with all-atom molecular dynamics of the membrane-embedded COX indicated that the coupling between Cu_A and heme *a* is dominated by the His204 pathway [166]. Also for the ba_3 oxidase from *T. thermophilus* pathways calculations identified the equivalent His157 pathway as the most efficient in mediating $Cu_A \rightarrow$ heme *b* ET [136]. Quantum mechanics-molecular mechanics

electron pathway calculations on the COX from *R. sphaeroides*, on the other hand, gave preeminence to the Cys 252 route over the His260 one [16], thus emphasizing that this issue remains controversial.

The intermolecular ET step from cytochrome *c* (Cyt-c) to Cu_A is more difficult to analyze as it requires good knowledge of the Cyt-c/COX complex structure. Numerous investigations focused in elucidating the interaction interface between the two proteins employing a variety of chemical, biochemical, spectroscopic and kinetic strategies have been published [167, 168]. The different studies revealed that the proximal His ligand of Cu_A is not surface-exposed and, therefore, cannot directly contact with the partially exposed heme edge of the Cyt-c electron donor. Instead, detailed studies on mammalian and *Pa. denitrificants* COX provided evidence of the involvement of a hydrophobic cluster around Cu_A as the dominant interaction site for correct positioning of the redox centers, with Trp104 as the most likely electron entry point in COX [45, 169–173]. Based on these observations, it was proposed that the Cyt-c $\rightarrow$ Cu_A electron pathway-comprises a through-space heme *c*-Trp104 step, followed of a H-bond element between Trp104 and the bridging ligand Cys200. It was argued that the high covalency of the Cu–S(Cys) bond would render this route more efficient than other possible pathways [86, 129]. Very recently, Shimada et. al reported a high resolution structure of the mammalian Cyt-c/COX complex [174]. The structure confirms, to a large extent, the contacts predicted before by indirect methods, but the Cyt-c $\rightarrow$ Cu_A pathway that emerges from the crystal structure does not include Trp104. Instead, the main pathway implicates Tyr105 as electron entry point, followed by a through-space jump to the weak axial ligand Met207. A similar route was predicted before based on pathways analysis of the Cyt-c_{552}/*Tt*-Cu_A complex structure obtained by a combination of NMR spectroscopy and docking calculations [61]. In this case the preferred pathway for the Cyt-c_{552} $\rightarrow$ Cu_A ET involves Phe88 and Met160 [136]. Interestingly, the Cu-S(Met) low covalency was found to be 10 times stronger in the π_u alternative GS relative to σ_u^*. For the subsequent Cu_A $\rightarrow$ heme *b* ET step the preferred pathway involves the sequence His157-Arg450-heme *b*. The covalency of the Cu–N(His157) bond is doubled in the σ_u^* state compared to the π_u state. Taken together, these results suggest that the alternative π_u and σ_u^* ground states of Cu_A might be involved in effectively coupling upstream and downstream ET with high directionality (Figure 18) [136]. Later, calculations suggested the absence of S(Met) character in the two alternative GS wave functions [80].

The caa_3 oxidase from *T. thermophilus* has a cytochrome *c* domain fused to subunit COX II [175], so that electron flow into the Cu_A site is intramolecular, and differs from the orientation reported by the NMR experiments. The optimal electron tunneling pathway (18.8 Å long) suggests that the Met ligand is the electron entry port to Cu_A. Finally, a crystal structure of the cytochrome *c*-COX bovine complex revealed a distance between the iron atom of heme *c* and the copper atom of Cu_A of 23.0 Å [174], with the most probable ET pathway reaching the copper site by the axial Met ligand. Recent NMR experiments on a selectively labelled Met residue help identifying net electron spin density onto the axial Met [137], giving a strong support to this proposed ET pathway.

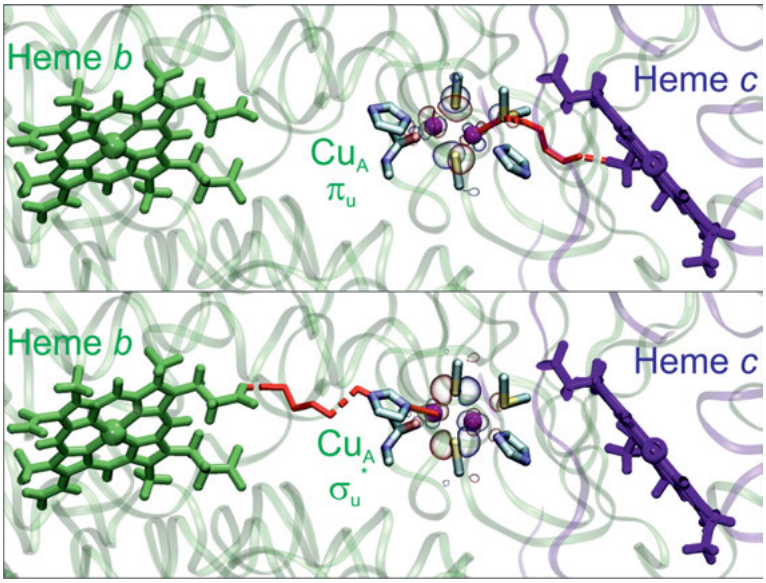


Figure 18. Proposed intermolecular Cyt-c → Cu_A (top) and intramolecular Cu_A → heme *b* (bottom) electron transfer pathways for the complex cytochrome c_{552} / ba_3-COX from *Thermus thermophilus*. The first step involves the alternative π_u GS of the mixed-valent Cu_A, and Met160 as electron entry point. The second reaction takes place from the σ_u^* GS of reduced Cu_A and implicates His157 as electron exit point. Adapted from [136] with permission of the National Academy of Sciences, USA, copyright 2012.

6.3. Regulation of the Electron Transfer Reaction

Electroprotonic energy transduction processes in biology, such as respiration and photosynthesis, rely upon complex protein-based redox machineries that utilize the energy provided by a cascade of downhill ET reactions to drive uphill proton translocation across a biological membrane. The energy stored in this electrochemical transmembrane potential is ultimately used for synthesizing ATP. Not surprisingly, the multiple inter- and intramolecular ET steps involved are characterized by very low driving forces, as a prerequisite for maximizing the mechanochemical work required for proton translocation and to avoid excessive heat release. Moreover, the complex architecture of these systems impose long electron donor-acceptor distances, typically well above 10 Å. The Cu_A electron hub of COX is not an exception; driving forces are only a few mV and distances are about 18–20 Å for both the upstream and downstream ET processes [61, 168, 172–174]. In spite of these unfavorable conditions, ET rates to and from Cu_A have been found to be fast [176–180], which can only be achieved by optimizing pathways that efficiently mediate electron superexchange, and by minimizing the associated reorganization energies [181]. It remains unclear, though, how the activity of COX and of the entire respiratory chain in general is regulated to translocate protons only when it is actually needed, avoiding detrimental overshooting of

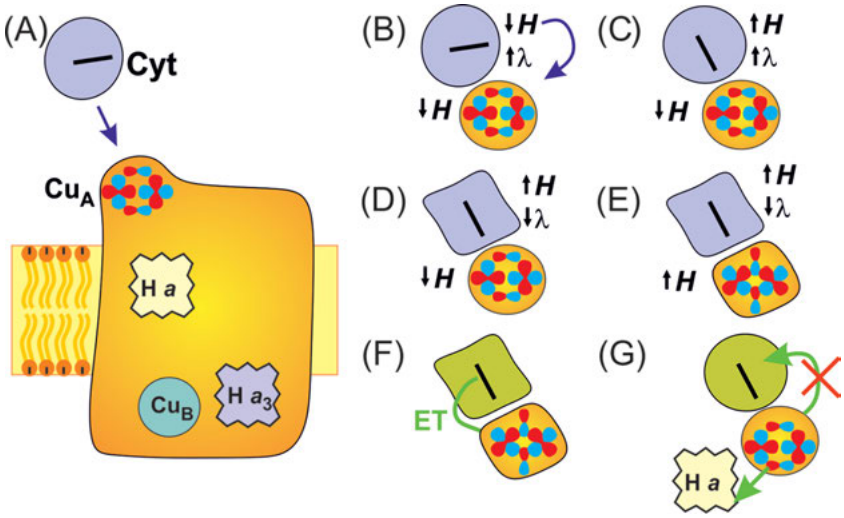


Figure 19. Schematic representation of the feedback inhibition mechanism proposed for the regulation of electron transfer reactions in the Cyt-c/COX complex. See text for a detailed description. Reprinted from [182] with permission from Elsevier; copyright 2014.

the transmembrane potential. Recently, Alvarez-Paggi et al. brought forward the hypothesis schematically depicted in Figure 19 for the negative feedback regulation of the Cyt-c → Cu_A ET reaction [168, 182]. The first step of the reaction (A → B) is the recognition and binding of Cyt-c to COX II. This process is electrostatically driven, but different experiments with the integral enzyme [169–174] and model systems [183–185] indicate that the initial complex is not optimized for ET that, therefore, requires some rearrangement of Cyt-c and interfacial water molecules before an effective electron pathway can be established (B → C). The interactions of Cyt-c with COX have been found to perturb the structure of Cyt-c [172] and these perturbation may result in disruption of a H-bond between the axial Met ligand of the heme iron and a Tyr residue, which results in a two-fold drop of the reorganization energy of Cyt-c in the complex with respect to the protein free in solution (C → D) [186]. These protein-protein interactions, in turn, might induce a slight distortion of either the Cu–Cu distance or dihedral angles in the Cu_A site, which were shown empirically [123] and computationally [136] to favor the population of the alternative π_u GS (D → E), thus further increasing the intermolecular ET rate (F). After the Cu_A site is reduced, the lower reorganization energy of the σ_u^* GS together with its enhanced electronic coupling with heme *a* favors this orbital as the redox active one for the subsequent ET step, thus conferring directionality to the overall process (G).

A temporary misbalance between the proton-pumping activity of the respiratory complexes and the gradient dissipation activity of ATP synthase may result in increased local electric fields. This, in turn, would result in slowed down reorientation of Cyt-c in the complex [184, 185], thus undermining its ability to achieve sufficiently high electronic coupling with the Cu_A electron acceptor.

Moreover, theoretical calculations predict that increasing electric fields of biologically meaningful magnitude may rise $\Delta E_{\sigma_u^*/\pi_u}$, thus reducing the accessibility of the π_u GS [136]. Therefore, the increase of the local electric field is expected to temporarily block or significantly slow the proton-coupled ET reactions of the Cyt-c/COX complex, until the proton electrochemical gradient is sufficiently dissipated by the activity of ATP-synthase. Phosphorylation of tyrosine residues of Cyt-c has been proposed as an additional inhibitory regulation mechanisms of the Cyt-c → Cu_A ET reaction [187].

7. NOT TOO FLOPPY, NOT TOO RIGID

7.1. Flexibility of the Cupredoxin Fold

Proteins in general are characterized by a remarkable structural and dynamical complexity that can be described in terms of multidimensional and hierarchical free energy landscapes. Exploration of the manifold of thermally accessible conformational substates throughout these landscapes implies a wide range of time scales, ranging from femtoseconds for bond vibrations, picoseconds to nanoseconds for side-chain rotations, and microseconds to seconds for collective motions of larger protein domains [188]. Most of these motions are slaved or coupled to fluctuations of either the bulk solvent or the hydration shell [189, 190]. The different hierarchies of molecular flexibility have been found to play a key functional role in the canonical and alternative functions of several proteins [168, 191–196].

In the specific case of Cu_A, but also in T1 Cu electron transfer proteins, it is the rigidity rather than the flexibility of the cupredoxin fold that is usually attributed with the special role of minimizing the ET reorganization energy of the metal site [26, 62, 86, 159, 160, 197–199].

Interestingly, the *Tt*Cu_A fragment presents an unusually high stability [200] characterized by folding free energies of –85 kJ mol^{-1} (T_m > 100 °C) and –65 kJ mol^{-1} (T_m = 83 °C) for the reduced and oxidized forms, respectively [201]. The T1 Cu protein azurin, in contrast, has significantly lower folding free energies (–52.2 and –40.0 kJ mol^{-1} for oxidized and reduced azurin, respectively) [202, 203], which is accompanied by an almost doubled reorganization energy with respect to *Tt*Cu_A [136, 159].

The low reorganization energies of copper proteins have been interpreted in terms of the so-called entatic/rack-induced state hypothesis, which assumes that the cupredoxin fold of the apoprotein contains a preformed rigid coordination site for the subsequent binding of copper ions [204–208]. This concept, however, does not consider the latter finding of chaperone-mediated uptake of copper in *Tt*Cu_A [96] and other copper proteins [69].

NMR studies revealed high degrees of structural heterogeneity in the metal-binding regions of the non-metallated *Tt*Cu_A and azurin-binding cupredoxin domains, particularly in the microsecond to second time domain. These dynamics are largely quenched upon copper binding to both apo-proteins, thus showing that a preformed metal-binding site is not required and, instead, metal binding

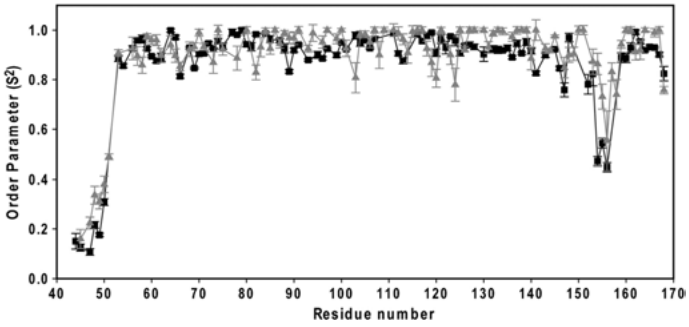


Figure 20. Backbone order parameters obtained by NMR wild type *Tt*Cu_A in its apo (non-metallated) form (black squares) and reduced holo (metalated) form (gray triangles). Adapted from [70] with permission of the National Academy of Sciences, USA, copyright 2012.

is a major contributor to the rigidity of electron transfer copper centers [70]. Thus, it was concluded that the high rigidity of the site is an imperative for the holoprotein to perform ET; while the flexibility of the apocupredoxin is necessary for the metal ions uptake from the metallochaperones [71, 209].

In general terms, NMR experiments confirm a high rigidity of the *Tt*Cu_A protein in the pico-to-nanosecond timescale, which is very similar in the holo- and apoproteins. Interestingly, the ligand loop 149–160 displays a significantly lower order parameter than other residues, thus revealing some flexibility for this crucial segment (Figure 20). In the microsecond-to-millisecond timescale two regions close to the metal binding site exhibit high mobility in the apoprotein, and this motion is significantly but not completely quenched in the holoprotein. For larger time scales up to seconds significant dynamic features were only observed in the absence of copper [70].

Overall, the NMR results confirm the view of a rigid cupredoxin fold that contributes to minimize the ET reorganization energies. At the same time, these studies reveal residual flexibility of specific segments within short-to-intermediate time scales that is expected to be coupled to solvent fluctuations and, thereby, to influence ET dynamics reactions through the viscoelastic properties of the solvent-protein milieu, as discussed in detail in the following section.

7.2. Frictionally Controlled Electron Transfer

Most experimental investigations of protein ET reactions are conducted *in vitro* using diluted and close to ideal protein solutions as these conditions facilitate spectroscopic and electrochemical determinations under well-defined and reproducible physicochemical conditions. The results of these experiments are usually treated within the framework of Marcus semi-classical theory. Extrapolation of these results to *in vivo* conditions, however, is not straightforward. Cells are crowded environments that include large structures and dissolved macromolecu-

les in concentrations as high as 450 g L^{-1}, which may affect a number of fundamental physicochemical parameters. For example, local microviscosities in the mitochondrial intermembrane space present values typically around 35–63 cP in healthy cells, and up to 400 cP under apoptotic conditions. This crowding effect, which also includes the presence of locally high electric fields imposed by the membrane potential, may affect the protein-solvent dynamics and, therefore, the ET reactions of Cu_A and of respiratory and photosynthetic complexes in general.

Recently, Zitare et al. assessed the effect of macromolecular crowding on the ET dynamics of Cu_A and T1 Cu copper proteins by systematic addition of well-established thickening or crowding agents, such as sucrose or polyethylene glycol [161, 182]. The experiments were conducted with four different types of metalloproteins: (a) the wild-type T1 Cu protein azurin, (b) the wild type *Tt*Cu_A, (c) the chimeras h*Tt*3L, *AtTt*3L, *Tt*1L-Entry, and *Tt*1L-Ligands, in which either one or the three loops of *Tt*Cu_A have been replaced by homologous sequences from other organism with preservation of the native *Tt*Cu_A ligand set and (d) the chimeras Cu_A-Ami and Cu_A-Az, in which the T1 mononuclear centers of amicyanin and azurin, respectively, have been introduced into the *Tt*Cu_A scaffold by loop engineering. The different proteins were adsorbed on electrodes previously coated with biocompatible films of alkane thiols with variable length, and heterogeneous ET rate constants (k_{ET}) were determined electrochemically. In all cases, when measured at low viscosities of ca. 1 cP, k_{ET} exhibits an exponential distance-dependence for the thicker alkane thiol films (longer electron tunneling distances), which is typical of a Marcus-like long-range non-adiabatical ET mechanism. For thinner films, however, k_{ET} shows a much softer dependence with the tunneling distance, thus suggesting a change of ET regime associated to the increasing electronic coupling.

Addition of crowding agents results in significant drops of k_{ET} for all the studied copper proteins, which correlate with the viscosity of the medium (η) as a power law of the form $k_{ET} \propto \eta^{-\gamma}$. The parameter γ is zero at electron tunneling distances ≥ 23 Å, and increases in a sigmoidal fashion upon shortening the electron pathway, i.e., upon increasing the electronic coupling. This variation reveals that k_{ET} values are determined by the interplay between electron tunneling times, which are a function of the film thickness, and solvent relaxation times, which are a function of the solution viscosity. Such a kinetic behavior is indicative of a frictionally controlled ET mechanism that is best described by Equation (2) [210]:

$$k_{ET} = \frac{\epsilon_s}{3A\gamma V_m \epsilon_{op}} \left(\frac{RT\lambda}{\pi^3} \right)^{1/2} exp\left(-\frac{\Delta G^{\#}_{ET} + \gamma \Delta G^{\#}_{s}}{RT} \right) \quad (2)$$

where ε_s and ε_{op} are the statical and optical dielectric constants, respectively, A is an empirical constant, V_m is the solvent molar volume, $\Delta G^{\#}_{app} = \Delta G^{\#}_{ET} + \gamma \Delta G^{\#}_{s}$ is the apparent activation free energy of the reaction, and R, T, and λ have the usual meaning. This equation reduces to the usual Marcus expression when $\gamma = 0$, i.e., when there is no frictional control of the ET rate and, therefore, $\Delta G^{\#}_{app} = \Delta G^{\#}_{ET} = (\Delta G^0 + \lambda)^2 / 4\lambda$. Interestingly, the parameter γ was found to be metal site specific, with an average value of 0.54 for wild-type and chimeric

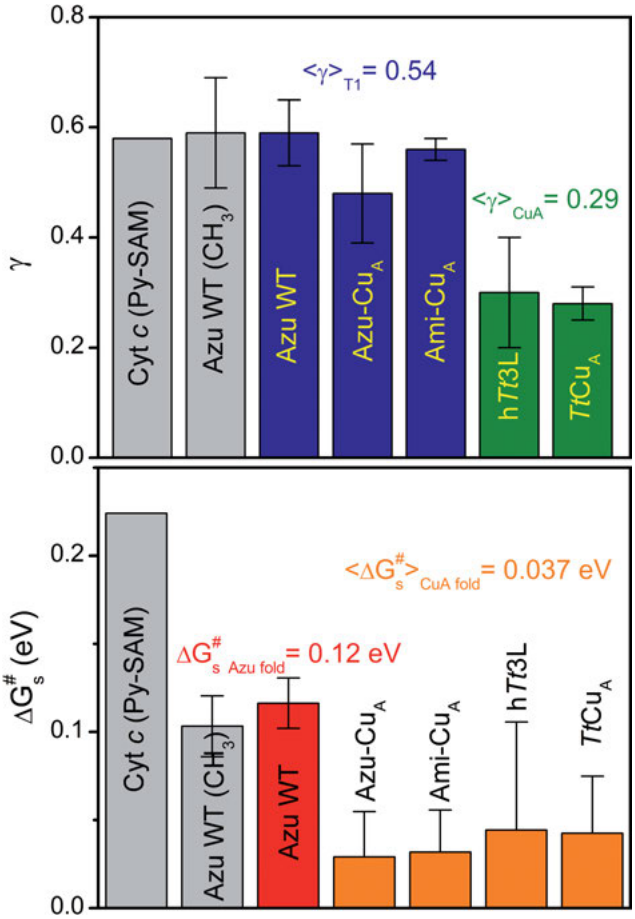


Figure 21. Empirical frictional parameter (γ) and activation free energy for the milieu frictional motion ($\Delta G_s^{\#}$) for different copper proteins, wild-type proteins azurin (Azu WT), and *Tt*Cu_A and the chimeric proteins Azu-Cu_A, Ami-Cu_A, and h*Tt*3L compared with those reported for the more flexible cytochrome *c* (Cyt c) [161, 182]. Reprinted with permission from Elsevier; copyright 2019.

T1 sites and of 0.29 for all binuclear Cu_A centers (Figure 21). These results are consistent with the rigidity of the Cu_2S_2 diamond of Cu_A, compared to the mononuclear T1 Cu sites, and indicate that the parameter γ is a direct measure of the dynamic coupling between the metal center and the protein/solvent milieu. The parameter $\Delta G_s^{\#}$, on the other hand, reports on the temperature-dependence of those protein-solvent motions that define the ET reaction coordinate and was found to be defined by the overall protein matrix rather than by the metal site. Thus $\Delta G_s^{\#}$ has a nearly identical value of 0.037 eV for all the wild-type and chimeric proteins based on the *Tt*-Cu_A fold, independently of the type of metal site, while for the azurin fold this value is 0.12 eV (Figure 21).

These studies show that the nearly two orders of magnitude higher viscosities of the cell interior compared to the usual *in vitro* conditions are sufficiently high to slow down nuclear motions relevant to the ET coordinate, thus breaking down the underlying assumptions of Marcus theory. This effect leads to a frictionally controlled ET mechanism that is characterized by significantly lower than expected ET rates for tunneling distances shorter than 24 Å. For a characteristic electron tunneling distance of 10 Å this may result in a 90 % drop of k_{ET} for the Cu_A site when the viscosity reaches the typical values of the mitochondrial intermembrane space, approx. 55 cP [161, 182].

8. FUTURE DIRECTIONS

The study of the electronic and geometric structure of the Cu_A center has been an intense matter of debate over several decades in the community of biological inorganic chemists. After the unveiling of its structure, a decade of new intense spectroscopic studies followed, and the chapter of research on Cu_A seemed to be closed. However, there are new exciting findings in the field that bring new questions to be addressed.

One long-standing question that still remains to be answered is why nature has selected Cu_A centers instead of the more ubiquitous Type 1 Cu sites in COX and in N_2OR. In the same way, it is not clear why there are no soluble, small Cu_A-containing proteins in the cupredoxin scaffolds equivalent to the blue T1 Cu proteins.

There are two findings that challenge some of the currently accepted paradigms. The existence of two alternative ground states in Cu_A also suggests possible pathways for electron transfer that deserve further exploration and pose new questions about the directionality of biological electron transfer. The on/off conformation of a His ligand in N_2O reductase also challenges the requirement of a rigid site, and suggest the possibility of a gated electron transfer event, that requires elucidation within the mechanism of this enzyme.

Finally, the recent report of a symmetric, novel Cu_A site in PmoD suggests that there might be different alternatives in nature to build this efficient electron transfer center beyond the known design of ligand loops in a cupredoxin scaffold. In other words, a new chapter for the study of Cu_A has been opened.

ACKNOWLEDGMENTS

We thank all our present and previous group members who, along the years, have contributed with their work, enthusiasm and creativity to study the CuA site. We particularly thank Luciano Abriata, Damian Alvarez-Paggi, María Eugenia Zaballa, Ulises Zitare, Andrés Espinoza-Cara, María Eugenia Llases, Jonathan Szuster, and Alcho Leguto. We thank ANPCyT, UBA, UNR, and CONICET for support.

ABBREVIATIONS AND DEFINITIONS

Ami	amicyanin
Az	azurin
COX	cytochrome *c* oxidase
Cu_A-Az	loop-engineered azurin able to form a Cu_A site
Cyt-c	cytochrome *c*
EPR	electron paramagnetic resonance
ET	electron transfer
EXAFS	extended X-ray absorption fine structure
GS	ground state
HOMO	highest occupied molecular orbital
IMS	mitochondrial intermembrane space
NOR	nitric oxide reductase
N_2OR	nitrous oxide reductase = NosZ
NMR	nuclear magnetic resonance
RAMO	redox-active molecular orbital
RR	resonance Raman
Sco	proteins involved in the synthesis of cytochrome *c* oxidase
Tt Cu_A	Cu_A-containing soluble subunit II of *T. thermophilus* ba_3 oxidase
*Tt*1L	chimeric protein built on *Tt* Cu_A with one loop replaced by one of a eukaryotic homologue
*Tt*3L	chimeric protein built on *Tt* Cu_A with the three loops surrounding the metal site replaced by those of a eukaryotic homologue
XAS	X-ray absorption spectroscopy

REFERENCES

1. H. Beinert, *Eur. J. Biochem.* **1997**, *245*, 521–32.
2. P. M. H. Kroneck, *J. Biol. Inorg. Chem.* **2018**, *23*, 27–39.
3. J. P. Hosler, S. Ferguson-Miller, D. A. Mills, *Annu. Rev. Biochem.* **2006**, *75*, 165–187.
4. G. T. Babcock, M. Wikström, *Nature* **1992**, *356*, 301–309.
5. M. Wikström, K. Krab, V. Sharma, *Chem. Rev.* **2018**, *118*, 2469–2490.
6. C. Dennison, G. W. Canters, *Recl. Trav. Chim. Pays-Bas* **1996**, *115*, 345–351.
7. A. C. Rosenzweig, *Nat. Struct. Biol.* **2000**, *7*, 169.
8. R. A. Scott, W. G. Zumft, C. L. Coyle, D. M. Dooley, *Proc. Natl. Acad. Sci. USA* **1989**, *86*, 4082–4086.
9. C. Carreira, S. R. Pauleta, I. Moura, *J. Inorg. Biochem.* **2017**, *177*, 423–434.
10. A. Wust, L. Schneider, A. Pomowski, W. G. Zumft, P. M. Kroneck, O. Einsle, *Biol. Chem.* **2012**, *393*, 1067–1077.
11. S. Al-Attar, S. de Vries, *FEBS Lett.* **2015**, *589*, 2050–2057.
12. Suharti, M. J. Strampraad, I. Schröder, S. de Vries, *Biochemistry* **2001**, *40*, 2632–2639.
13. T. Soulimane, G. Buse, G. P. Bourenkov, H. D. Bartunik, R. Huber, M. E. Than, *EMBO J* **2000**, *19*, 1766–1776.
14. T. Tsukihara, H. Aoyama, E. Yamashita, T. Tomizaki, H. Yamaguchi, K. Shinzawa-Itoh, R. Nakashima, R. Yaono, S. Yoshikawa, *Science* **1996**, *272*, 1136–1144.

15. B. E. Ramirez, B. G. Malmström, J. R. Winkler, H. B. Gray, *Proc. Natl. Acad. Sci. USA* **1995**, *92*, 11949–11951.
16. M. F. Lucas, D. L. Rousseau, V. Guallar, *Biochim. Biophys.-Bioenerg.* **2011**, *1807*, 1305–1313.
17. E. Pilet, A. Jasaitis, U. Liebl, M. H. Vos, *Proc. Natl. Acad. Sci. USA* **2004**, *101*, 16198–16203.
18. V. Tipmanee, J. Blumberger, *J. Phys. Chem. B* **2012**, *116*, 1876–1883.
19. T. Tiefenbrunn, W. Liu, Y. Chen, V. Katritch, C. D. Stout, J. A. Fee, V. Cherezov, *PLOS ONE* **2011**, *6*, 1–12.
20. A. Pomowski, W. G. Zumft, P. M. Kroneck, O. Einsle, *Nature* **2011**, *477*, 234–237.
21. T. Tsukihara, K. Shimokata, Y. Katayama, H. Shimada, K. Muramoto, H. Aoyama, M. Mochizuki, K. Shinzawa-Itoh, E. Yamashita, M. Yao, Y. Ishimura, S. Yoshikawa, *Proc. Natl. Acad. Sci. USA* **2003**, *100*, 15304–15309.
22. M. Prudêncio, A. S. Pereira, P. Tavares, S. Besson, I. Cabrito, K. Brown, B. Samyn, B. Devreese, J. Van Beeumen, F. Rusnak, G. Fauque, J. J. G. Moura, M. Tegoni, C. Cambillau, I. Moura, *Biochemistry* **2000**, *39*, 3899–3907.
23. P. Stach, O. Einsle, W. Schumacher, E. Kurun, P. M. Kroneck, *J. Inorg. Biochem.* **2000**, *79*, 381–385.
24. K. Brown, M. Tegoni, M. Prudêncio, A. S. Pereira, S. Besson, J. J. Moura, I. Moura, C. Cambillau, *Nat. Struct. Biol.* **2000**, *7*, 191–195.
25. C. Immoos, M. G. Hill, D. Sanders, J. A. Fee, C. E. Slutter, J. H. Richards, H. B. Gray, *J. Biol. Inorg. Chem.* **1996**, *1*, 529–531.
26. E. I. Solomon, X. Xie, A. Dey, *Chem. Soc. Rev.* **2008**, *37*, 623–38.
27. J. Liu, S. Chakraborty, P. Hosseinzadeh, Y. Yu, S. Tian, I. Petrik, A. Bhagi, Y. Lu, *Chem. Rev.* **2014**, *114*, 4366–4469.
28. O. S. Fisher, G. E. Kenney, M. O. Ross, S. Y. Ro, B. E. Lemma, S. Batelu, P. M. Thomas, V. C. Sosnowski, C. J. DeHart, N. L. Kelleher, T. L. Stemmler, B. M. Hoffman, A. C. Rosenzweig, *Nat. Commun.* **2018**, *9*, 4276.
29. M. O. Ross, O. S. Fisher, M. N. Morgada, M. D. Krzyaniak, M. R. Wasielewski, A. J. Vila, B. M. Hoffman, A. C. Rosenzweig, *J. Am. Chem. Soc.* **2019**, *141*, 4678–4686.
30. C. A. Elvehjem, *Biochem. J.* **1930**, *24*, 415–426.
31. D. Keilin, E. F. Hartree, *Nature* **1955**, *176*, 200.
32. P. Lappalainen, R. Aasa, B. G. Malmström, M. Saraste, *J. Biol. Chem.* **1993**, *268*, 26416–26421.
33. C. E. Slutter, D. Sanders, P. Wittung, B. G. Malmström, R. Aasa, J. H. Richards, H. B. Gray, J. A. Fee, *Biochemistry* **1996**, *35*, 3387–3395.
34. C. von Wachenfeldt, S. de Vries, J. van der Oost, *FEBS Lett.* **1994**, *340*, 109–113.
35. M. Paumann, B. Lubura, G. Regelsberger, M. Feichtinger, G. Köllensberger, C. Jakopitsch, P. G. Furtmüller, G. A. Peschek, C. Obinger, *J. Biol. Chem.* **2004**, *279*, 10293–10303.
36. H. K. Abicht, M. A. Schärer, N. Quade, R. Ledermann, E. Mohorko, G. Capitani, H. Hennecke, R. Glockshuber, *J. Biol. Chem.* **2014**, *289*, 32431–32444.
37. J. van der Oost, P. Lappalainen, A. Musacchio, A. Warne, L. Lemieux, J. Rumbley, R. B. Gennis, R. Aasa, T. Pascher, B. G. Malmström, *EMBO J.* **1992**, *11*, 3209–3217.
38. M. Hay, J. H. Richards, Y. Lu, *Proc. Natl. Acad. Sci. USA* **1996**, *93*, 461–464.
39. C. Dennison, E. Vijgenboom, S. de Vries, J. van der Oost, G. W. Canters, *FEBS Lett.* **1995**, *365*, 92–94.
40. L. H. Jones, A. Liu, V. L. Davidson, *J. Biol. Chem.* **2003**, *278*, 47269–47274.
41. R. P. Houser, V. G. Young, W. B. Tolman, *J. Am. Chem. Soc.* **1996**, *118*, 2101–2102.
42. M. Gennari, J. Pécaut, S. DeBeer, F. Neese, M.-N. Collomb, C. Duboc, *Angew. Chem. Int. Ed.* **2011**, *50*, 5662–5666.

43. S. K. Ghosh, S. K. Saha, M. C. Ghosh, R. N. Bose, J. W. Reed, E. S. Gould, *Inorg. Chem.* **1992**, *31*, 3358–3362.
44. C. Belle, W. Rammal, J.-L. Pierre, *J. Inorg. Biochem.* **2005**, *99*, 1929–1936.
45. S. Iwata, C. Ostermeier, B. Ludwig, H. Michel, *Nature* **1995**, *376*, 660–669.
46. T. Tsukihara, H. Aoyama, E. Yamashita, T. Tomizaki, H. Yamaguchi, K. Shinzawa-Itoh, R. Nakashima, R. Yaono, S. Yoshikawa, *Science* **1995**, *269*, 1069–1074.
47. M. Wilmanns, P. Lappalainen, M. Kelly, E. Sauer-Eriksson, M. Saraste, *Proc. Natl. Acad. Sci. USA* **1995**, *92*, 11955–11959.
48. J. A. Farrar, F. Neese, P. Lappalainen, P. M. H. Kroneck, M. Saraste, W. G. Zumft, A. J. Thomson, *J. Am. Chem. Soc.* **1996**, *118*, 11501–11514.
49. M. N. Morgada, L. A. Abriata, U. Zitare, D. Alvarez-Paggi, D. H. Murgida, A. J. Vila, *Angew. Chem. Int. Ed.* **2014**, *53*, 6188–6192.
50. E. T. Adman, *Adv. Protein Chem.* **1991**, *42*, 145–197.
51. C. Dennison, *Nat. Prod. Rep.* **2008**, *25*, 15–24.
52. P. A. Williams, N. J. Blackburn, D. Sanders, H. Bellamy, E. A. Stura, J. A. Fee, D. E. McRee, *Nat. Struct. Biol.* **1999**, *6*, 509–516.
53. J. M. Guss, H. D. Bartunik, H. C. Freeman, *Acta Crystallogr. B* **1992**, *48*, 790–811.
54. M. Kelly, P. Lappalainen, G. Talbo, T. Haltia, J. van der Oost, M. Saraste, *J. Biol. Chem.* **1993**, *268*, 16781–16787.
55. H. B. Gray, B. G. Malmstrom, R. J. Williams, *J. Biol. Inorg. Chem.* **2000**, *5*, 551–9.
56. N. J. Blackburn, M. E. Barr, W. H. Woodruff, J. van der Ooost, S. de Vries, *Biochemistry* **1994**, *33*, 10401–10407.
57. N. Yano, K. Muramoto, A. Shimada, S. Takemura, J. Baba, H. Fujisawa, M. Mochizuki, K. Shinzawa-Itoh, E. Yamashita, T. Tsukihara, S. Yoshikawa, *J. Biol. Chem.* **2016**, 23882–23894.
58. X. Wang, S. M. Berry, Y. Xia, Y. Lu, *J. Am. Chem. Soc.* **1999**, *121*, 7449–7450.
59. X. Xie, S. I. Gorelsky, R. Sarangi, D. K. Garner, H. J. Hwang, K. O. Hodgson, B. Hedman, Y. Lu, E. I. Solomon, *J. Am. Chem. Soc.* **2008**, *130*, 5194–5205.
60. U. A. Zitare, J. Szuster, M. C. Santalla, M. E. Llases, M. N. Morgada, A. J. Vila, D. H. Murgida, *Inorg. Chem.* **2019**, *58*, 2149–2157.
61. L. Muresanu, P. Pristovsek, F. Löhr, O. Maneg, M. D. Mukrasch, H. Rüterjans, B. Ludwig, C. Lücke, *J. Biol. Chem.* **2006**, *281*, 14503–14513.
62. S. Yanagisawa, M. J. Banfield, C. Dennison, *Biochemistry* **2006**, *45*, 8812–8822.
63. L. A. Abriata, A. J. Vila, *J. Inorg. Biochem.* **2014**, *132*, 18–20.
64. T. Kajikawa, K. Kataoka, T. Sakurai, *Biochem. Biophys. Res. Commun.* **2012**, *422*, 152–156.
65. J. S. Kretchmer, N. Boekelheide, J. J. Warren, J. R. Winkler, H. B. Gray, T. F. Miller, *Proc. Natl. Acad. Sci. USA* **2018**, *115*, 6129–6134.
66. R. J. Williams, *Eur. J. Biochem.* **1995**, *234*, 363–381.
67. E. I. Solomon, R. G. Hadt, B. E. R. Snyder, *Isr. J. Chem.* **2016**, *56*, 649–659.
68. H. Nar, A. Messerschmidt, R. Huber, M. van de Kamp, G. W. Canters, *FEBS Lett.* **1992**, *306*, 119–124.
69. N. J. Robinson, D. R. Winge, *Annu. Rev. Biochem.* **2010**, *79*, 537–562.
70. M.-E. Zaballa, L. A. Abriata, A. Donaire, A. J. Vila, *Proc. Natl. Acad. Sci. USA* **2012**, *109*, 9254–9259.
71. M. N. Morgada, L. A. Abriata, C. Cefaro, K. Gajda, L. Banci, A. J. Vila, *Proc. Natl. Acad. Sci. USA* **2015**, *112*, 11771–11776.
72. P. M. Kroneck, W. E. Antholine, D. H. Kastrau, G. Buse, G. C. Steffens, W. G. Zumft, *FEBS Lett.* **1990**, *268*, 274–276.
73. W. E. Antholine, D. H. Kastrau, G. C. Steffens, G. Buse, W. G. Zumft, P. M. Kroneck, *Eur. J. Biochem.* **1992**, *209*, 875–881.

74. P. M. Kroneck, W. A. Antholine, J. Riester, W. G. Zumft, *FEBS Lett.* **1988**, *242*, 70–74.
75. J. M. Charnock, A. Dreusch, H. Körner, F. Neese, J. Nelson, A. Kannt, H. Michel, C. D. Garner, P. M. H. Kroneck, W. G. Zumft, *Eur. J. Biochem.* **2000**, *267*, 1368–1381.
76. M. T. Hay, M. C. Ang, D. R. Gamelin, E. I. Solomon, W. E. Antholine, M. Ralle, N. J. Blackburn, P. D. Massey, X. Wang, A. H. Kwon, Y. Lu, P. O. Box, R. V. September, *Inorg. Chem.* **1998**, *37*, 191–198.
77. H. Robinson, M. C. Ang, Y.-G. Gao, M. T. Hay, Y. Lu, A. H. J. Wang, *Biochemistry* **1999**, *38*, 5677–5683.
78. H. J. Hwang, M. Ang, Y. Lu, *J. Biol. Inorg. Chem.* **2004**, *9*, 489–494.
79. K. M. Clark, S. Tian, W. A. van der Donk, Y. Lu, *Chem. Commun.* **2016**, *53*, 224–227.
80. M.-L. Tsai, R. G. Hadt, N. M. Marshall, T. D. Wilson, Y. Lu, E. I. Solomon, *Proc. Natl. Acad. Sci. USA* **2013**, *110*, 14658–14663.
81. T. D. Wilson, Y. Yu, Y. Lu, *Coord. Chem. Rev.* **2013**, *257*, 260–276.
82. L. Li, A. Song, Y. Xie, Z. Huang, E. de Waal, K. Urszula, G. W. Canters, *Chin. Sci. Bull.* **2001**, *46*, 1608–1611.
83. L. Komorowski, S. Anemüller, G. Schäfer, *J. Bioenerg. Biomembr.* **2001**, *33*, 27–34.
84. M.-E. Llases, M.-N. Lisa, M. N. Morgada, E. Giannini, P. M. Alzari, A. J. Vila, *FEBS J.* **2019**.
85. S. I. Gorelsky, X. Xie, Y. Chen, J. A. Fee, E. I. Solomon, *J. Am. Chem. Soc.* **2006**, *128*, 16452–16453.
86. D. R. Gamelin, D. W. Randall, M. T. Hay, R. P. Houser, T. C. Mulder, G. W. Canters, S. de Vries, W. B. Tolman, Y. Lu, E. I. Solomon, *J. Am. Chem. Soc.* **1998**, *120*, 5246–5263.
87. M. Gennari, J. Pécaut, M.-N. Collomb, C. Duboc, *Dalton Trans.* **2012**, *41*, 3130–3133.
88. U. Zitare, D. Alvarez-Paggi, M. N. Morgada, L. A. Abriata, A. J. Vila, D. H. Murgida, *Angew. Chem. Int. Ed.* **2015**, *54*, 9555–9559.
89. S. J. Stohs, D. Bagchi, *Free Radic. Biol. Med.* **1995**, *18*, 321–336.
90. K. Jomova, M. Valko, *Toxicology* **2011**, *283*, 65–87.
91. B.-E. Kim, T. Nevitt, D. J. Thiele, *Nat. Chem. Biol.* **2008**, *4*, 176–185.
92. R. A. Festa, D. J. Thiele, *Curr. Biol.* **2011**, *21*, R877–R883.
93. K. A. Jett, S. C. Leary, *J. Biol. Chem.* **2018**, *293*, 4644–4652.
94. S. Schimo, I. Wittig, K. M. Pos, B. Ludwig, *PLoS ONE* **2017**, *12*, 1–19.
95. A. Timón-Gómez, E. Nývltová, L. A. Abriata, A. J. Vila, J. Hosler, A. Barrientos, *Semin. Cell Dev. Biol.* **2018**, *76*, 163–178.
96. L. A. Abriata, L. Banci, I. Bertini, S. Ciofi-Baffoni, P. Gkazonis, G. A. Spyroulias, A. J. Vila, S. Wang, *Nat. Chem. Biol.* **2008**, *4*, 599.
97. L. Banci, I. Bertini, G. Cavallaro, S. Ciofi-Baffoni, *FEBS J.* **2011**, *278*, 2244–2262.
98. S. C. Leary, B. A. Kaufman, G. Pellecchia, G. H. Guercin, A. Mattman, M. Jaksch, E. A. Shoubridge, *Hum. Mol. Genet.* **2004**, *13*, 1839–1848.
99. P. Wunsch, M. Herb, H. Wieland, U. M. Schiek, W. G. Zumft, *J. Bacteriol.* **2003**, *185*, 887–896.
100. S. R. Pauleta, I. Moura, *Encycl. Inorg. Bioinorg. Chem.* **2017**, 1–11.
101. Z. N. Baker, P. A. Cobine, S. C. Leary, *Metallomics* **2017**, *9*, 1501–1512.
102. L. Banci, I. Bertini, K. S. McGreevy, A. Rosato, *Nat. Prod. Rep.* **2010**, *27*, 695–710.
103. A. D. Smith, B. L. Logeman, D. J. Thiele, *Annu. Rev. Microbiol.* **2017**, *71*, 597–623.
104. L. Banci, I. Bertini, V. Calderone, S. Ciofi-Baffoni, S. Mangani, M. Martinelli, P. Palumaa, S. Wang, *Proc. Natl. Acad. Sci. USA* **2006**, *103*, 8595–8600.
105. D. M. Glerum, A. Shtanko, A. Tzagoloff, *J. Biol. Chem.* **1996**, *271*, 20531–20535.
106. N. R. Mattatall, J. Jazairi, B. C. Hill, *J. Biol. Chem.* **2000**, *275*, 28802–28809.

107. T. Nittis, G. N. George, D. R. Winge, *Biochemistry* **2001**, *276*, 42520–42526.
108. D. M. Glerum, A. Shtanko, A. Tzagoloff, G. Moira, S. Andrey, T. Alexander, D. Moira Glerum, A. Shtanko, A. Tzagoloff, *J. Biol. Chem.* **1996**, *24*, 14504–14509.
109. L. Banci, I. Bertini, S. Ciofi-Baffoni, T. Hadjiloi, M. Martinelli, P. Palumaa, *Proc. Natl. Acad. Sci. USA* **2008**, *105*, 6803–6808.
110. Y. C. Horng, S. C. Leary, P. A. Cobine, F. B. J. Young, G. N. George, E. A. Shoubridge, D. R. Winge, *J. Biol. Chem.* **2005**, *280*, 34113–34122.
111. S. C. Leary, F. Sasarman, T. Nishimura, E. A. Shoubridge, *Hum. Mol. Genet.* **2009**, *18*, 2230–2240.
112. C. Abajian, L. A. Yatsunyk, B. E. Ramirez, A. C. Rosenzweig, *J. Biol. Chem.* **2004**, *279*, 53584–53592.
113. D. A. Stroud, M. J. Maher, C. Lindau, F. N. Vögtle, A. E. Frazier, E. Surgenor, H. Mountford, A. P. Singh, M. Bonas, S. Oeljeklaus, B. Warscheid, C. Meisinger, D. R. Thorburn, M. T. Ryan, *Hum. Mol. Genet.* **2015**, *24*, 5404–5415.
114. A. Ghosh, A. T. Pratt, S. Soma, S. G. Theriault, A. T. Griffin, P. P. Trivedi, V. M. Gohil, *Hum. Mol. Genet.* **2016**, *25*, 660–671.
115. P. Wunsch, H. Korner, F. Neese, R. J. van Spanning, P. M. Kroneck, W. G. Zumft, *FEBS Lett.* **2005**, *579*, 4605–4609.
116. C. Feng, *Coord. Chem. Rev.* **2012**, *256*, 393–411.
117. S. P. Bennett, M. J. Soriano-Laguna, J. M. Bradley, D. A. Svistunenko, D. J. Richardson, A. J. Gates, N. E. Le Brun, *Chem. Sci.* **2019**, *10*, 4985–4993.
118. C. Greenwood, B. C. Hill, D. Barber, D. G. Eglinton, A. J. Thomson, *Biochem. J.* **1983**, *215*, 303–316.
119. P. M. Kroneck, W. A. Antholine, J. Riester, W. G. Zumft, *FEBS Lett.* **1989**, *248*, 212–213.
120. C. R. Andrew, J. Sanders-Loehr, *Acc. Chem. Res.* **1996**, *29*, 365–372.
121. C. R. Andrew, J. Han, S. de Vries, J. van der Oost, B. A. Averill, T. M. Loehr, J. Sanders-Loehr, *J. Am. Chem. Soc.* **1994**, *116*, 10805–10806.
122. S. E. Wallace-Williams, C. A. James, S. de Vries, M. Saraste, P. Lappalainen, J. van der Oost, M. Fabian, G. Palmer, W. H. Woodruff, *J. Am. Chem. Soc.* **1996**, *118*, 3986–3987.
123. A. J. Leguto, M. A. Smith, M. N. Morgada, U. A. Zitare, D. H. Murgida, K. M. Lancaster, A. J. Vila, *J. Am. Chem. Soc.* **2019**, *141*, 1373–1381.
124. M. G. Savelieff, Y. Lu, *J. Biol. Inorg. Chem.* **2010**, *15*, 461–483.
125. G. Henkel, A. Müller, S. Weissgräber, G. Buse, T. Soulimane, G. C. M. Steffens, H.-F. Nolting, *Angew. Chem. Int. Ed.* **1995**, *34*, 1488–1492.
126. M. Karpefors, C. E. Slutter, J. A. Fee, R. Aasa, B. Källebring, S. Larsson, T. Vänngård, *Biophys. J.* **1996**, *71*, 2823–2829.
127. F. Neese, W. G. Zumft, W. E. Antholine, P. M. H. Kroneck, *J. Am. Chem. Soc.* **1996**, *118*, 8692–8699.
128. K. R. Williams, D. R. Gamelin, L. B. LaCroix, R. P. Houser, W. B. Tolman, T. C. Mulder, S. de Vries, B. Hedman, K. O. Hodgson, E. I. Solomon, *J. Am. Chem. Soc.* **1997**, *119*, 613–614.
129. S. DeBeer George, M. Metz, R. K. Szilagyi, H. Wang, S. P. Cramer, Y. Lu, W. B. Tolman, B. Hedman, K. O. Hodgson, E. I. Solomon, *J. Am. Chem. Soc.* **2001**, *123*, 5757–5767.
130. H. J. Hwang, Y. Lu, *Proc. Natl. Acad. Sci. USA* **2004**, *101*, 12842–12847.
131. U. Kolczak, J. Salgado, G. Siegal, M. Saraste, G. W. Canters, *Biospectroscopy* **1999**, *5*, S19–32.
132. J. Salgado, G. Warmerdam, L. Bubacco, G. W. Canters, *Biochemistry* **1998**, *37*, 7378–7389.
133. C. Dennison, A. Berg, G. W. Canters, *Biochemistry* **1997**, *36*, 3262–3269.

134. I. Bertini, K. L. Bren, A. Clemente, J. A. Fee, H. B. Gray, C. Luchinat, B. G. Malmström, J. H. Richards, D. Sanders, C. E. Slutter, *J. Am. Chem. Soc.* **1996**, *118*, 11658–11659.
135. L. A. Abriata, G. N. Ledesma, R. Pierattelli, A. J. Vila, *J. Am. Chem. Soc.* **2009**, *131*, 1939–1946.
136. L. A. Abriata, D. Alvarez-Paggi, G. N. Ledesma, N. J. Blackburn, A. J. Vila, D. H. Murgida, *Proc. Natl. Acad. Sci. USA* **2012**, *109*, 17348–17353.
137. M. N. Morgada, M.-E. Llases, L. A. Abriata, E. Giannini, M.-N. Lisa, M.-A. Castro, D. H. Murgida, P. M. Alzari, A. J. Vila, unpublished data.
138. C. O. Fernández, J. A. Cricco, C. E. Slutter, J. H. Richards, H. B. Gray, A. J. Vila, *J. Am. Chem. Soc.* **2001**, *123*, 11678–11685.
139. M. H. Olsson, U. Ryde, *J. Am. Chem. Soc.* **2001**, *123*, 7866–7876.
140. L. Ryden, *Prog. Clin. Biol. Res.* **1988**, *274*, 349–366.
141. V. Zickermann, M. Verkhovsky, J. Morgan, M. Wikström, S. Anemüller, E. Bill, G. C. M. Steffens, B. Ludwig, *Eur. J. Biochem.* **1995**, *234*, 686–693.
142. H. J. Hwang, S. M. Berry, M. J. Nilges, Y. Lu, *J. Am. Chem. Soc.* **2005**, *127*, 7274–7275.
143. G. N. Ledesma, D. H. Murgida, H. K. Ly, H. Wackerbarth, J. Ulstrup, A. J. Costa-Filho, A. J. Vila, *J. Am. Chem. Soc.* **2007**, *129*, 11884–11885.
144. S. I. Chan, P. M. Li, *Biochemistry* **1990**, *29*, 1–12.
145. C. E. Slutter, R. Langen, D. Sanders, S. M. Lawrence, P. Wittung, A. J. Di Bilio, M. G. Hill, J. A. Fee, J. H. Richards, J. R. Winkler, B. G. Malmström, *Inorg. Chim. Acta* **1996**, *243*, 141–145.
146. K. Fujita, N. Nakamura, H. Ohno, B. S. Leigh, K. Niki, H. B. Gray, J. H. Richards, *J. Am. Chem. Soc.* **2004**, *126*, 13954–13961.
147. S. Gupta, A. Warne, M. Saraste, S. Mazumdar, *Biochemistry* **2001**, *40*, 6180–6189.
148. D. Alvarez-Paggi, U. A. Zitare, J. Szuster, M. N. Morgada, A. J. Leguto, A. J. Vila, D. H. Murgida, *J. Am. Chem. Soc.* **2017**, *139*, 9803–9806.
149. D. Alvarez-Paggi, L. A. Abriata, D. H. Murgida, A. J. Vila, *Chem. Commun.* **2013**, *49*, 5381–5383.
150. D. Lukoyanov, S. M. Berry, Y. Lu, W. E. Antholine, C. P. Scholes, *Biophys. J.* **2002**, *82*, 2758–2766.
151. Y. Zhen, B. Schmidt, U. G. Kang, W. Antholine, S. Ferguson-Miller, *Biochemistry* **2002**, *41*, 2288–2297.
152. K. Wang, L. Geren, Y. Zhen, L. Ma, S. Ferguson-Miller, B. Durham, F. Millett, *Biochemistry* **2002**, *41*, 2298–2304.
153. N. J. Blackburn, M. Ralle, E. Gomez, M. G. Hill, A. Pastuszyn, D. Sanders, J. A. Fee, *Biochemistry* **1999**, *38*, 7075–7084.
154. S. Y. New, N. M. Marshall, T. S. A. Hor, F. Xue, Y. Lu, *Chem. Commun.* **2012**, *48*, 4217–4219.
155. N. M. Marshall, D. K. Garner, T. D. Wilson, Y.-G. Gao, H. Robinson, M. J. Nilges, Y. Lu, *Nature* **2009**, *462*, 113–116.
156. A. J. Di Bilio, M. G. Hill, N. Bonander, B. G. Karlsson, R. M. Villahermosa, B. G. Malmström, J. R. Winkler, H. B. Gray, *J. Am. Chem. Soc.* **1997**, *119*, 9921–9922.
157. P. Brzezinski, *Biochemistry* **1996**, *35*, 5611–5615.
158. S. Larsson, B. Källebring, P. Wittung, B. G. Malmström, *Proc. Natl. Acad. Sci. USA* **1995**, *92*, 7167–7171.
159. O. Farver, Y. Lu, M. C. Ang, I. Pecht, *Proc. Natl. Acad. Sci. USA* **1999**, *96*, 899–902.
160. O. Farver, H. J. Hwang, Y. Lu, I. Pecht, *J. Phys. Chem. B* **2007**, *111*, 6690–6694.
161. N. J. Blackburn, S. de Vries, M. E. Barr, R. P. Houser, W. B. Tolman, D. Sanders, J. A. Fee, *J. Am. Chem. Soc.* **1997**, *119*, 6135–6143.

162. U. A. Zitare, J. Szuster, M. F. Scocozza, A. Espinoza-Cara, A. J. Leguto, M. N. Morgada, A. J. Vila, D. H. Murgida, *Electrochim. Acta* **2019**, *294*, 117–125.
163. J. J. Regan, B. E. Ramirez, J. R. Winkler, H. B. Gray, B. G. Malmström, *J. Bioenerg. Biomembr.* **1998**, *30*, 35–39.
164. B. G. Malmström, *J. Biol. Inorg. Chem.* **1998**, *3*, 339–343.
165. D. M. Medvedev, I. Daizadeh, A. A. Stuchebrukhov, *J. Am. Chem. Soc.* **2000**, *122*, 6571–6582.
166. M.-L. Tan, I. Balabin, J. N. Onuchic, *Biophys. J.* **2004**, *86*, 1813–1819.
167. S. Yoshikawa, A. Shimada, *Chem. Rev.* **2015**, *115*, 1936–1989.
168. D. Alvarez-Paggi, L. Hannibal, M. A. Castro, S. Oviedo-Rouco, V. Demicheli, V. Tórtora, F. Tomasina, R. Radi, D. H. Murgida, *Chem. Rev.* **2017**, *117*, 13382–13460.
169. H. Witt, F. Malatesta, F. Nicoletti, M. Brunori, B. Ludwig, *Eur. J. Biochem.* **1998**, *251*, 367–373.
170. H. Witt, F. Malatesta, F. Nicoletti, M. Brunori, B. Ludwig, *J. Biol. Chem.* **1998**, *273*, 5132–5136.
171. I. Bertini, G. Cavallaro, A. Rosato, *J. Biol. Inorg. Chem.* ***2005***, *10*, 613–624.
172. K. Sakamoto, M. Kamiya, M. Imai, K. Shinzawa-Itoh, T. Uchida, K. Kawano, S. Yoshikawa, K. Ishimori, *Proc. Natl. Acad. Sci. USA* **2011**, *108*, 12271–12276.
173. W. Sato, S. Hitaoka, K. Inoue, M. Imai, T. Saio, T. Uchida, K. Shinzawa-Itoh, S. Yoshikawa, K. Yoshizawa, K. Ishimori, *J. Biol. Chem.* **2016**, *291*, 15320–15331.
174. S. Shimada, K. Shinzawa-Itoh, J. Baba, S. Aoe, A. Shimada, E. Yamashita, J. Kang, M. Tateno, S. Yoshikawa, T. Tsukihara, *EMBO J.* **2017**, *36*, 291–300.
175. J. A. Lyons, D. Aragao, O. Slattery, A. V. Pisliakov, T. Soulimane, M. Caffrey, *Nature* **2012**, *487*, 514–8.
176. O. Farver, Y. Chen, J. A. Fee, I. Pecht, *FEBS Lett.* **2006**, *580*, 3417–3421.
177. O. Farver, Ó. Einarsdóttir, I. Pecht, *Eur. J. Biochem.* **2000**, *267*, 950–954.
178. L. M. Geren, J. R. Beasley, B. R. Fine, A. J. Saunders, S. Hibdon, G. J. Pielak, B. Durham, F. Millett, *J. Biol. Chem.* **1995**, *270*, 2466–2472.
179. S. H. Speck, D. Dye, E. Margoliash, *Proc. Natl. Acad. Sci. USA* **1984**, *81*, 347–351.
180. E. A. E. Garber, E. Margoliash, *Biochim. Biophys. Acta-Bioenerg.* **1990**, *1015*, 279–287.
181. H. B. Gray, J. R. Winkler, *Q. Rev. Biophys.* **2003**, *36*, 341–372.
182. D. Alvarez-Paggi, U. Zitare, D. H. Murgida, *Biochim. Biophys. Acta* **2014**, *1837*, 1196–1207.
183. D. Alvarez-Paggi, W. Meister, U. Kuhlmann, I. Weidinger, K. Tenger, L. Zimányi, G. Rákhely, P. Hildebrandt, D. H. Murgida, *J. Phys. Chem. B* **2013**, *117*, 6061–6068.
184. D. Alvarez-Paggi, D. F. Martín, P. M. DeBiase, P. Hildebrandt, M. A. Martí, D. H. Murgida, *J. Am. Chem. Soc.* **2010**, *132*, 5769–5778.
185. A. Kranich, H. K. Ly, P. Hildebrandt, D. H. Murgida, *J. Am. Chem. Soc.* **2008**, *130*, 9844–9848.
186. D. Alvarez-Paggi, M. A. Castro, V. Tórtora, L. Castro, R. Radi, D. H. Murgida, *J. Am. Chem. Soc.* **2013**, *135*, 4389–4397.
187. M. Hüttemann, P. Pecina, M. Rainbolt, T. H. Sanderson, V. E. Kagan, L. Samavati, J. W. Doan, I. Lee, *Mitochondrion* **2011**, *11*, 369–381.
188. K. Henzler-Wildman, D. Kern, *Nature* **2007**, *450*, 964–972.
189. H. Frauenfelder, S. Sligar, P. Wolynes, *Science* **1991**, *254*, 1598–1603.
190. H. Frauenfelder, G. Chen, J. Berendzen, P. W. Fenimore, H. Jansson, B. H. McMahon, I. R. Stroe, J. Swenson, R. D. Young, *Proc. Natl. Acad. Sci. USA* **2009**, *106*, 5129–5134.
191. V. C. Nashine, S. Hammes-Schiffer, S. J. Benkovic, *Curr. Opin. Chem. Biol.* **2010**, *14*, 644–651.
192. G. Wei, W. Xi, R. Nussinov, B. Ma, *Chem. Rev.* **2016**, *116*, 6516–6551.

193. X. J. Jordanides, M. J. Lang, X. Song, G. R. Fleming, *J. Phys. Chem. B* **1999**, *103*, 7995–8005.
194. J. T. King, K. J. Kubarych, *J. Am. Chem. Soc.* **2012**, *134*, 18705–18712.
195. D. Laage, T. Elsaesser, J. T. Hynes, *Chem. Rev.* **2017**, *117*, 10694–10725.
196. M.-C. Bellissent-Funel, A. Hassanali, M. Havenith, R. Henchman, P. Pohl, F. Sterpone, D. van der Spoel, Y. Xu, A. E. Garcia, *Chem. Rev.* **2016**, *116*, 7673–7697.
197. D. W. Randall, D. R. Gamelin, L. B. LaCroix, E. I. Solomon, *J. Biol. Inorg. Chem.* ***2000***, *5*, 16–29.
198. J. R. Winkler, P. Wittung-Stafshede, J. Leckner, B. G. Malmström, H. B. Gray, *Proc. Natl. Acad. Sci. USA* **1997**, *94*, 4246–4249.
199. K. M. Lancaster, O. Farver, S. Wherland, E. J. Crane, 3rd, J. H. Richards, I. Pecht, H. B. Gray, *J. Am. Chem. Soc.* **2011**, *133*, 4865–4873.
200. A. Sujak, N. J. M. Sanghamitra, O. Maneg, B. Ludwig, S. Mazumdar, *Biophys. J.* **2007**, *93*, 2845–2851.
201. P. Wittung-Stafshede, B. G. Malmström, D. Sanders, J. A. Fee, J. R. Winkler, H. B. Gray, *Biochemistry* **1998**, *37*, 3172–3177.
202. J. Leckner, P. Wittung, N. Bonander, B. G. Karlsson, B. G. Malmström, *J. Biol. Inorg. Chem.* **1997**, *2*, 368–371.
203. J. Chaboy, S. Díaz-Moreno, I. Díaz-Moreno, M. A. De la Rosa, A. Díaz-Quintana, *Chem. Biol.* **2011**, *18*, 25–31.
204. L. B. LaCroix, S. E. Shadle, Y. Wang, B. A. Averill, B. Hedman, K. O. Hodgson, E. I. Solomon, *J. Am. Chem. Soc.* **1996**, *118*, 7755–7768.
205. B. G. Malmström, *Eur. J. Biochem.* **1994**, *223*, 711–718.
206. S. Larsson, *J. Biol. Inorg. Chem.* **2000**, *5*, 560–564.
207. A. Messerschmidt, L. Prade, S. J. Kroes, J. Sanders-Loehr, R. Huber, G. W. Canters, *Proc. Natl. Acad. Sci. USA* **1998**, *95*, 3443–3448.
208. C. Buning, G. W. Canters, P. Comba, C. Dennison, L. Jeuken, M. Melter, J. Sanders-Loehr, *J. Am. Chem. Soc.* **2000**, *122*, 204–211.
209. S. A. Pérez-Henarejos, L. A. Alcaraz, A. Donaire, *Arch. Biochem. Biophys.* **2015**, *584*, 134–148.
210. H. Yue, D. Khoshtariya, D. H. Waldeck, J. Grochol, P. Hildebrandt, D. H. Murgida, *J. Phys. Chem. B* **2006**, *110*, 19906–19913.

5

The Tetranuclear Copper-Sulfide Center of Nitrous Oxide Reductase

Sofia R. Pauleta,[1*] ***Marta S. P. Carepo,***[2] ***and Isabel Moura***[2]

[1]Microbial Stress Lab, UCIBIO, REQUIMTE, Departamento de Química, Faculdade de Ciências e Tecnologia, Universidade Nova de Lisboa, PT-2829–516 Caparica, Portugal <srp@fct.unl.pt>

[2]Biological Chemistry Lab, LAQV, REQUIMTE, Departamento de Química, Faculdade de Ciências e Tecnologia, Universidade Nova de Lisboa, PT-2829–516 Caparica, Portugal <marta.carepo@fct.unl.pt> <isabelmoura@fct.unl.pt>

[*]SRP planned and wrote the manuscript, with contributions from MSPC and IM.

Metal Ions in Life Sciences, Volume 20 Guest Editors: Martha E. Sosa Torres and Peter M. H. Kroneck
Series Editors: Astrid Sigel, Eva Freisinger, and Roland K. O. Sigel

https://doi.org/10.1515/9783110589757-005

Abstract: Nitrous oxide reductase catalyzes the reduction of nitrous oxide (N_2O) to dinitrogen (N_2) and water at a catalytic tetranuclear copper sulfide center, named CuZ, overcoming the high activation energy of this reaction. In this center each Cu atom is coordinated by two imidazole rings of histidine side-chains, with the exception of one named Cu_{IV}. This enzyme has been isolated with CuZ in two forms CuZ(4Cu1S) and CuZ(4Cu2S), which differ in the Cu_I-Cu_{IV} bridging ligand, leading to considerable differences in their spectroscopic and catalytic properties. The Cu atoms in CuZ(4Cu1S) can be reduced to the [$4Cu^{1+}$] oxidation state, and its catalytic properties are compatible with the nitrous oxide reduction rates of whole cells, while in CuZ(4Cu2S) they can only be reduced to the [$1Cu^{2+}$-$3Cu^{1+}$] oxidation state, which has a very low turnover number. The catalytic cycle of this enzyme has been explored and one of the intermediates, CuZ^0, has recently been identified and shown to be in the [$1Cu^{2+}$-$3Cu^{1+}$] oxidation state. Contrary to CuZ(4Cu2S), CuZ^0 is rapidly reduced intramolecularly by the electron transferring center of the enzyme, CuA, to [$4Cu^{1+}$] by a physiologically relevant redox partner.

The three-dimensional structure of nitrous oxide reductase with the CuZ center either as CuZ(4Cu1S) or as CuZ(4Cu2S) shows that it is a unique functional dimer, with the CuZ of one subunit receiving electrons from CuA of the other subunit. The complex nature of this center has posed some questions relative to its assembly, which are only partially answered, as well as to which is the active form of CuZ *in vivo*.

The structural, spectroscopic, and catalytic features of the two forms of CuZ will be addressed here, as well as its assembly. The understanding of its catalytic features, activation, and assembly is essential to develop strategies to decrease the release of nitrous oxide to the atmosphere and to reduce its concentration in the stratosphere, as well as to serve as inspiration to synthetic inorganic chemists to develop new models of this peculiar and challenging copper sulfide center.

Keywords: activation · biogenesis · catalytic mechanism · CuZ center · denitrification · greenhouse gas · nitrous oxide · nitrous oxide reductase

1. INTRODUCTION

Nitrous oxide (N_2O) is a powerful greenhouse gas with a major impact on global warming due to its long lifetime in the atmosphere, global warming potential (300 times higher than CO_2), and implications on the ozone layer depletion occurring in the stratosphere [1–6]. Nitrous oxide reductase (N_2OR) is the enzyme that catalyzes the reduction of N_2O to N_2 and water (Eq. (1)), transforming a powerful oxidant in two inert molecules, and overcoming a high energetic barrier that is associated with this chemical reaction [7, 8].

$$N_2O + 2\ e^- + 2\ H^+ \rightarrow N_2 + H_2O,\ E^{0\prime} = 1.35\ V,\ pH\ 7.0 \quad (1)$$

The active center of N_2OR is a tetranuclear copper center, CuZ, that is quite unique and complex not only in its geometry but also in its spectroscopic and catalytic properties [9, 10]. The complexity associated with CuZ is also due to the presence of a μ_4-sulfide ligand bridging the four copper atoms and to the

nature of another ligand bridging Cu_I and Cu_{IV}, either a solvent derived ligand on an "open" edge, or a second sulfide (S^{2-}). The binding of this sulfur ligand and its implications towards activity is still a matter of controversy and contributes to the uncertainty that is still associated with the catalytic mechanism proposed so far. Therefore, although the structure, spectroscopic, and catalytic properties of CuZ have been extensively studied there are several aspects of its activation, structure *in vivo*, and assembly that remain to be answered. Some of these aspects will be discussed here.

The uniqueness of the CuZ arrangement in Nature is even more puzzling since it is known that there is no need for tetranuclearity to obtain N_2O reduction, as binuclear copper model compounds have been shown to present activity towards N_2O [11]. Furthermore, contrary to other known enzymes able to catalyze reactions involving gases, such as hydrogenase, nitrogenase, CO dehydrogenase or even nitrite reductases, that show versatility towards the transition metals present in their active center to achieve catalysis [12–16], in the case of N_2OR this is not observed, and the CuZ center is the only known metal center among these enzymes found so far in Nature [17]. The specificity presented by the N_2OR active center must therefore be related with the properties and reactivity of N_2O that will be briefly discussed in Section 2.1.

2. NITROUS OXIDE

2.1. Properties and Reactions

N_2O reduction to N_2 is a highly exergonic reaction, with a $\Delta G^{0\prime}$ of –339.5 kJ/mol [18], however, it has a high energetic activation barrier of 250 kJ/mol [19] due to the electronic delocalization that stabilizes the molecule and because it corresponds to a spin-forbidden process [7]. N_2O has a dipole moment of 0.166 Debye, with double bond distances shorter than average, which are 1.128 Å and 1.184 Å, for N–N and N–O, respectively [7, 8]. The high reduction potential observed catalogues N_2O as a powerful oxidant and in Nature only N_2OR is able to overcome this activation barrier reducing this gas to the inert N_2 and water.

The atmospheric concentration of N_2O increased 22 % compared with the values estimated for the pre-industrial era, and in 2017 was approximately 330 ppb, corresponding to an increase of 0.25 % each year [5, 20, 21]. The photolytic action of sunlight in the stratosphere promotes N_2O decomposition to N_2 and $O(^1D)$. A minor fraction of N_2O reacts with $O(^1D)$ forming NO [5, 22, 23]. This reaction is an important source of reactive nitrogen species that will react with ozone accounting for the greenhouse effect of N_2O.

Due to the driving force to find efficient ways to decompose N_2O, hetero- and homogeneous catalysis using metal complexes have been described (some examples can be found in [24–26]), mainly involving non-transition metals, considering that N_2O is a weak ligand. Model compounds that mimic the CuZ center are also being tested towards N_2O reduction and so far, there are already several

multicopper clusters with the ability to react with this molecule. One example is the dissymmetric mixed-valent Cu(II)Cu(I) complex $[2{\cdot}(H_2O)(OTf)]^{1+}$ (OTf = trifluoromethanesulfonate anion, $CF_3SO_3^-$) published by Torelli and co-workers, containing OTf and H_2O as exchangeable ligands. This compound reduces N_2O to N_2, leading to a doubly bridged $[Cu_2(\mu\text{-SR})(\mu\text{-OH})]$ complex as the final product [11]. Another example is the $Cu_2^{1+}(\mu\text{-S})$ complex synthesized by Hillhouse and co-workers, $\{(IPr^*)Cu\}_2(\mu\text{-S})$ (IPr* = 1,3-bis(2,6-(diphenylmethyl)-4-methylphenyl)imidazole-2-ylidene) [27]. This complex reacts with N_2O giving rise to a mixture of six compounds, with $[(IPr^*)Cu]_2(\mu\text{-SO}_4)$, as the major product of the reaction, and it is suggested that the role of the tetranuclearity observed for CuZ is a way to protect the μ_4-sulfur ligand from oxidation or expulsion [28]. Tolman's group synthesized a localized mixed valent $Cu(II)Cu(I)_2$ cluster bridged by disulfide, $[L_3Cu_3(\mu_3\text{-}S_2)]X_2$ (L = 1,4,7-trimethyl-triazacyclononane, $X = O_3SCF_3^-$ or SbF_6^-). The cluster reacts with N_2O to yield N_2, and computational analysis implicates a transition state structure that features μ-1,1-bridging of N_2O via its O-atom to a $[L_2Cu_2(\mu\text{-}S_2)]^+$ fragment [29]. The first model compound for the μ-sulfido-tetracopper CuZ center, with the ability to reduce N_2O, is one in which two molecules of a reduced amidinate-supported $[Cu_4(\mu_4\text{-S})]$ cluster are able to reduce N_2O, one cluster acting as an activator and the other as a reductant [30].

2.2 Sources of Nitrous Oxide

Several studies have shown that N_2O emissions can come from natural sources like oceans, forests, and savannas, as well as from anthropogenic sources. The anthropogenic sources of N_2O are identified as the agricultural soils, mainly the production of forages and nitrogen-fixing crops, livestock manure, as well as manure used as fertilizers, fossil fuel combustion, adipic and nitric acid industrial production, waste incineration, biomass burning, and waste water treatment plants [31–35].

From 1970 to 2012 according to the Emission Database for Global Atmospheric Research (EDGAR), N_2O emission expressed in CO_2 equivalents increased in Asia Pacific, Sub-Sahara Africa, Latin America, and the Caribbean region, and decreased in the European Union and North America. The decrease in N_2O emissions observed in both Europe and North America is mainly a result of a reduction in N_2O emissions from mobile combustion, as well as from adipic and nitric acid industrial production together with motor emission control technologies for on-road vehicles. Agriculture soils associated to nitrogen fertilizers remains one of the major sources of N_2O emissions around the world [34, 35].

Microbial metabolic activities associated with the nitrogen biogeochemical cycle are responsible for the direct release of N_2O, as well as for the anthropogenic release arising from agriculture due to the use of nitrogen fertilizers [36–38]. The nitrogen cycle has five major pathways: nitrification, denitrification, nitrogen fixation, anaerobic ammonium oxidation (Anammox) and dissimilatory nitrate reduction to ammonium (DNRA) [39] (Figure 1).

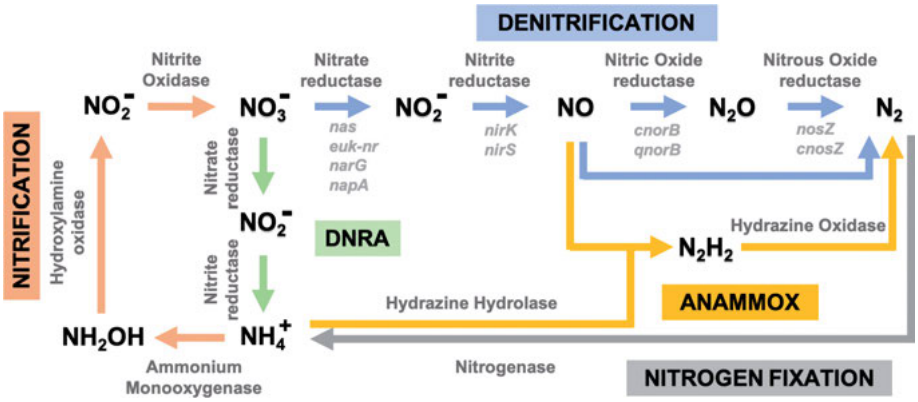


Figure 1. Major pathways of the nitrogen cycle. Different pathways are highlighted in different colors: denitrification in blue, nitrification in light brown, nitrogen fixation in grey, anammox in gold yellow, and dissimilatory nitrate reduction to ammonium (DNRA) in green. The enzymes that catalyze each step are written above the corresponding arrow and the genes encoding the enzymes in denitrification, that is the major N_2O source, are written below the arrows. Figure adapted from [17, 142].

The N_2O release is mainly associated with denitrification and nitrification, as well as with DNRA pathway, though in a lesser extent [39–43]. Denitrification, the process corresponding to anaerobic respiration of nitrate, accounts for the major release of N_2O, since this molecule is the last intermediate in the four-step reduction of nitrate to N_2, and it is the product of a reaction catalyzed by the enzyme NO reductase that is responsible for the reduction of NO to N_2O, forming the N–N bond [32]. Not all microorganisms can perform the complete denitrification [44–46], reducing N_2O to N_2 in a reaction catalyzed by N_2OR, due to either the absence of the gene encoding N_2OR, *nosZ* (Section 3.1.) [47, 48], or by environmental conditions associated with low pH [49–51], presence of dioxygen (O_2) [52–54], carbon dioxide (CO_2) [55, 56], or sulfide (S^{2-}) [57].

In the nitrification pathway three types of microorganisms are involved: (i) the ammonium oxidizers that are responsible for the oxidation of ammonium (NH_4^+) to nitrite (NO_2^-), (ii) the nitrite oxidizers that oxidize nitrite to nitrate (NO_3^-), and (iii) the complete ammonium oxidizers that perform all steps from the oxidation of ammonium to nitrate [39]. In the nitrification pathway, occurring under oxic conditions, N_2O is formed during the oxidation of hydroxylamine (NH_2OH), when nitrite is present in low concentrations, while ammonium exists in high concentrations [58, 59]. Formation of ammonium is accomplished through two different processes, the DNRA or the direct reduction of dinitrogen (N_2) to ammonium (nitrogen fixation). The bacteria and archaea involved in these processes have nitrogenases and in this step N_2O is formed in small amounts concomitantly with ammonium when nitrate/nitrite is being reduced [42, 60, 61]. Recently, it was found that chemodenitrification processes can also contribute to the formation of N_2O, in which Fe(II) reacts with nitrite forming nitric oxide (NO), which can further react with Fe(II) forming N_2O [62] (Section 3.1).

3. NITROUS OXIDE REDUCTASE

N_2OR catalyzes the N_2O reduction to N_2 according to Equation (1). This reaction is the last step of the metabolic denitrification pathway (Section 2.2, Figure 1) and is used by some microorganisms to produce a proton-driven force for ATP synthesis, under anoxic conditions [63]. N_2OR is a copper enzyme containing two different metal centers: (i) the binuclear mixed-valent CuA center [64], that acts as the electron transferring center, and (ii) the catalytic center, CuZ. This is a unique catalytic tetranuclear copper sulfide center, tailored by Nature to overcome the high activation energy of N_2O reduction and is the subject of this chapter.

3.1 Clade I and Clade II of Nitrous Oxide Reductase

N_2OR is encoded in the genome by *nosZ*. The comparative analysis of sequenced genomes encoding N_2OR led to the division of these organisms into two clades according to the gene composition of the *nosZ* operon. In Figure 2, the gene organization of the *nosZ* operon for these two clades is presented for a representative microorganism of each clade.

Clade I N_2ORs have been isolated from proteobacteria of the α-, β-, and γ-division [65], with most of these genomes also encoding genes for the other enzymes of the denitrification pathway. These N_2ORs have been extensively characterized and the majority of the spectroscopic, kinetic, and structural data reported in the literature have been acquired for the Clade I enzymes.

Clade II N_2ORs are not so well characterized and have been shown to be encoded in the genome of some Gram-positive bacteria, such as *Geobacillus thermodenitrificans* and other *Bacillus* species, and in proteobacteria of the δ- and ε-division [66–68]. Contrary to Clade I, these microorganisms have been classified as canonical non-denitrifiers because the genes *nirK* and *nirS*, which encode the copper-type and cytochrome cd_1-type nitrite reductases (see Figure 1), respectively, associated with denitrification are absent from their genomes [18, 47, 60, 61].

Clade II N_2OR has only been isolated from *Wolinella succinogenes* [69], and it was shown to bind a *c*-type heme in an additional C-terminal domain. This feature is a unique trade of *W. succinogenes* N_2OR [69–74], as the analysis of the primary sequence of other members of Clade II does not reveal the presence of the canonical *c*-type heme binding motif, $-C(X)_{2-4}CH-$. Nevertheless, several physiological studies have been performed using *Dechloromonas aromatica* and *Anaeromyxobacter dehalogenans* [61, 62]. These organisms have been gaining interest as they can convert nitrite to N_2 by linking abiotic to biotic reactions, in a process that involves the abiotic reduction of nitrite by Fe(II), as mentioned in Section 2.2, generating N_2O, which is then reduced to N_2 by N_2OR [62]. Thus, although these organisms do not encode the canonical denitrifying genes, they should no longer be classified as non-denitrifiers.

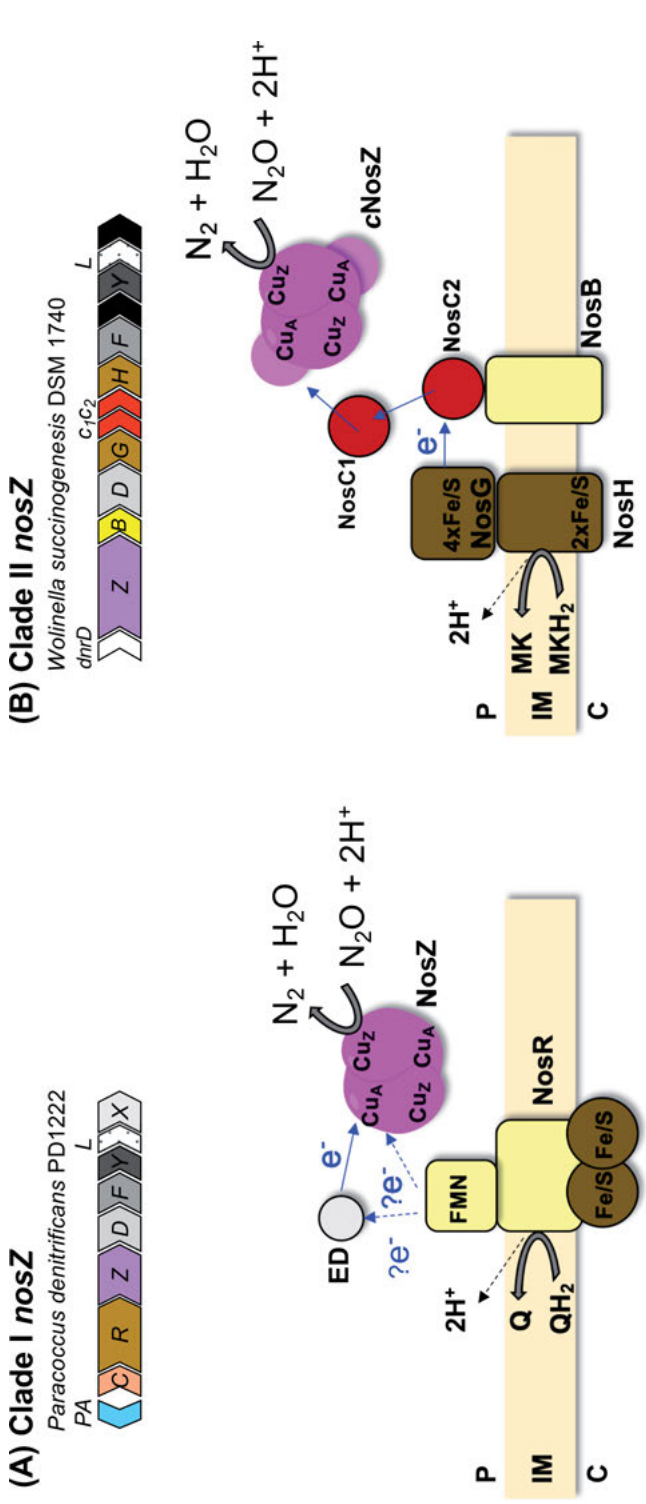


Figure 2. The canonical gene cluster and electron transfer pathway from the quinol/menaquinol pool to Clade I and Clade II N_2OR. Panel (**A**) shows the *nosZ* gene cluster organization of Clade I N_2OR and the proposed electron transfer pathway from NosR to a small electron donor protein (ED) that then transfers the electron to CuA center of N_2OR. The possibility of a direct route from NosR to N_2OR is also represented. Panel (**B**) shows the *nosZ* gene cluster organization of Clade II N_2OR and the proposed electron transfer pathway from membrane-associated NosG/NosH to NosC2, then to NosC1 and finally to the *c*-type heme domain of *c*NosZ of *W. succinogenes.* Figure was prepared based on [77, 143]. PA = pseudoazurin, C = thioredoxin-like protein, *c* = *c*-type cytochrome, Fe/S = iron-sulfur cluster protein. The arrows in black in the gene cluster organization correspond to hypothetical proteins, *dnr* − dissimilative nitrate respiration regulator, Q = quinone, MK = menaquinone.

One major difference between Clade I and Clade II N_2OR is the enzyme's affinity for N_2O, with Clade I N_2OR exhibiting a lower affinity (around 20 μM) than Clade II N_2OR (0.1 μM), though a higher reduction rate is observed for the former [61, 75]. Another feature that distinguishes the enzymes of these two clades is the signal peptide for protein transport to the periplasm in Gram-negative bacteria. Clade I N_2OR is transported to the periplasm by the twin-arginine translocation (Tat) system in the folded state, while Clade II N_2OR is transported through the secretory pathway (Sec) in the unfolded state. In fact, proteins that require *c*-type heme attachment are transported to the periplasm through this system, though this is not the case for all members of Clade II. The reasoning for this difference is still unknown. Moreover, Clade II N_2ORs from Gram-negative bacteria are predicted to be membrane-associated, but none of these enzymes has yet been isolated.

Besides these differences, Clade I and Clade II N_2ORs also differ in the genes that constitute their *nosZ* operons. The genes proposed to be involved in CuZ center assembly (*nosDFYL*) (Section 4.5) are present in both clades, and there are others that differ from each clade. The Clade II *nosZ* operon has two genes, *nosHG*, homologues to the ones encoding the quinol dehydrogenase NapHG [76], and two other encoding *c*-type cytochromes, which together are proposed to compose the electron transfer chain from menaquinol to N_2OR [69, 77] (Figure 2B). In the case of Clade I, *nosR* encodes an Fe-S protein with a periplasmic FMN-binding domain [78, 79], NosR. This protein has been proposed to play a role in maintaining N_2OR in an active state and/or to be part of the electron transfer chain from the quinol pool to N_2OR [18, 78, 79] (Figure 2A). Another gene that is associated with the *nosZ* operon of Clade I N_2OR α-proteobacteria is *nosX*, which encodes a flavoprotein belonging to the ApbE protein family [80]. This protein was shown to be involved in the maturation of NosR by donating its FAD group [79], and explains why a *nosXnirX Paracoccus denitrificans* mutant strain was not able to reduce N_2O [80]. In the other Clade I microorganisms this role is played by a protein of the same family but its gene is usually distantly located in the genome from the *nosZ* operon [75, 79].

3.2 Structure, Biochemical and Catalytic Properties of Nitrous Oxide Reductase

Clade I N_2OR has been isolated from different species and extensively characterized with its structure determined in different oxidation states, in the presence of iodide (Figure 3) and of N_2O [81–84]. These enzymes have also been studied using different spectroscopic techniques, and its kinetic parameters have been determined using either artificial or putative physiological electron donors. On the other hand, the N_2OR from *W. succinogenes* is the only Clade II enzyme isolated so far [70] and its structure is not yet known (only a model structure was proposed considering that this enzyme could be regarded as an electron transfer complex between Clade I N_2OR and a small electron transfer protein [85]).

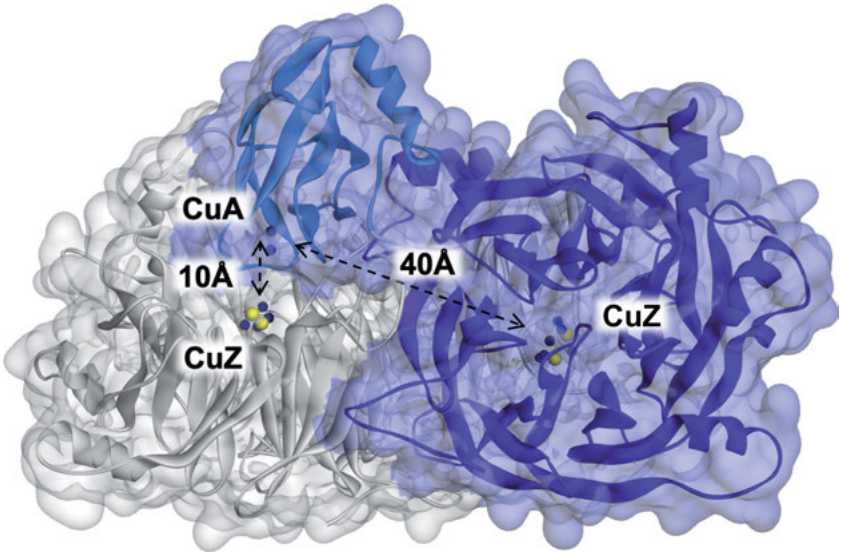


Figure 3. Structure of Clade I *P. stutzeri* N_2OR functional homodimer. The backbone of one monomer is represented with the identified secondary structure colored in blue (N-terminal domain) or light blue (C-terminal domain) with its transparent surface in light blue, and the other monomer is similarly represented in grey. The copper and sulfur atoms of CuA and CuZ centers are represented as blue and yellow spheres, respectively. The distances between CuA and CuZ centers of the two monomers are given. The Figure was prepared with Biovia Discovery Studio using PDB ID 3SBP for *P. stutzeri* N_2OR.

The first report of a Clade I N_2OR dates back to 1972, when a copper enzyme was purified from *Alcaligenes faecalis* [86], but its catalytic activity was only discovered 10 years after for a similar protein isolated from *Pseudomonas stutzeri* [87]. Since then, there have been several reports on its purification from denitrifying microorganisms, with the more extensively studied enzymes being the ones isolated from *P. denitrificans*, *P. stutzeri,* and *Marinobacter hydrocarbonoclasticus*. These are periplasmic dimeric enzymes that bind 12 copper atoms per dimer, arranged in two multicopper centers: a binuclear CuA electron transfer center and a tetranuclear copper sulfide center, CuZ, the active center at which N_2O binds and its reduction occurs.

The analysis of the Clade I *P. denitrificans* N_2OR primary sequence together with its X-ray structure revealed that these two copper centers are organized in two different domains of the enzyme: the N-terminal domain has a β-propeller structure with seven blades (Figure 3), in the center of which the CuZ center is inserted, while the CuA center binds in the loop region between the β8 and β9 strands of the β-barrel C-terminal domain [82]. The two centers in the monomer are 40 Å apart but only at 10 Å, when considering different monomers (head-to-tail arrangement). Thus, the structural arrangement of N_2OR results in a functional dimer since the electron transfer has to occur between CuA and CuZ located in different monomers (Figure 3). The CuA center is

very similar to the CuA centers present in cytochrome *c* oxidase [88] and quinol CuA nitric oxide reductase [89]. It is a binuclear copper center with two bridging cysteines binding the two copper atoms, through their Sγ atom, with the other ligands being two Nε2 atoms of the imidazole ring of two histidines, the Sδ atom of a methionine and the carbonyl of a tryptophan residue. The properties of the CuA center will not be further addressed here, as they are the focus of Chapter 4 in this book [64].

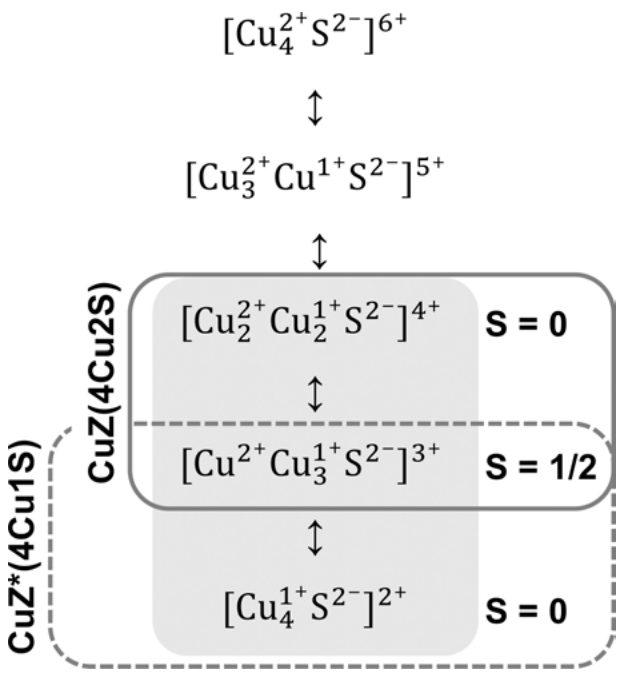
The specific activity of Clade I N_2OR isolated from different organisms has been reported to range from 1 and 10 μmol of N_2O $min^{-1}mg^{-1}$ N_2OR, without activation [90–94], while the one of Clade II *W. succinogenes* N_2OR is 160 μmol of N_2O $min^{-1}mg^{-1}$ N_2OR [70, 73]. The electrons required for the catalytic reduction of N_2O can be donated by small periplasmic *c*-type cytochromes [95–97] or type 1 copper proteins [95, 98], depending on the microorganism, and mitochondrial cytochrome *c* has also been used in some cases as a non-physiological electron donor [99–101]. Besides the *in vitro* assays, whole cell studies have shown that during N_2O reduction oxidation of a cytochrome in the case of *Rhodobacter capsulatus*, *Rhodobacter sphaeroides*, and *P. denitrificans* [96, 102] occurs, and that a *R. capsulatus* cytochrome c_2 knock-out strain was unable to reduce N_2O [97]. This strengthens the hypothesis that small electron transfer proteins are involved in the electron transfer pathway during N_2O reduction. Nevertheless, the involvement of these small electron donor proteins does not exclude that *in vivo* there are other proteins also involved in the electron transfer chain (Figure 2), such as NosR. This hypothesis has been strengthened by the observation that in *Pseudomonas aeruginosa* the denitrification enzymes form a supramolecular complex that includes also NosR [103, 104].

4. THE CATALYTIC CuZ CENTER

4.1 Structure of the CuZ Center

The catalytic center of N_2OR, "CuZ", being a tetranuclear copper center bridged by a sulfur atom, has only been observed in N_2OR so far, in contrast to CuA [105, 106]. Its nuclearity foresees that five different oxidation states can be reached in this center (Scheme 1), but only three have been observed by spectroscopic techniques so far: [$2Cu^{2+}$-$2Cu^{1+}$], [$1Cu^{2+}$-$3Cu^{1+}$] and [$4Cu^{1+}$]. The first two were detected in isolated N_2ORs, while the later can only be obtained *in vitro* through a prolonged incubation with reduced viologens, as will be discussed in Section 4.3. In addition, the CuZ center can be observed in either of the forms, CuZ(4Cu1S) or CuZ(4Cu2S), that have different coordination spheres, as well as different spectroscopic, redox, and kinetic properties.

It is important to mention that N_2OR has never been isolated with the CuZ center in a single form, but usually the preparations are richer in either CuZ(4Cu2S), when isolated under anoxic conditions [93], or in CuZ(4Cu1S)



Scheme 1. Possible oxidation states of the CuZ core center. The oxidation states that have been observed and characterized for CuZ(4Cu1S) and CuZ(4Cu2S) are shaded in grey.

when oxic conditions have been used, or the cells had been stored for a longer time at low temperature prior to enzyme isolation [93, 107].

The tetranuclear core structure of the CuZ was revealed for the first time when the X-ray structure of *M. hydrocarbonoclasticus* N_2OR together with the one of *P. denitrificans* N_2OR, both with the CuZ center mainly as CuZ(4Cu1S), was reported [81, 82]. The CuZ center is a tetranuclear μ_4-sulfide-bridged copper center, adopting a distorted tetrahedral geometry. The Cu_I, Cu_{II}, and Cu_{III} atoms are each coordinated by two histidines, while Cu_{IV} is coordinated by only one (Figure 4). These copper atoms are coordinated by either Nε2 (His80, His128, His270, His325, His376) or Nδ1 (His79 and His437) of the histidine imidazole ring (numbering of the residues according to the primary sequence of *P. denitrificans* N_2OR) and there is also a solvent-derived molecule bridging Cu_I and Cu_{IV} (Figure 4). These histidine residues are located in the blades (His79, His80, His128, His325 and His376) or in the top (His270 and His437) of the β-propeller N-terminal domain [81–84].

The structure of CuZ(4Cu2S) in the $[2Cu^{2+}\text{-}2Cu^{1+}]$ oxidation state was reported 10 years later for *P. stutzeri* N_2OR, isolated under anoxic conditions [84], and revealed a second sulfur atom bridging Cu_I and Cu_{IV} instead of a solvent-derived molecule [84] (Figure 4).

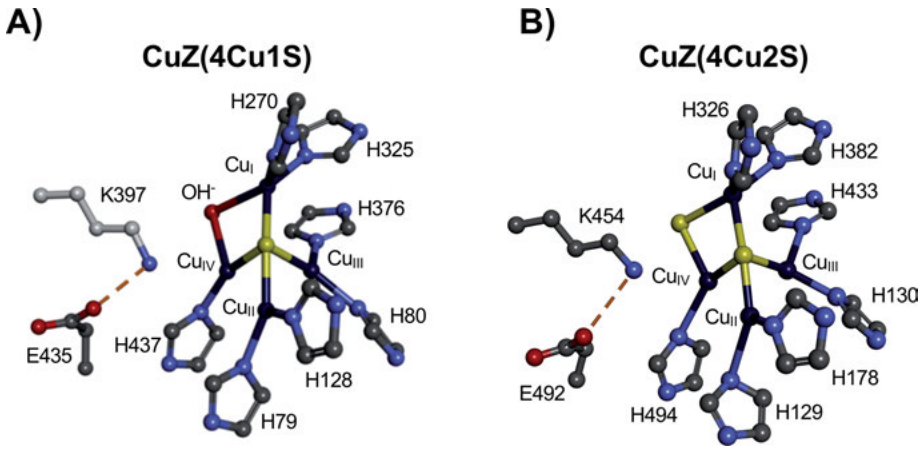


Figure 4. Structure of the different forms of the CuZ center. (**A**) Coordination of CuZ(4Cu1S) in *P. denitrificans* N_2OR and (**B**) coordination of CuZ(4Cu2S) in *P. stutzeri* N_2OR. Figures were prepared with Biovia Discovery Studio using PDB ID 1FWX for *P. denitrificans* N_2OR and PDB ID 3SBP for *P. stutzeri* N_2OR. Color scheme for the atoms: carbon in grey, Cu in dark blue, N in light blue, S in yellow and O in red.

The structural difference between CuZ(4Cu1S) and CuZ(4Cu2S) makes their redox, spectroscopic, and kinetic properties to be different, as discussed in the following Sections.

4.2 Redox and Spectroscopic Properties of CuZ(4Cu1S) and CuZ(4Cu2S)

In the case of CuZ(4Cu1S), the oxidation state of the four copper atoms was determined by Cu K-edge X-ray absorption spectroscopy to be [$1Cu^{2+}$-$3Cu^{1+}$] [108]. Several spectroscopic techniques have been used to characterize this center [108–112], showing that it exhibits an absorption maximum at 640 nm ($\varepsilon_{640nm} \approx 3.5\ mM^{-1}cm^{-1}$ per monomer) and a broad axial electron paramagnetic resonance (EPR) signal ($g_{\parallel} = 2.16$ and $g_{\perp} \approx 2.04$) with poorly resolved hyperfine-split lines in the parallel region [107, 108, 113–117] (Table 1). Theoretical calculations using the structure of this center were instrumental to interpret the spectroscopic data and show that the spin density, $S_{Total} = ½$, is mainly delocalized over Cu_I (26 %) and Cu_{IV} (12 %), with 26 % on the bridging sulfur [118]. The nature of the Cu_I-Cu_{IV} edge was identified examining the density functional theory (DFT) models of CuZ(4Cu1S) with different edge ligands (bridging H_2O, H_2O bound to Cu_I, bridging OH^-, or OH^- bound to Cu_{IV}, and H-bonded to a protonated lysine residue) [110–112]. The data showed that the spectroscopic properties of CuZ(4Cu1S) were better explained considering a model in which a OH^- ligand occupies the Cu_I-Cu_{IV} edge, closer to Cu_I (2.00 Å) than to Cu_{IV} (2.09 Å), and in which Lys397 and Glu435 are H-bonded to each other [112, 119] (numbering according to *P. denitrificans* N_2OR amino acid sequence).

Table 1. The different oxidation states, spectroscopic data, and activities towards N_2O for the different forms of the CuZ center of N_2OR.

"CuZ" form	Oxidation state (Cu_I – Cu_{IV} ligand)	Absorption maximum	Spin state	EPR	Turnover number (N_2O)	Ref.
CuZ(4Cu1S)	[$1Cu^{2+}$:$3Cu^{1+}$:S:OH] (bridging OH^-)	640 nm (~3.5 mM^{-1} cm^{-1})[a]	S=1/2	$g_{\parallel}$ = 2.160, $g_{\perp}$ = 2.040 $A_{\parallel}$ = 6.1 mT/$A_{\parallel}$ = 2.4 mT[b]	0	[93, 107, 108, 111, 113, 116, 122]
	[$4Cu^{1+}$:S] (empty)	No bands	S=0	Silent	321 s^{-1}	[90, 112, 119, 123, 124]
CuZ^0(4Cu1S)	[$1Cu^{2+}$:$3Cu^{1+}$:S:OH] (Cu_{IV}-OH^-)	680 nm (~2.0 mM^{-1} cm^{-1})[a]	S=1/2	$g_{\parallel}$ = 2.177, $g_{\perp}$ = 2.05 $A_{\parallel}$ = 4.2 mT[c]	321 s^{-1}	[118, 121]
CuZ(4Cu2S)	[$2Cu^{2+}$:$2Cu^{1+}$:2S] (bridging S^{2-})	545 nm (~5.0 mM^{-1} cm^{-1})[a]	S=0	Silent	0	[90, 93, 116, 119]
	[$1Cu^{2+}$:$3Cu^{1+}$:2S] (bridging SH^-)	670 nm (~3.0–4.4 mM^{-1} cm^{-1})[a]	S=1/2	$g_{\parallel}$ = 2.150, $g_{\perp}$ = 2.035 $A_{\parallel}$ = 5.6 mT[d]	0.6 h^{-1}	[90, 93, 110, 116, 119, 123, 125, 126]

N.D. – Not determined.
[a]Extinction coefficients given per N_2OR monomer.
[b]With a 5:2 ratio.
[c]Considering two identical hyperfine coupling constants.
[d]Considering three identical hyperfine coupling constants.

CuZ(4Cu1S), in the [$1Cu^{2+}$-$3Cu^{1+}$] oxidation state, cannot be oxidized by potassium ferricyanide and it is also not easily reduced *in vitro* by just sodium dithionite (–471 mV versus SHE at pH 7.0 [120]). Nevertheless, the [$4Cu^{1+}$] oxidation state is observed after a prolonged incubation with reduced methyl or benzyl viologen (Table 1), but the redox potential of the [$1Cu^{2+}$-$3Cu^{1+}$]/[$4Cu^{1+}$] pair could not be determined by potentiometry as it is an irreversible process [121].

Contrary to CuZ(4Cu1S), CuZ(4Cu2S) can exist in either the [$2Cu^{2+}$-$2Cu^{1+}$] or the [$1Cu^{2+}$-$3Cu^{1+}$] oxidation state, in an equilibrium that has a reduction potential of +60 mV, at pH 7.5 [93], but the [$4Cu^{1+}$] oxidation state has never been observed so far, not even *in vitro*.

In the [$2Cu^{2+}$-$2Cu^{1+}$] oxidation state, CuZ(4Cu2S) is characterized by an absorption maximum at 550 nm ($\varepsilon_{550nm} \approx 5.0$ $mM^{-1}cm^{-1}$ per monomer), while the [$1Cu^{2+}$-$3Cu^{1+}$] oxidation state exhibits an absorption maximum around 670 nm (ε_{670nm} 3.0–4.4 $mM^{-1}cm^{-1}$ per monomer) (Table 1) [93, 94, 119]. Of these two oxidation states, the only EPR-active state of CuZ(4Cu2S) is [$1Cu^{2+}$-$3Cu^{1+}$], since magnetic circular dichroism (MCD) showed that the [$2Cu^{2+}$-$2Cu^{1+}$] oxidation state was diamagnetic. The EPR spectrum of CuZ(4Cu2S) in the [$1Cu^{2+}$-$3Cu^{1+}$] oxidation state exhibits an axial signal ($g_{\parallel} > g_{\perp} > 2.0$), with five evenly spaced hyperfine lines in the $g_{\parallel}$ region [119]. This signal was interpreted considering three identical ^{63,65}Cu hyperfine coupling constants of 5.6 mT [119] (Table 1), with the spin density being distributed over Cu_I (17 %), Cu_{II} (11 %), and Cu_{IV} (10 %), with the remaining spin density being over μ_4-sulfide (34 %), μ_2-SH^- (16 %), and Cu_{III} (6 %) [119].

The protonation state of the bridging sulfur ligand was assessed in the two oxidation states by studying the pH profile of the CuZ(4Cu2S) resonance Raman spectra in combination with DFT calculations. The observation of D_2O-isotope sensitive vibration modes, identified as S-H bending modes, indicated that the Cu_I-Cu_{IV} edge is a sulfide (μ_2S^{2-}), with a p$K_a \leq 3$, in the [$2Cu^{2+}$-$2Cu^{1+}$] oxidation state, and a hydrosulfide (μ_2SH^-) with a p$K_a \geq 11$ in the [$1Cu^{2+}$-$3Cu^{1+}$] oxidation state [119] (Table 1).

4.3 Activation of the CuZ Center and Binding of Nitrous Oxide

As mentioned in Section 3.2., the specific activity of isolated Clade I N_2OR is very low (1–10 μmol of N_2O $min^{-1}mg^{-1}$ N_2OR [90–94]), and does not explain the N_2O reduction rates observed for whole cells or crude cell extracts (48–72 μmol of N_2O $min^{-1}mg^{-1}$ N_2OR [75, 90, 127]). This observation led to the hypothesis that this enzyme requires activation and is isolated in an unready state [10]. On the contrary, the specific activity of Clade II *W. succinogenes* N_2OR was reported to be 160 μmol of N_2O $min^{-1}mg^{-1}$ N_2OR [70, 73], which might mean that Clade II N_2ORs are isolated in a ready state. However, kinetic studies on other enzymes from this clade are needed to confirm this hypothesis.

The specific activity of Clade I N_2ORs with CuZ(4Cu1S) increases during incubation with reduced methyl viologen and reaches a maximum value of 200 μmol

of N_2O $min^{-1}mg^{-1}$ N_2OR after 3–5 h, in the case of *M. hydrocarbonoclasticus* N_2OR [99, 119]. During this time, by monitoring the EPR signal of the enzyme, it could be shown that the activation step is the reduction of CuZ(4Cu1S) from the $[1Cu^{2+}\text{-}3Cu^{1+}]$ to the $[4Cu^{1+}]$ oxidation state [124, 128], as the EPR signal decreases, evidencing the formation of a diamagnetic species. The rate constant of this activation process was estimated to be 1.2×10^{-3} s^{-1} at pH 7.0 for *M. hydrocarbonoclasticus* N_2OR [124], indicating that it is too slow to be part of the catalytic cycle of the enzyme with a turnover number of 321 s^{-1} (turnover number of *M. hydrocarbonoclasticus* N_2OR with CuZ(4Cu1S) in the $[4Cu^{1+}]$ oxidation state) (Table 1) [123, 128].

In the case of CuZ(4Cu2S), the only catalytically competent oxidation state is $[1Cu^{2+}\text{-}3Cu^{1+}]$, but its turnover number is much lower, 0.6 h^{-1} [123] (Table 1). Thus, it can be proposed that *in vivo* an activation step must occur prior to catalysis and that while N_2O is available, N_2OR is kept active by (a) still unidentified protein(s), avoiding inactivation after each catalytic cycle.

In fact, there has been a debate between the different research groups studying N_2OR as whether the $[4Cu^{1+}]$ oxidation state is ever attained *in vivo*. Although, this question has not yet been answered, the analysis of the kinetic parameters of N_2OR with CuZ(4Cu1S) in the $[4Cu^{1+}]$ oxidation state together with the estimation of the reduction rates of whole cells can help to clarify this question. In the case of *M. hydrocarbonoclasticus* N_2OR, a V_{max} of 200 μmol of N_2O $min^{-1}mg^{-1}$ N_2OR and a K_M of 18 μM were estimated from the kinetic assays using reduced methyl viologen as electron donor [99]. A similar affinity constant was determined for these cells in the reduction of exogenously added N_2O. Moreover, the estimated V_{max} for the isolated N_2OR explains the high N_2O-reduction rate observed for the cells growing in the presence of nitrate, when considering the yield of the enzyme purification [75]. Hence, these data corroborated the hypothesis that N_2OR with the CuZ center in the $[4Cu^{1+}]$ oxidation state can be the ready state of this enzyme *in vivo*.

One other argument for CuZ(4Cu1S) being part of the catalytic cycle of the enzyme *in vivo* was the observation that iodide, a proposed inhibitor of N_2OR, binds to the CuZ center at the $Cu_I\text{-}Cu_{IV}$ edge [83], where a solvent-derived molecule has been modeled in the structure of *Achromobacter cycloclastes* N_2OR with CuZ(4Cu1S) in the $[1Cu^{2+}\text{-}3Cu^{1+}]$ oxidation state [81–83] (Figure 5). This observation was used to model the binding of N_2O to the "CuZ" center by DFT calculations, with the μ-1,3-N_2O coordination at the $Cu_I\text{-}Cu_{IV}$ edge of CuZ(4Cu1S) in the $[4Cu^{1+}]$ oxidation state being found to be the most favorable binding mode for the substrate (Figure 5) [111, 117, 118, 124]. However, it can be argued that there has not been any other experimental evidence for the binding of the substrate to CuZ(4Cu1S). In an attempt to experimentally observe the binding of N_2O to CuZ(4Cu2S), the crystals of *P. stutzeri* N_2OR with CuZ(4Cu2S) in the $[2Cu^{2+}\text{-}2Cu^{1+}]$ oxidation state were pressurized with N_2O [84]. It was observed that the substrate was over the CuZ center [84], at a distance such that it cannot be considered to be coordinating any of the copper atoms. Moreover, it was later shown that in this oxidation state CuZ(4Cu2S) does not react with N_2O [123], so this structural rearrangement is not relevant for the catalytic cycle.

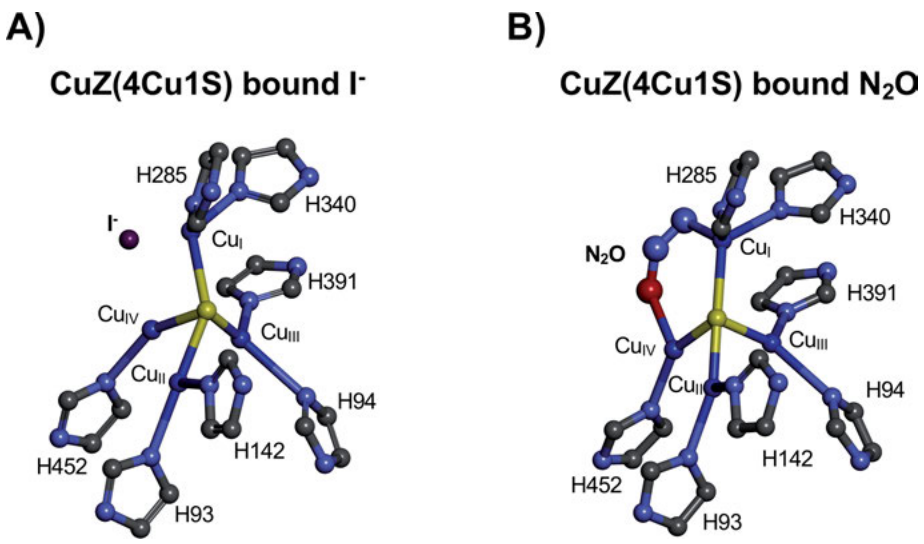


Figure 5. Structure of the CuZ center of *A. cycloclastes* N_2OR. (**A**) Coordination of iodide-bound CuZ(4Cu1S) and (**B**) of N_2O-bound CuZ(4Cu1S) (N_2O was modelled into the structure with data from [124]). Figures were prepared with Biovia Discovery Studio using PDB ID 2IWK for *A. cycloclastes* inhibitor-bound N_2OR and PDB ID 2IWF for *A. cycloclastes* N_2O-bound N_2OR. Color scheme for the atoms: carbon in grey, Cu in dark blue, N in light blue, S in yellow, O in red and I in violet.

In conclusion, even if CuZ(4Cu1S) is obtained when N_2OR is isolated under conditions that can be considered to damage its catalytic center (presence of dioxygen), this is the only form of the CuZ with high specific activity. One of the remaining questions to be answered is how CuZ(4Cu2S) is activated and whether this process involves the displacement of the μ_2-sulfur bridging ligand and whether NosR or another membrane-associated protein complex plays a role in the activation mechanism and catalytic activity of N_2OR.

4.4 Catalytic Cycle of CuZ(4Cu1S)

The catalytic cycle of CuZ(4Cu1S) has been proposed recently (Figure 6). The catalysis starts with the CuZ center in the fully reduced state, [$4Cu^{1+}$] (intermediate 1), reacting with the substrate, N_2O, to form intermediate 2. During this process, as mentioned in Section 4.3, N_2O binds with its terminal N to Cu_I in a linear configuration [118], and elongation of the N–O bond leads to the rearrangement of its structure, so that it binds at the Cu_I-Cu_{IV} edge in a μ-1,3-N_2O coordination forming a 139º N–N–O bond angle. In this intermediate 2, the oxygen atom is H-bonded to the protonated amino group of Lys397 [124, 129], but this intermediate has not yet been isolated.

The release of N_2 requires two electrons that will be transferred, via Cu_{IV}, from the fully reduced CuZ(4Cu1S) [$4Cu^{1+}$], in a proton-coupled process, with

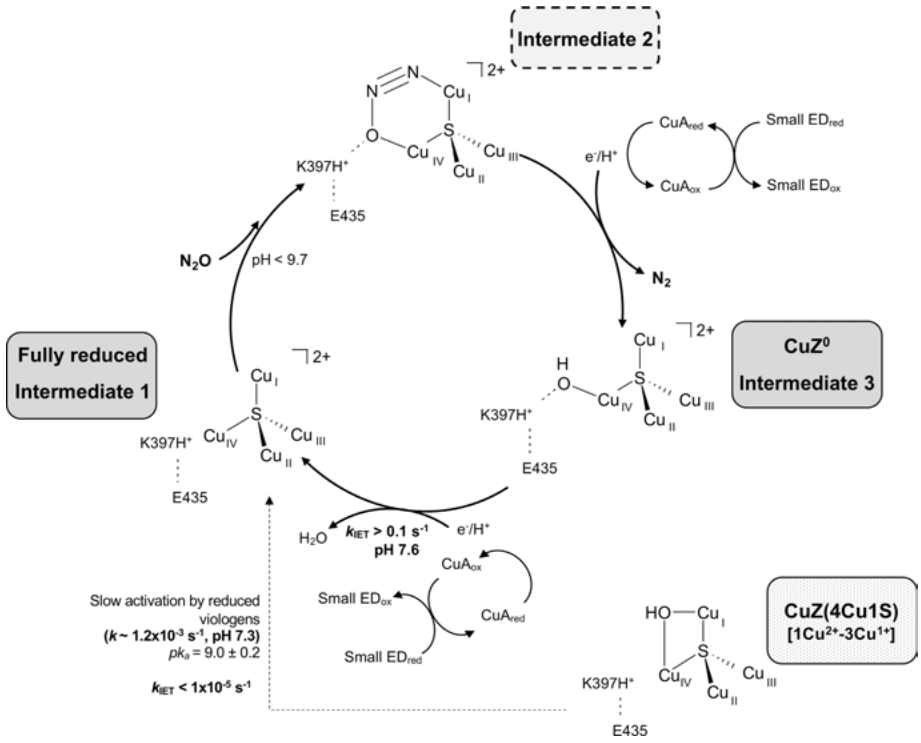


Figure 6. Catalytic cycle of N_2O reduction by N_2OR with CuZ(4Cu1S). Intermediates 1 and 3 have been trapped and characterized, while intermediate 2 remains to be characterized. The scheme also shows the reduction of the $[1Cu^{2+}\text{-}3Cu^{1+}]$ oxidation state to $[4Cu^{1+}]$ by reduced viologens (activation), which is not part of the catalytic cycle. Residues are numbered according to *P. denitrificans* N_2OR mature amino acid sequence. ED = electron donor.

all four copper atoms of CuZ(4Cu1S) being involved [118]. The cleavage of the N–O bond requires one electron, with one proton being simultaneously transferred from Lys397 to the oxygen, that becomes coordinated to Cu_{IV} as a hydroxide. The second electron is transferred, cleaving the Cu_I–N bond and leading to the release of N_2. Re-protonation of Lys397, with a proton from the solvent, coupled with electron transfer from the CuA center leads to the formation of CuZ^0, intermediate 3. In this intermediate, the protonated form of Lys397 is H-bonded to the hydroxide ligand of Cu_{IV}, and to Glu435 to stabilize the CuZ^0 intermediate [118]. The catalytic cycle is closed by the rapid intramolecular electron transfer via the CuA center ($k_{IET} > 0.1\ s^{-1}$) [118] (Figure 6).

The first intermediate, CuZ(4Cu1S) in the $[4Cu^{1+}]$ oxidation state, has all its copper atoms in a d^{10} configuration and thus it is spectroscopically silent, but upon stoichiometric reaction with N_2O the intermediate CuZ^0 is observed. This species is characterized by an absorption maximum at 680 nm ($\varepsilon_{680nm} \approx 2.0\ mM^{-1}cm^{-1}$) and it reacts rapidly with the substrate without further activation [118, 121], indi-

cating that it is directly involved in the catalytic cycle of the enzyme. CuZ^0 has g values similar to those observed for CuZ(4Cu1S) in the $[1Cu^{2+}\text{-}3Cu^{1+}]$ oxidation state: $g_{\parallel} = 2.177 > g_{\perp} = 2.05 > 2.0$ [121], with two equal $^{63,65}Cu$ hyperfine coupling constants ($A_{\parallel} = 4.2$ mT) to account for the 5-line hyperfine pattern in the parallel region (Table 1). This means that the total spin in CuZ^0 is equally distributed between Cu_I and Cu_{IV}, while as it was mentioned before in CuZ(4Cu1S) in the same oxidation state Cu_I has a higher spin density.

The analysis of resonance Raman and MCD spectroscopic data used to characterize CuZ^0 corroborated that the four copper atoms are in the $[1Cu^{2+}\text{-}3Cu^{1+}]$ oxidation state [118], and that Cu_I and Cu_{IV} have a different coordination sphere from the one in CuZ(4Cu1S) in the same oxidation state. On the basis of DFT calculations the most plausible structure is the one presented in Figure 6, with a hydroxo ligand terminally coordinating Cu_{IV} with a H bonded to a protonated lysine [118]. Thus, in CuZ^0 there is no bridging ligand between Cu_I and Cu_{IV}.

CuZ^0 can be reduced intramolecularly by CuA *in vitro*, in an experiment using a more physiologically relevant electron donor than methyl viologen, sodium ascorbate (with a reduction potential of $+60$ mV) [118]. These experiments showed that the intramolecular electron transfer rate constant is higher than $0.1\ s^{-1}$ [118].

4.5. CuZ Center Assembly

The CuZ center is a complex metal center with four copper atoms bound to an inorganic sulfur and coordinated by the imidazole rings of seven histidine residues that are bound to the center of a seven-bladed β-propeller domain. Thus, it is not solvent-exposed as other copper centers and attempts to heterologously produce N_2OR in microorganisms that do not have the *nosZ* operon in their genome have failed [130]. This indicates that the assembly machinery of the CuZ center might be as complex as the ones of Fe-S or *c*-type heme-containing proteins.

The presence of conserved genes in the *nos*Z operons of both Clade I and Clade II N_2ORs (Figure 2), *nosDFYL*, [9] has led to the hypothesis that the encoded proteins could play a role in the activity or metal center assembly of these enzymes. The metal center assembly of N_2OR from either clade has not been extensively explored but it is proposed to occur in the periplasm of Gram-negative bacteria, to which the apo-N_2OR is transported either in the folded or unfolded state by the Tat or the Sec system, respectively (Section 3.1). In fact, the X-ray structure of Clade I *Shewanella denitrificans* apo-N_2OR has been reported [130]. In this structure, the two domains have a similar fold as in the holo-N_2OR, and in the absence of a structural Ca^{2+} ion the regions surrounding the copper centers show structural disorder. Notably, the binding of Ca^{2+} and ordering of the protein backbone prevent further access to the two copper centers. Most likely, the polypeptide chain requires some flexibility for proper copper insertion.

In the case of the CuA center, it has been proposed that its assembly is dependent on the same molecular system as the one involved in the assembly of CuA in cytochrome *c* oxidase [131], while the CuZ center assembly follows a different route. One evidence that there is a different route for CuZ assembly is the fact that N_2OR purified from *P. stutzeri* mutant strains with *nosDFY* knocked out only presents spectroscopic features of CuA [115, 132]. A similar result was obtained when analyzing the spectroscopic properties of N_2OR isolated from a *Pseudomonas putida* strain expressing only *nosZnosR* [133].

The product of *nosDFY* has been proposed to form an ABC transporter used for sulfur mobilization, transport, and insertion into the CuZ [134, 135]. The cytoplasmic component, NosF, is proposed to hydrolyze ATP to transport sulfur in an energy-dependent manner across NosY. NosF has a "Walker A" motif in its primary sequence and other features of an ATPase [134, 136], with its activity being Mg^{2+}-dependent with a K_M of 3 mM for the hydrolysis of ATP and 10 mM for GTP [136]. NosY is a six-span membrane protein proposed to form a pore connecting the cytoplasmic NosF with the periplasmic NosD [131]. NosD has not yet been isolated, but the analysis of its primary sequence proposes that it has two β-helical domains with 4/5 parallel β-helix repeats with homology to proteins belonging to the carbohydrate-binding and sugar hydrolase protein family [135].

NosL is a small periplasmic outer-membrane bound protein classified as a copper chaperone [137–139]. It has two homologous domains with a β-β-α-β topology and binds specifically Cu^{1+} in a 1:1 stoichiometry [138, 140]. Although it was proposed that NosL was not essential for the CuZ center assembly [133, 141], its presence in the *P. aeruginosa* supramolecular complex involving N_2OR and other enzymes participating in the denitrification pathway [103, 104], contradicted those initial findings. Recent work by the team of Gates and Le Brun has shown that in fact NosL is a dedicated copper chaperon essential for CuZ assembly under certain environmental conditions, such as low copper concentrations [140].

In conclusion, little is known about N_2OR biogenesis. It is expected that in the near future a better characterization of the proteins proposed to be involved in CuA and CuZ center assemblies will shed more light into this process, together with the isolation of the intermediate forms of N_2OR during this process. Moreover, it cannot be ruled out that more proteins and enzymes might be involved in this process, besides the ones mentioned here, considering the complex nature of N_2OR and in particular of the CuZ center.

5. GENERAL CONCLUSIONS AND FUTURE PERSPECTIVES

Due to the growing importance of N_2O emissions to global warming, the development of mitigation strategies is extremely important to control N_2O emissions. N_2OR catalyzes the reduction of N_2O to N_2 and studies characterizing the CuZ

active center of this enzyme, together with environmental stresses able to inactivate N_2OR function or production, represent an important contribution for the design of successful strategies to control atmospheric N_2O release by anthropogenic and natural sources.

The high complexity associated with the CuZ of N_2OR cannot be separated from the large activation barrier needed to overcome N_2O reduction to generate N_2 and water. N_2OR is a functional dimer with a unique catalytic copper center that can be isolated in two different forms: CuZ(4Cu1S) and CuZ(4Cu2S). These two centers have different specific activities, with CuZ(4Cu2S) reacting with N_2O with a low turnover number (0.6 h^{-1}), only in the [$1Cu^{2+}$-$3Cu^{1+}$] oxidation state, while CuZ(4Cu1S), with a high turnover number activity in the [$4Cu^{1+}$] oxidation state, is considered to be the catalytic form that explains the high N_2O reduction rate of whole cells. This reduced form of CuZ(4Cu1S) reacts stoichiometrically with N_2O to complete the catalytic cycle, and in the absence of reducing power an intermediate is formed, CuZ^0. This intermediate is in the [$1Cu^{2+}$-$3Cu^{1+}$] oxidation state with a hydroxide bound to Cu_{IV}, that can be displaced by the substrate, when there are enough available electrons to continue the catalytic cycle.

In the future, the elucidation of the coordination sphere of CuZ(4Cu1S) in the [$4Cu^{1+}$] oxidation state and of CuZ^0, will be instrumental to understand their high reactivity and point out changes in the positioning of residues that could be involved in electron transfer and/or in stabilizing *in vivo* the active state of CuZ. In fact, it is still unknown whether N_2OR requires activation *in vivo* and how its active form is maintained. The isolation of N_2OR with CuZ(4Cu2S), which has a low specific activity, raised the hypothesis that this could be a protective form of the enzyme when the substrate or electrons are not available. However, the mechanism to either remove the μ_2-briding sulfur, or to increase its specific activity is still unknown. There is the possibility that accessory proteins, such as NosR, could play a role in these processes, but this hypothesis requires experimental validation.

The biogenesis of N_2OR that involves the assembly of two metal centers, CuA and "CuZ" needs to be further explored. The CuA center being similar to the one present in cytochrome *c* oxidase, could be assembled using a similar pathway without the need for specific proteins, however, the same is not expected for the CuZ center. The assembly machinery of CuZ is proposed to be encoded in the *nosZ* operon, by *nosDFYL*, but only a few of these proteins have been characterized and their involvement in Cu or S assembly/transport/delivery remains to be identified.

The better and more complete understanding of the CuZ center activity and assembly will for sure be of use to design strategies to mitigate the concentration of N_2O in the stratosphere, as well as control its emissions from soils and water environments and have an impact on the climate change.

ACKNOWLEDGMENTS

We would like to thank C. Carreira, S. Dell'Acqua, I. Cabrito and E. Johnston, for their contribution for the work presented here. We would also like to acknowledge the contributions of J. J. G. de Moura, C. Cambillau, E. Solomon, and O. Einsle for the fruitful discussions and collaborative work during the past years.

The authors would like to thank Fundação para a Ciência e Tecnologia (FCT) for the financial support provided (PTDC/BIA-PRO/098882/2008 to SRP and PTDC/BBB-BQB/0129/2014 to IM). This work was further supported by the Associate Laboratory for Green Chemistry- LAQV and Unidade de Ciências Biomoleculares Aplicadas-UCIBIO, which is financed by national funds from FCT/MCTES (UID/QUI/50006/2019 and UID/Multi/04378/2019, respectively). MSPC acknowledges FCT/MCTES for funding her "Research Position" (signed with FCT NOVA in accordance with DL.57/2016 and Lei 57/2017).

ABBREVIATIONS AND DEFINITIONS

Anammox	anaerobic ammonium oxidation
ATP	adenosine 5′-triphosphate
DFT	density functional theory
DNRA	dissimilatory nitrate reduction to ammonium
EPR	electron paramagnetic resonance
FAD	flavin adenine dinucleotide
FMN	flavin mononucleotide
GTP	guanosine 5′-triphosphate
N_2OR	nitrous oxide reductase
MCD	magnetic circular dichroism
Sec	secretory pathway
SHE	standard hydrogen electrode
Tat	twin-arginine translocation

REFERENCES

1. Climate Change 2007: The Physical Science Basis, Cambridge University Press, **2007**.
2. Climate Change 2014: Mitigation of Climate Change, Cambridge University Press, **2014**.
3. A. R. Mosier, *Biol. Fertil. Soils* **1998**, *27*, 221–229.
4. M. J. Prather, *Science* **1998**, *279*, 1339–1341.
5. A. R. Ravishankara, J. S. Daniel, R. W. Portmann, *Science* **2009**, *326*, 123–125.
6. W. C. Trogler, *J. Chem. Educ.* **1995**, *72*, 973–976.
7. G. A. Vaughan, P. B. Rupert, G. L. Hillhouse, *J. Am. Chem. Soc.* **1987**, *109*, 5538–5539.
8. L. Pauling, *Proc. Natl. Acad. Sci. USA* **1932**, *18*, 498–499.

9. S. R. Pauleta, S. Dell'Acqua, I. Moura, *Coord. Chem. Rev.* **2013**, *257*, 332–349.
10. S. R. Pauleta, C. Carreira, I. Moura, in *Metalloenzymes in Denitrification: Applications and Environmental Impacts*, Eds I. Moura, J. J. G. Moura, S. R. Pauleta, L. Maia, Royal Society of Chemistry, Cambridge, **2017**, pp. 141–169.
11. C. Esmieu, M. Orio, S. Torelli, L. Le Pape, J. Pecaut, C. Lebrun, S. Menage, *Chem. Sci.* **2014**, *5*, 4774–4784.
12. F. Mus, D. R. Colman, J. W. Peters, E. S. Boyd, *Free Radic. Biol. Med.* **2019**. 10.1016/j.freeradbiomed.2019.01.050.
13. J. W. Peters, G. J. Schut, E. S. Boyd, D. W. Mulder, E. M. Shepard, J. B. Broderick, P. W. King, M. W. Adams, *Biochim. Biophys. Acta* **2015**, *1853*, 1350–1369.
14. J. H. Jeoung, J. Fesseler, S. Goetzl, H. Dobbek, *Met. Ions Life Sci.* **2014**, *14*, 37–69.
15. S. Horrell, D. Kekilli, R. W. Strange, M. A. Hough, *Metallomics* **2017**, *9*, 1470–1482.
16. L. B. Maia, J. J. Moura, *J. Biol. Inorg. Chem.* **2015**, *20*, 403–433.
17. S. R. Pauleta, M. S. P. Carepo, I. Moura, *Coord. Chem. Rev.* **2019**, *387*, 436–449.
18. W. G. Zumft, P. M. Kroneck, *Adv. Microb. Physiol.* **2007**, *52*, 107–227.
19. K. Jones, in *Comprehensive Inorganic Chemistry*, Eds J. C. Bailar, H. J. Emelöus, R. Nyholm, A. F. Trottman-Dickenson, Pergamon Press, Oxford, **1975**, pp. 147–388.
20. D. J. Wuebbles, *Science* **2009**, *326*, 56–57.
21. W. Ye, L. Bian, C. Wang, R. Zhu, X. Zheng, M. Ding, *J. Environ. Sci.* **2016**, *47*, 193–200.
22. K. Lassey, M. Harvey, *Water & Atmosphere* **2007**, *15*, 10–11.
23. R. W. Portmann, J. S. Daniel, A. R. Ravishankara, *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **2012**, *367*, 1256–1264.
24. A. Miyamoto, S. Baba, M. Mori, Y. Murakami, *J. Physic. Chem.* **1981**, *85*, 3117–3122.
25. R. Zeng, M. Feller, Y. Ben-David, D. Milstein, *J. Am. Chem. Soc.* **2017**, *139*, 5720–5723.
26. A. Dandekar, M. A. Vannice, *Appl. Catal. B* **1999**, *22*, 179–200.
27. J. Zhai, A. S. Filatov, G. L. Hillhouse, M. D. Hopkins, *Chem. Sci.* **2016**, *7*, 589–595.
28. S. Bagherzadeh, N. P. Mankad, *Chem. Commun.* **2018**, *54*, 1097–1100.
29. I. Bar-Nahum, A. K. Gupta, S. M. Huber, M. Z. Ertem, C. J. Cramer, W. B. Tolman, *J. Am. Chem. Soc.* **2009**, *131*, 2812–2814.
30. B. J. Johnson, W. E. Antholine, S. V. Lindeman, M. J. Graham, N. P. Mankad, *J. Am. Chem. Soc.* **2016**, *138*, 13107–13110.
31. N. Gruber, J. N. Galloway, *Nature* **2008**, *451*, 293–296.
32. J. N. Galloway, A. R. Townsend, J. W. Erisman, M. Bekunda, Z. Cai, J. R. Freney, L. A. Martinelli, S. P. Seitzinger, M. A. Sutton, *Science* **2008**, *320*, 889–892.
33. J. N. Galloway, F. J. Dentener, D. G. Capone, E. W. Boyer, R. W. Howarth, S. P. Seitzinger, G. P. Asner, C. C. Cleveland, P. A. Green, E. A. Holland, D. M. Karl, A. F. Michaels, J. H. Porter, A. R. Townsend, C. J. Vöosmarty, *Biogeochemistry* **2004**, *70*, 153–226.
34. Inventory of U. S. Greenhouse Gas Emissions and Sinks: 1990–2017, United States Environmental Protection Agency, **2019**.
35. Annual European Union greenhouse gas inventory 1990–2014 and inventory report 2016, European Environmental Agency, **2016**.
36. T. J. Griffis, Z. Chen, J. M. Baker, J. D. Wood, D. B. Millet, X. Lee, R. T. Venterea, P. A. Turner, *Proc. Natl. Acad. Sci. USA* **2017**, *114*, 12081–12085.
37. I. Shcherbak, N. Millar, G. P. Robertson, *Proc. Natl. Acad. Sci. USA* **2014**, *111*, 9199–9204.
38. W. H. Schlesinger, *Proc. Natl. Acad. Sci. USA* **2009**, *106*, 203–208.
39. L. Y. Stein, M. G. Klotz, *Curr. Biol.* **2016**, *26*, R94–98.
40. L. Y. Stein, Y. L. Yung, *Annu. Rev. Earth and Planet. Sci.* **2003**, *31*, 329–356.
41. L. Y. Stein, M. G. Klotz, *Biochem. Soc. Trans.* **2011**, *39*, 1826–1831.

42. S. Hallin, L. Philippot, F. E. Loffler, R. A. Sanford, C. M. Jones, *Trends Microbiol.* **2018**, *26*, 43–55.
43. G. Braker, R. Conrad, *Adv. Appl. Microbiol.* **2011**, *75*, 33–70.
44. H. Shoun, S. Fushinobu, L. Jiang, S. W. Kim, T. Wakagi, *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **2012**, *367*, 1186–1194.
45. K. Maeda, A. Spor, V. Edel-Hermann, C. Heraud, M. C. Breuil, F. Bizouard, S. Toyoda, N. Yoshida, C. Steinberg, L. Philippot, *Sci. Rep.* **2015**, *5*, 9697.
46. S. A. Higgins, A. Welsh, L. H. Orellana, K. T. Konstantinidis, J. C. Chee-Sanford, R. A. Sanford, C. W. Schadt, F. E. Loffler, *Appl. Environ. Microbiol.* **2016**, *82*, 2919–2928.
47. D. R. Graf, C. M. Jones, S. Hallin, *PLoS One* **2014**, *9*, e114118.
48. L. Philippot, J. Andert, C. M. Jones, D. Bru, S. Hallin, *Global Change Biol.* **2010**, *17*, 1497–1504.
49. B. Liu, P. T. Mørkved, A. Frostegård, L. R. Bakken, *FEMS Microbiol. Ecol.* **2010**, *72*, 407–417.
50. R. N. Van Den Heuvel, S. E. Bakker, M. S. M. Jetten, M. M. Hefting, *Geobiology* **2011**, *9*, 294–300.
51. L. Bergaust, Y. Mao, L. R. Bakken, A. Frostegård, *Appl. Environ. Microbiol.* **2010**, *76*, 6387–6396.
52. P. Lycus, M. J. Soriano-Laguna, M. Kjos, D. J. Richardson, A. J. Gates, D. A. Milligan, A. Frostegard, L. Bergaust, L. R. Bakken, *Proc. Natl. Acad. Sci. USA* **2018**, *115*, 11820–11825.
53. L. Bergaust, R. J. van Spanning, A. Frostegard, L. R. Bakken, *Microbiology* **2012**, *158*, 826–834.
54. J. Hassan, Z. Qu, L. L. Bergaust, L. R. Bakken, *PLoS Comput. Biol.* **2016**, *12*, e1004621.
55. R. Wan, Y. Chen, X. Zheng, Y. Su, M. Li, *Environ. Sci. Technol.* **2016**, *50*, 9915–9922.
56. R. Wan, L. Wang, Y. Chen, X. Zheng, Y. Su, X. Tao, *Sci. Total Environ.* **2018**, *643*, 1074–1083.
57. J. H. Park, H. S. Shin, I. S. Lee, J. H. Bae, *Environ. Technol.* **2002**, *23*, 53–65.
58. R. Yu, M. J. Kampschreur, M. C. van Loosdrecht, K. Chandran, *Environ. Sci. Technol.* **2010**, *44*, 1313–1319.
59. P. Wunderlin, J. Mohn, A. Joss, L. Emmenegger, H. Siegrist, *Water Res.* **2012**, *46*, 1027–1037.
60. R. A. Sanford, D. D. Wagner, Q. Wu, J. C. Chee-Sanford, S. H. Thomas, C. Cruz–Garcia, G. Rodriguez, A. Massol-Deya, K. K. Krishnani, K. M. Ritalahti, S. Nissen, K. T. Konstantinidis, F. E. Loffler, *Proc. Natl. Acad. Sci. USA* **2012**, *109*, 19709–19714.
61. S. Yoon, S. Nissen, D. Park, R. A. Sanford, F. E. Loffler, *Appl. Environ. Microbiol.* **2016**, *82*, 3793–3800.
62. J. R. Onley, S. Ahsan, R. A. Sanford, F. E. Loffler, *Appl. Environ. Microbiol.* **2017**. 10.1128/AEM.01985–17.
63. W. G. Zumft, *Microbiol. Molec. Biol. Rev.* **1997**, *61*, 533–616.
64. M. N. Morgada, D. H. Murgida, A. J. Vila, Chapter 4 of this book.
65. C. Carreira, S. R. Pauleta, I. Moura, *J. Inorg. Biochem.* **2017**, *177*, 423–434.
66. C. M. Jones, A. Welsh, I. N. Throback, P. Dorsch, L. R. Bakken, S. Hallin, *FEMS Microbiol. Ecol.* **2011**, *76*, 541–552.
67. D. Mania, K. Heylen, R. J. van Spanning, A. Frostegard, *Environ. Microbiol.* **2014**, *16*, 3196–3210.
68. D. Mania, K. Heylen, R. J. van Spanning, A. Frostegard, *Environ. Microbiol.* **2016**, *18*, 2937–2950.
69. J. Simon, O. Einsle, P. M. Kroneck, W. G. Zumft, *FEBS Lett.* **2004**, *569*, 7–12.

70. S. Teraguchi, T. C. Hollocher, *J. Biol. Chem.* **1989**, *264*, 1972–1979.
71. C. S. Zhang, T. C. Hollocher, A. F. Kolodziej, W. H. Orme-Johnson, *J. Biol. Chem.* **1991**, *266*, 2199–2202.
72. C. Zhang, A. M. Jones, T. C. Hollocher, *Biochem. Biophys. Res. Commun.* **1992**, *187*, 135–139.
73. C. S. Zhang, T. C. Hollocher, *Biochim. Biophys. Acta* **1993**, *1142*, 253–261.
74. M. Luckmann, D. Mania, M. Kern, L. R. Bakken, A. Frostegard, J. Simon, *Microbiology* **2014**, *160* 1749–1759.
75. C. Carreira, O. Mestre, R. F. Nunes, I. Moura, S. R. Pauleta, *PeerJ* **2018**, *6*, e5603, DOI: 5610.7717/peerj.5603.
76. M. Kern, J. Simon, *Biochim. Biophys. Acta* **2009**, *1787*, 646–656.
77. S. Hein, S. Witt, J. Simon, *Environ. Microbiol.* **2017**, *19*, 4913–4925.
78. P. Wunsch, W. G. Zumft, *J. Bacteriol.* **2005**, *187*, 1992–2001.
79. L. Zhang, C. Trncik, S. L. A. Andrade, O. Einsle, *Biochim. Biophys. Acta* **2017**, *1858*, 95–102.
80. P. Wunsch, H. Korner, F. Neese, R. J. van Spanning, P. M. Kroneck, W. G. Zumft, *FEBS Lett.* **2005**, *579*, 4605–4609.
81. K. Brown, K. Djinovic-Carugo, T. Haltia, I. Cabrito, M. Saraste, J. J. G. Moura, I. Moura, M. Tegoni, C. Cambillau, *J. Biol. Chem.* **2000**, *275*, 41133–41136.
82. K. Brown, M. Tegoni, M. Prudêncio, A. S. Pereira, S. Besson, J. J. G. Moura, I. Moura, C. Cambillau, *Nat. Struct. Biol.* **2000**, *7*, 191–195.
83. K. Paraskevopoulos, S. V. Antonyuk, R. G. Sawers, R. R. Eady, S. S. Hasnain, *J. Mol. Biol.* **2006**, *362*, 55–65.
84. A. Pomowski, W. G. Zumft, P. M. Kroneck, O. Einsle, *Nature* **2011**, *477*, 234–237.
85. S. Dell'acqua, I. Moura, J. J. Moura, S. R. Pauleta, *J. Biol. Inorg. Chem.* **2011**, *16*, 1241–1254.
86. T. Matsubara, H. Iwasaki, *J. Biochem.* **1972**, *71*, 747–750.
87. W. G. Zumft, T. Matsubara, *FEBS Lett.* **1982**, *148*, 107–112.
88. H. Beinert, *Eur. J. Biochem.* **1997**, *245*, 521–532.
89. Suharti, M. J. Strampraad, I. Schroder, S. de Vries, *Biochemistry* **2001**, *40* 2632–2639.
90. C. L. Coyle, W. G. Zumft, P. M. Kroneck, H. Korner, W. Jakob, *Eur. J. Biochem.* **1985**, *153*, 459–467.
91. C. K. SooHoo, T. C. Hollocher, *J. Biol. Chem.* **1991**, *266*, 2203–2209.
92. S. Ferretti, J. G. Grossmann, S. S. Hasnain, R. R. Eady, B. E. Smith, *Eur. J. Biochem.* **1999**, *259*, 651–659.
93. T. Rasmussen, B. C. Berks, J. N. Butt, A. J. Thomson, *Biochem. J.* **2002**, *364*, 807–815.
94. S. Dell'Acqua, S. R. Pauleta, J. J. Moura, I. Moura, *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **2012**, *367*, 1204–1212.
95. B. C. Berks, D. Baratta, D. J. Richardson, S. J. Ferguson, *Eur. J. Biochem.* **1993**, *212*, 467–476.
96. M. Itoh, K. Matsuura, T. Satoh, *FEBS Lett.* **1989**, *251*, 104–108.
97. D. J. Richardson, L. C. Bell, A. G. McEwan, J. B. Jackson, S. J. Ferguson, *Eur. J. Biochem.* **1991**, *199*, 677–683.
98. J. W. B. Moir, S. J. Ferguson, *Microbiology* **1994**, *140*, 389–397.
99. S. Dell'acqua, S. R. Pauleta, E. Monzani, A. S. Pereira, L. Casella, J. J. Moura, I. Moura, *Biochemistry* **2008**, *47*, 10852–10862.
100. K. Fujita, J. M. Chan, J. a. Bollinger, M. L. Alvarez, D. M. Dooley, *J. Inorg. Biochem.* **2007**, *101*, 1836–1844.
101. T. Rasmussen, T. Brittain, B. C. Berks, N. J. Watmough, A. J. Thomson, *Dalton Trans.* **2005**, 3501–3506.

102. F. C. Boogerd, H. W. van Verseveld, A. H. Stouthamer, *FEBS Lett.* **1980**, *113*, 279–284.
103. J. M. Borrero-de Acuna, M. Rohde, J. Wissing, L. Jansch, M. Schobert, G. Molinari, K. N. Timmis, M. Jahn, D. Jahn, *J. Bacteriol.* **2016**, *198*, 1401–1413.
104. J. M. Borrero-de Acuna, K. N. Timmis, M. Jahn, D. Jahn, *Microb. Biotechnol.* **2017**, *10*, 1523–1534.
105. S. Yoshikawa, A. Shimada, K. Shinzawa-Itoh, *Met. Ions Life Sci.* **2015**, *15*, 89–130.
106. T. Tosha, Y. Shiro, *IUBMB Life* **2013**, *65*, 217–226.
107. M. Prudencio, A. S. Pereira, P. Tavares, S. Besson, I. Cabrito, K. Brown, B. Samyn, B. Devreese, J. Van Beeumen, F. Rusnak, G. Fauque, J. J. Moura, M. Tegoni, C. Cambillau, I. Moura, *Biochemistry* **2000**, *39*, 3899–3907.
108. P. Chen, S. DeBeer George, I. Cabrito, W. E. Antholine, J. J. Moura, I. Moura, B. Hedman, K. O. Hodgson, E. I. Solomon, *J. Am. Chem. Soc.* **2002**, *124*, 744–745.
109. M. L. Alvarez, J. Ai, W. Zumft, J. Sanders-Loehr, D. M. Dooley, *J. Am. Chem. Soc.* **2001**, *123*, 576–587.
110. P. Chen, I. Cabrito, J. J. G. Moura, I. Moura, E. I. Solomon, *J. Am. Chem. Soc.* **2002**, *124*, 10497–10507.
111. P. Chen, S. I. Gorelsky, S. Ghosh, E. I. Solomon, *Angew. Chem.* **2004**, *43*, 4132–4140.
112. S. Ghosh, S. I. Gorelsky, George, S. DeBeer, J. M. Chan, I. Cabrito, D. M. Dooley, J. J. G. Moura, I. Moura, E. I. Solomon, *J. Am. Chem. Soc.* **2007**, *129*, 3955–3965.
113. J. A. Farrar, A. J. Thomson, M. R. Cheesman, D. M. Dooley, W. G. Zumft, *FEBS Lett.* **1991**, *294*, 11–15.
114. V. S. Oganesyan, T. Rasmussen, S. Fairhurst, A. J. Thomson, *Dalton Trans.* **2004**, 996–1002.
115. D. M. Dooley, M. A. McGuirl, A. C. Rosenzweig, J. A. Landin, R. A. Scott, W. G. Zumft, F. Devlin, P. J. Stephens, *Inorg. Chem.* **1991**, *30*, 3006–3011.
116. J. A. Farrar, W. G. Zumft, A. J. Thomson, *Proc. Natl. Acad. Sci. USA* **1998**, *95*, 9891–9896.
117. E. I. Solomon, D. E. Heppner, E. M. Johnston, J. W. Ginsbach, J. Cirera, M. Qayyum, M. T. Kieber-Emmons, C. H. Kjaergaard, R. G. Hadt, L. Tian, *Chem. Rev.* **2014**, *114*, 3659–3853.
118. E. M. Johnston, C. Carreira, S. Dell'Acqua, S. G. Dey, S. R. Pauleta, I. Moura, E. I. Solomon, *J. Am. Chem. Soc.* **2017**, *139*, 4462–4476.
119. E. M. Johnston, S. Dell'Acqua, S. R. Pauleta, I. Moura, E. I. Solomon, *Chem. Sci.* **2015**, *6*, 5670–5679.
120. S. G. Mayhew, *Eur. J. Biochem.* **1978**, *85*, 535–547.
121. S. Dell'Acqua, S. R. Pauleta, P. M. P. de Sousa, E. Monzani, L. Casella, J. J. G. Moura, I. Moura, *J. Biol. Inorg. Chem.* **2010**, *15*, 967–976.
122. T. Rasmussen, B. C. Berks, J. Sanders-Loehr, D. M. Dooley, W. G. Zumft, A. J. Thomson, *Biochemistry* **2000**, *39*, 12753–12756.
123. E. M. Johnston, S. Dell'Acqua, S. Ramos, S. R. Pauleta, I. Moura, E. I. Solomon, *J. Am. Chem. Soc.* **2014**, *136* 614–617.
124. S. Ghosh, S. I. Gorelsky, P. Chen, I. Cabrito, J. J. G. Moura, I. Moura, E. I. Solomon, *J. Am. Chem. Soc.* **2003**, *125*, 15708–15709.
125. K. Yamaguchi, A. Kawamura, H. Ogawa, S. Suzuki, *J. Biochem.* **2003**, *134*, 853–858.
126. S. W. Snyder, T. C. Hollocher, *J. Biol. Chem.* **1987**, *262*, 6515–6525.
127. H. Körner, K. Frunzke, K. Döhler, W. G. Zumft, *Arch. Microbiol.* **1987**, *148*, 20–24.
128. J. M. Chan, J. Bollinger, C. L. Grewell, D. M. Dooley, *J. Am. Chem. Soc.* **2004**, *126*, 3030–3031.
129. M. Z. Ertem, C. J. Cramer, F. Himo, P. E. Siegbahn, *J. Biol. Inorg. Chem.* **2012**, *17*, 687–698.
130. L. K. Schneider, O. Einsle, *Biochemistry* **2016**, *55* 1433–1440.

131. S. R. Pauleta, I. Moura, in *Encyclopedia of Inorganic and Bioinorganic Chemistry*, Eds R. A. Scott, John Wiley & Sons, Ltd, New Jersey, **2017**.
132. J. Riester, W. G. Zumft, P. M. Kroneck, *Eur. J. Biochem.* **1989**, *178*, 751–762.
133. P. Wunsch, M. Herb, H. Wieland, U. M. Schiek, W. G. Zumft, *J. Bacteriol.* **2003**, *185*, 887–896.
134. W. G. Zumft, A. Viebrock-Sambale, C. Braun, *Eur. J. Biochem.* **1990**, *192*, 591–599.
135. W. G. Zumft, *J. Mol. Microbiol. Biotechnol.* **2005**, *10*, 154–166.
136. U. Honisch, W. G. Zumft, *J. Bacteriol.* **2003**, *185*, 1895–1902.
137. L. M. Taubner, M. A. McGuirl, D. M. Dooley, V. Copie, *Biochemistry* **2006**, *45*, 12240–12252.
138. M. A. McGuirl, J. A. Bollinger, N. Cosper, R. A. Scott, D. M. Dooley, *J. Biol. Inorg. Chem.* **2001**, *6*, 189–195.
139. L. M. Taubner, M. A. McGuirl, D. M. Dooley, V. Copie, *J. Biomol. NMR* **2004**, *29*, 211–212.
140. S. P. Bennett, M. J. Soriano-Laguna, Justin M. Bradley, D. A. Svistunenko, D. J. Richardson, A. J. Gates, N. E. Le Brun, *Chemical Science* **2019**, *10*, 4985–4993.
141. A. Dreusch, D. M. Burgisser, C. W. Heizmann, W. G. Zumft, *Biochim. Biophys. Acta* **1997**, *1319*, 311–318.
142. M. Giles, N. Morley, E. M. Baggs, T. J. Daniell, *Front. Microbiol.* **2012**, *3*, 407.
143. J. Simon, M. G. Klotz, *Biochim. Biophys. Acta* **2013**, *1827*, 114–135.

6

Cytochrome P_{450} · The Dioxygen-Activating Heme Thiolate

F. Miguel Castro Martínez, [1] ***R. Daniel Páez López,*** [2] ***Pedro D. Sarmiento Pavía,*** [2] ***Martha E. Sosa Torres,*** [2] ***and Peter M. H. Kroneck*** [3]

[1]Departamento de Física y Química Teórica, Facultad de Química, Universidad Nacional Autónoma de México, Ciudad Universitaria, 04510, CDMX, México
<miguel.castro.m@gmail.com>

[2]Departamento de Química Inorgánica y Nuclear, Facultad de Química, Universidad Nacional Autónoma de México, Ciudad Universitaria, 04510, CDMX, México
<richie_daniels@comunidad.unam.mx>
<pedrodavids@comunidad.unam.mx>
<mest@unam.mx>

[3]Department of Biology, University of Konstanz, D-78457 Konstanz, Germany
<peter.kroneck@uni-konstanz.de>

Metal Ions in Life Sciences, Volume 20 Guest Editors: Martha E. Sosa Torres and Peter M. H. Kroneck
Series Editors: Astrid Sigel, Eva Freisinger, and Roland K. O. Sigel

https://doi.org/10.1515/9783110589757-006

Abstract: Cytochromes P_{450} (CYPs) are heme *b*-binding enzymes and belong to Nature's most versatile catalysts. They participate in countless essential life processes, and exist in all domains of life, Bacteria, Archaea, and Eukarya, and in viruses. CYPs attract the interest of researchers active in fields as diverse as biochemistry, chemistry, biophysics, molecular biology, pharmacology, and toxicology. CYPs fight chemicals such as drugs, poisonous compounds in plants, carcinogens formed during cooking, and environmental pollutants. They represent the first line of defense to detoxify and solubilize poisonous substances by modifying them with dioxygen. The heme iron is proximally coordinated by a thiolate residue, and this ligation state represents the active form of the enzyme. The Fe(III) center displays characteristic UV/Vis and EPR spectra (Soret maximum at 418 nm; g-values at 2.41, 2.26, 1.91). The Fe(II) state binds the inhibitor carbon monoxide (CO) to produce a Fe(II)–CO complex, with the major absorption maximum at 450 nm, hence, its name P450. CYPs are flexible proteins in order to allow a vast range of substrates to enter and products to leave. Two extreme forms exist: substrate-bound (closed) and substrate-free (open). CYPs share a sophisticated catalytic cycle that involves a series of consecutive transformations of the heme thiolate active site, with the strong oxidants compound I and II as key intermediates. Each of these high-valent Fe(IV) species has its characteristic features and chemical properties, crucial for the activation of dioxygen and cleavage of strong C–H bonds.

Keywords: biodiversity · C–H bond cleavage · compound I · compound II · cytochrome P_{450} · dioxygen activation · directed evolution · drug metabolism · heme *b* · monooxygenase · thiolate

1. INTRODUCTION

The cytochromes P_{450} (CYPs, P450s) belong to Nature's most versatile catalysts in view of both their vast substrate range, and the diverse types of molecular transformations performed across the CYP superfamily. Actually, their substrate range exceeds that of other enzymes found in Nature. They represent a superfamily of monooxygenase enzymes with enormous potential for industrial applications. The CYPs perform highly specific oxidative chemistry, utilizing a sophisticated catalytic reaction mechanism, with the transient reaction cycle intermediate compound I (Cpd I) responsible for the majority of P450 oxidative reactions [1–17]. The catalytic versatility of P450s makes these enzymes particularly valuable for applications in synthetic biology [18–20]. This major advance comes at a time when P450 engineering has facilitated the elucidation of several mammalian P450 structures and generated P450 variants with novel substrate specificity and reactivity [21–27].

In this chapter we intend to introduce the reader to several basic structural and functional features of the thiolate-ligated heme enzyme cytochrome P_{450}, nominated "Molecule of the Month" in October 2006 (D. Goodsell, PDB database, doi:10.2210/rcsb_pdb/mom_2006_10). It participates in countless essential life processes, and has attracted the interest of numerous scientists active in fields as diverse as biochemistry, chemistry, biophysics, molecular biology, pharmacology, and toxicology. Sequencing has produced so far over 300,000 CYP sequences; nomenclature has been assigned to more than 41,000 CYP sequences and the

majority of the remainder has been sorted by BLAST (Basic Local Alignment Search Tool) searches into clans, families, and subfamilies [28]. CYP enzymes play a prominent role in phase I metabolism of approximately three quarters of drug metabolism reactions in humans, with only five isoforms accounting for the majority of these reactions [29–31]. The versatility of CYPs includes the ability to metabolize innumerable substrates of both physiological and xenobiotic importance, and to be markedly altered in activity by a wide variety of inducers and inhibitors. However, they can also cause harm. On the one hand, CYPs, such as CYP3A4, detoxify a broad spectrum of diverse drugs, but on the other hand, CYP3A4 can also modify the antibiotic erythromycin, and it is supposed to be partially responsible for the toxicity of large doses of acetaminophen (Tylenol) [30–33].

These remarkable enzymes can conduct a wide range of reactions, however, the most common is the oxidation of the substrate R-H, whereby one oxygen atom is inserted into the substrate, while the other is reduced to water, first formulated by Howard Mason [34, 35] by using $^{18}O_2$ (Equation 1).

$$\text{R-H} + 2\text{H}^+ + {}^{18}\text{O}_2 + 2\text{e}^- \rightarrow \text{R-}{}^{18}\text{OH} + \text{H}_2{}^{18}\text{O} \quad (1)$$

The hydrogen donor in many of these reactions is reduced nicotinamide adenine dinucleotide phosphate (NADPH) or reduced nicotinamide adenine dinucleotide (NADH), via electron transfer from enzymes such as the NADPH-cytochrome P_{450} oxidoreductase in the endoplasmic reticulum, or from ferredoxin plus ferredoxin reductase in mitochondria. Note that there exists a fungal cytochrome P_{450} nitric oxide reductase ($P450_{Nor}$). This enzyme is part of the denitrification pathway and receive electrons directly from NADH to reduce nitric oxide (NO) to nitrous oxide (N_2O) [36].

The cytochrome P_{450} enzymes were originally discovered in the 1940s, during the exploration of a class of unusual reactions involving dioxygen. These "mixed-function" oxidation reactions (Equation 1) proved to be significant in the metabolism of chemicals as diverse in structure as drugs, steroids, and pesticides [37]. It was Tsuneo Omura and Ryo Sato in 1962 who discovered a novel cytochrome in liver microsomes, with the initial observation of a pigment which bound carbon monoxide (CO). It showed a characteristic absorption band centered at 450 nm, which led the authors to coin the term cytochrome P_{450} [38]. A further decisive step forward was achieved by Minor Coon and coworkers who succeeded to solubilize the microsomal enzyme system into fractions containing cytochrome P_{450}, a cytochrome P_{450} reductase, and a heat-stable factor [39–41]. Today, thanks to genome sequencing we know that CYP enzymes are evolutionarily conserved and exist in all domains of life, Bacteria, Archaea, and Eukarya, and even in viruses [42]. However, they have not been found in *Escherichia coli*. In view of its ability to metabolize dioxygen and to bind carbon monoxide, there might have been an early role during chemical evolution for cytochrome P_{450} [43, 44]. Nearly all eukaryotes require sterols in their membranes for essential fluidity and lipid-packing properties. Sterol 14α-demethylase (CYP51) synthesizes these sterols and, therefore, is fundamental to eukaryotic life. Among human's

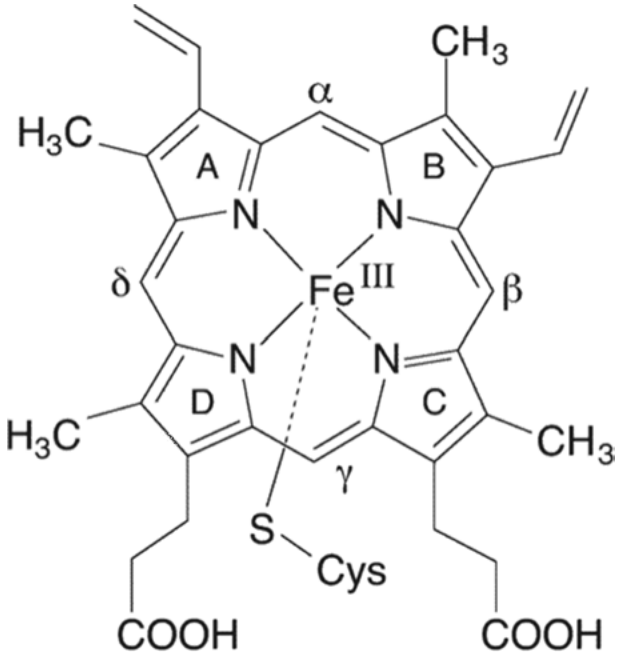


Figure 1. Structure of heme *b* (Fe protoporphyrin IX) with proximal cysteinate ligand. Adapted with permission from [8] (B. Meunier, S. P. de Visser, S. Shaik, *Chem. Rev.* **2004**, *104*, 3947–3980); copyright (2004) American Chemical Society.

functional CYP genes, DNA variants can cause a number of inborn errors of metabolism, other clinical diseases, and important differences in drug response.

Perhaps *the* defining feature of cytochrome P_{450}, from a bioinorganic point of view, is the combination of heme iron and thiolate sulfur (Cys-S^-) at the active site (Figure 1). The UV/Vis and EPR spectroscopic features of cytochrome P_{450} led researchers to describe it as perhaps the most curious amongst the cytochromes. Early on most ideas involved a thiolate-coordinated heme prosthetic group in order to explain the UV/Vis (maxima at 416, 535, 570, 650 nm) and EPR (g-values at 2.41, 2.26, 1.91) properties of cytochrome P_{450} [45–47]. Today it is well known that only thiolate-ligated heme proteins can activate inert C–H bonds, suggesting a critical role for the ligand in this process. Strong electron donation from the axial thiolate has been proposed to facilitate both oxygen and C–H bond activation [48–55].

In view of the vast literature accumulated and still accumulating in the field, we suggest for introduction the article by the world renowned biochemist and P450 pioneer Ronald Estabrook entitled "A Passion For P450s (Remembrances of the Early History of Research on Cytochrome P450)" published in 2003 [56]. To get a deeper understanding and insight into the complex topic, these excellent monographs are recommended: (i) *The Ubiquitous Roles of Cytochrome P450 Proteins* [1], (ii) *Cytochrome P450: Structure, Mechanism, and Biochemistry* [2],

and (iii) *Monooxygenase, Peroxidase and Peroxygenase. Properties and Mechanisms of Cytochrome P450* [3]. Broader perspectives and more inclusive of the work of experts in this dynamic research area are available in a number of very informative and authoritative reviews [4–17]. We apologize to all the researchers and their associates who have contributed to our current knowledge of the P450s but are not mentioned in this chapter. The discussion will be narrowly focused to some extent on the coordination chemistry and properties of the protoporphyrin IX heme prosthetic group ligated to an axial sulfur ligand in cytochrome P_{450} and its role in the activation of dioxygen, with not as much emphasis on the tremendous functional versatility of CYPs and their remarkable ability to metabolize innumerable substrates. Two structurally and functionally related enzymes, chloroperoxidase (CPO) and NO synthase (NOS), will be briefly mentioned when discussing the reaction cycle of CYP. Both have an Fe-protoporphyrin IX center coordinated to Cys-S^-, and they either activate dioxygen (O_2) in the case of NOS, or hydrogen peroxide (H_2O_2) in the case of CPO [7, 10]. These omissions are not intentional but merely the consequence of time and space. To stay updated with the rapidly advancing field we strongly recommend the website of the Centre for Cytochrome P450 Biodiversity at the Swansea University Medical School (www.p450swansea.co.uk/home) and the CYP homepage maintained by David Nelson (http://drnelson.uthsc.edu/CytochromeP450.html) [28]. Yet hopefully this chapter will illustrate what one may learn about studying the structure and function of such an important and intensively studied enzyme.

2. THE CYTOCHROME P_{450} SUPERFAMILY: NATURE'S MOST VERSATILE CATALYSTS

The range and number of P450s throughout the biosphere is striking and as such cytochrome P_{450} represents one of the largest enzyme families [14, 15]. To keep track of the steadily increasing mass of information on P450s, the CYP nomenclature was developed. The protein is designated as CYPxyz, CYP stands for cytochrome P_{450}, xyz is a code relating the CYP to a family, subfamily, and isoform [57–60]. Sequencing in all areas of the tree of life yielded over 300,000 cytochrome P_{450} sequences that have been mined and collected. The gene analysis of different CYPs indicates that these metalloproteins exhibit the same so-called "P450 signature" motif of 10 amino acids, including the invariant cysteine residue that ligates the heme iron to the protein, Phe-XX-Gly-X_bXX-Cys-X-Gly. The amino acid X_b is usually a basic amino acid that plays a key role in interactions with the reductase partner. Nearly 800 fungal species have been now sequenced, and over 3,000 bacterial sequences have been named, mostly from terrestrial or freshwater sources. Interestingly, 25 % of the seawater bacterial P450s did not match known P450 families, indicating marine bacterial CYPs are not as well sampled as land/freshwater based bacterial CYPs. In some cases evolution resulted in gene loss leading to no P450s. Future sequencing plans are expected to produce more than one million cytochrome P_{450} sequences by 2020 [28].

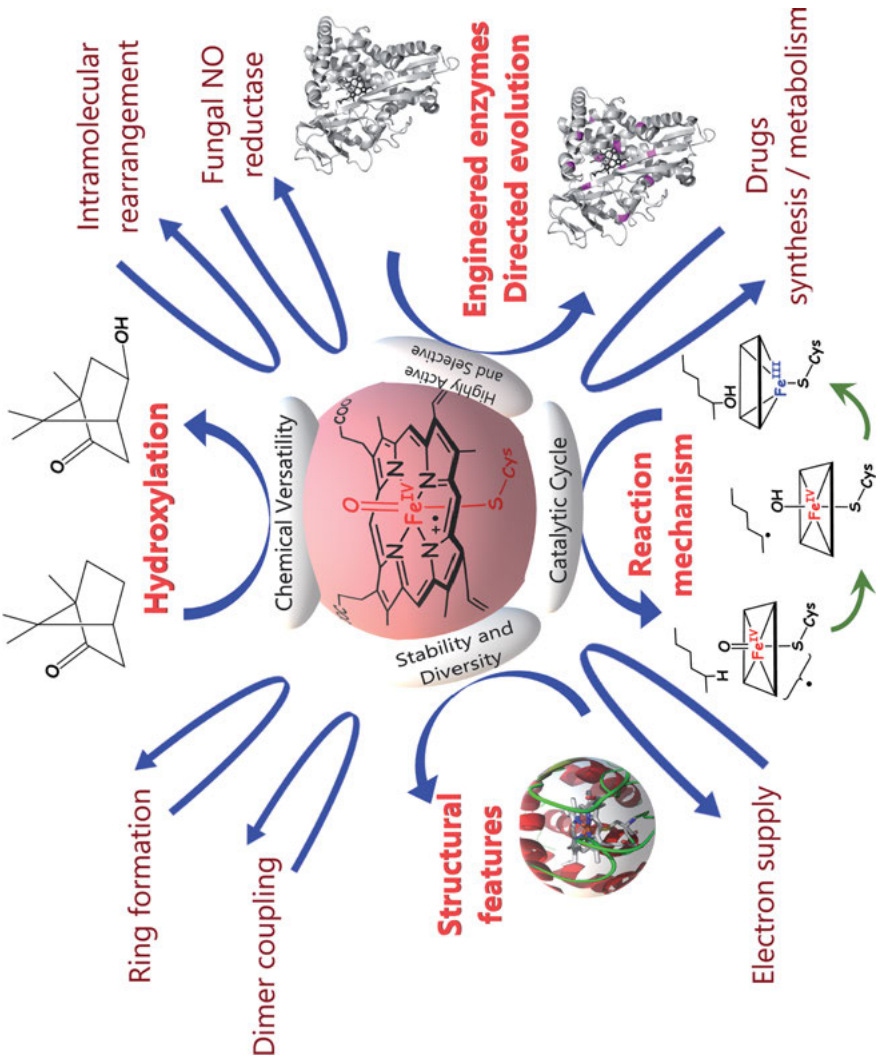


Figure 2. P450, the versatile catalyst with a vast substrate range and diverse types of molecular transformations.

The P450 enzymes, which are divided into gene families and subfamilies on the basis of the amount of identity between amino acid sequences [28], can be broadly classified in two main classes depending on their global substrate specificity. The first class includes P450 enzymes that have a strong preference for a single substrate, and originate mostly from either plants or bacteria, and also include P450s of the steroidogenesis and eicosanoid pathways in animals. The second class comprises P450 enzymes featuring a broad specificity for substrates, often very different in size and chemical nature, mostly comprising the P450s that are involved in the drug and pollutant metabolism in vertebrates. Consequently, P450 enzymes are frequently designated as unspecific monooxygenases (EC1.14.14.1) [58–60].

CYPs perform a rich and unique chemistry involving dioxygen, they react with hydrocarbons (hydroxylation, epoxidation, desaturation, allylic rearrangement), alcohols (oxidation), aldehydes (oxidation, deformylation), ethers (O-dealkylation), thioethers (sulfoxidation), amines (N-dealkylation, N-oxygenation), or halocarbons (reductive and oxidative dehalogenation) [61]. Notably, the proteins can undergo remarkable structural changes resulting in modifications of substrate specificity [62]. Many times these changes can be directly attributed to small differences in the amino acid sequence of a distinct cytochrome P_{450}. The classic stoichiometry of monooxygenases represents a sort of 'half-way' point on the pathway for the reduction of O_2 to produce two molecules of H_2O observed in respiratory energy conservation by cytochrome *c* oxidase [63]. Monooxygenases, also often referred to as mixed-function oxidases, require only two electrons and two protons to reductively cleave atmospheric dioxygen, producing only a single water molecule in the process while saving the second atom for the incorporation and formal oxidation of the organic substrate molecule (Equation 1) [34, 35, 61].

2.1. Redox Partners

CYPs belong to the group of external monooxygenases and thus receive the necessary electrons for cleavage of dioxygen and substrate hydroxylation from different redox partners. Even though the sequence conservation among P450 proteins of different families may be less than 20 %, their general topography and structural fold are highly conserved and point to a common mechanism of oxygen activation. In contrast to that, the most variable regions are represented by the flexible substrate recognition regions, which enable CYPs to be versatile biocatalysts with such a remarkable variety of chemical reactions catalyzed. The vast majority of P450s depends on one or more redox partners, however, some P450s do not require any other protein component to achieve the reductive activation of molecular oxygen. The fatty acid monooxygenase, cytochrome P450BM-3 from *Bacillus megaterium,* uses a FAD/FMN reductase for electron delivery, however, the P450 moiety and the FAD/FMN reductase are fused together in a single polypeptide chain of 119,000 daltons, whereas in other CYPs the reductase and P450 are separate polypeptides. The P450BM-3 has become an excellent model system for the investigation of microsomal CYPs with the distinct advantage that large amounts of the enzyme can be prepared from *E. coli* recombinant expression systems [64].

Although there exist a large number of P450s, they all share a common catalytic cycle as discussed in Section 4. Many P450s are promiscuous with respect to the redox partner, and very often catalysis can be supported by non-native redox partners. CYPs have been grouped into several classes depending on the topology of the protein components involved in the electron transfer to the P450 enzyme. Typically, the protein components contained the following redox centers: ferredoxin (Fdx; Fe-S cluster), ferredoxin reductase (FdR, FAD), cytochrome P450 reductase (CPR; FAD, FMN), flavodoxin (Fldx, FMN), 2-oxoacid:ferredoxin oxidoreductase (OFOR; thiamine pyrophosphate, Fe-S cluster), and phthalate-family oxygenase reductase (PFOR; FMN, Fe-S cluster). There also exist P450-only systems, which do not require external equivalents, such as thromboxane synthase (CYP5), prostacyclin synthase (CYP8), and allene oxide synthase (CYP74A). Obviously, the identification of new CYPS by genome sequencing projects will lead to the discovery of more and more P450 enzymes with novel catalytic properties. However, only a relatively small number of the many P450 enzymes have been characterized so far with regard to their three-dimensional structure and catalytic spectrum, thus, in some instances, the unusual P450 properties described may actually be rather common within the superfamily [65–68].

From a structural point of view, less information has been accumulated about the P450-redox partner interaction and inter-protein electron transfer reactions. One major problem is the difficulty in obtaining detailed structural data on electron transfer complexes, as they tend to be weak and possibly dynamic. For rapid turnover, the redox partner must bind, deliver an electron, and then dissociate. Such functional requirements preclude formation of stable long-lived complexes. The first crystal structure of a P450-redox partner complex between the heme- and FMN-binding domains of cytochrome P450BM-3 from *B. megaterium*, has been determined by the P450 pioneer Thomas Poulos and his associates in 1999 [69]. Being a soluble multidomain electron transfer protein, this enzyme represents an excellent model system for studying structure–function relationships in CYPs and the mechanism of electron transfer, thus can be regarded a prototype for the complex between eukaryotic microsomal P450s and P450 reductase. Further information about redox partner dynamics and electron transfer pathways will be given in Section 3.

2.2. Engineering and Directed Evolution

Nature, the best chemist on Earth, solves the difficult problem of being alive and enduring for billions of years, under an astonishing range of environmental conditions. Most of the complex chemistry that makes life possible is the work of Nature's macromolecular protein catalysts, the enzymes [70–73]. One prime example is cytochrome P_{450}, a catalyst with a remarkable substrate range and an enormous potential for industrial applications (Figure 2). It is engaged in highly specific oxidative chemistry, clearly their catalytic versatility makes these enzymes particularly attractive for applications in bioengineering [18–20, 24–27, 71–79]. Notably most P450-catalyzed reactions are relatively slow, with $k_{cat} < 100\ min^{-1}$. One remarkable exception is the ω-2 hydroxylation of myristic acid catalyzed by

CYP102A1 with a $k_{cat} > 4000\ min^{-1}$. In general, the application of CYPs for the production of bulk chemicals is limited in view of their low activity and stability and their cofactor dependency, thus, the practical application of CYPs may be limited to the production of fine chemicals. Perhaps the most successful application of CYPs on an industrial scale is the bioconversion process for pravastatin production, a potent LDL-cholesterol-lowering drug [25].

Synthetic chemists have produced a tremendous amount of compounds, which feed, clothe, house, entertain, and cure us. However, they struggle to match the efficiency and selectivity that biological systems, such as microorganisms achieve with enzymes: "Lord, I fall upon my knees and pray that all my syntheses may not be inferior to those conducted by bacteria" (J. B. S. Haldane) [70]. In many cases, synthetic processes rely on precious metals, toxic reagents and solvents, and require extreme conditions from an energetic point of view. Last but not least, they produce unwanted byproducts. Enzymes, on the other hand, such as cytochrome P_{450}, document how a protein can orient substrates for reaction, exclude a ligand from an active site, activate a metal or an organic cofactor, or suppress competing reactions to draw out new and admirable synthetic capabilities.

According to Frances Arnold, winner of the 2018 Nobel Prize in Chemistry awarded for the directed evolution of enzymes, synthetic chemists have been drawing inspiration from Nature for decades, and now is the time for protein engineers to use inspiration from synthetic chemistry to generate new enzymes that will improve on and replace synthetic catalysts and reaction pathways, and create enzymes for reactions that neither biology nor synthetic chemistry has conquered [20, 71–77]. To demonstrate the power of directed evolution, Arnold and coworkers directed the evolution of a cytochrome P_{450} anti-Markovnikov oxygenase. It catalyzes the conversion of alkenes into the anti-Markovnikov carbonyl compounds. Using iron, dioxygen, and a recyclable cofactor (NADPH), the laboratory-evolved cytochrome P_{450} enzyme catalyzes the anti-Markovnikov oxidation of different substituted styrenes, including hindered substrates such as internal and 1,1-disubstituted alkenes [18].

Another useful tool, bioconjugation, has been applied to CYPs, i.e., the chemical modification of the protein to investigate its functional and structural properties in greater detail. These bioconjugates can be used to address specific questions and the chemical strategies employed are diverse and depend on the problem under investigation, such as drug metabolism, catalysis, or biosensing [78]. In order to develop alternative methods to promote P450 catalysis, arguably, the most limiting factor to overcome is the timely delivery of electrons and the minimization of uncoupled reactions [79]. Uncoupling resulting from inefficient electron transfer is problematic not only due to the waste of reducing equivalents, but additionally for its generation of reactive oxygen species, which are detrimental to the enzyme. One promising strategy takes advantage of light energy to drive P450 reactions as recently reviewed by Cheruzel and coworkers [26]. This approach utilizes a light-harvesting unit with the appropriate excited state properties and reduction potential to initiate the P450 catalytic mechanism. Two separate paths have been pursued: the first relies on the *in situ* reduction of dioxygen and subsequent peroxide shunt pathway (discussed in Section 4). The second mimics the natural pathway through the stepwise delivery of electrons

and activation of dioxygen at the heme active site. While the first approach is limited to applications in hydrogen peroxide-tolerant P450 members or evolved strains, widespread use of the second requires the rapid injection of electrons to catalytically competent heme domains [26].

3. STRUCTURAL FEATURES OF CYTOCHROMES P_{450}

As of February 2019 the Protein Data Bank (PDB) lists 942 entries with the name P450 in the structure title (Homo sapiens, 186 structures). CYPs were first discovered in liver microsomes [37–41]. Detailed structure–function studies were limited since they are anchored to the membrane via an N-terminal hydrophobic tail which presented problems with purification and recombinant expression. As a consequence, most of the early mechanistic and biophysical investigations centered on soluble enzymes and $P450_{cam}$ (CYP101A1) from *Pseudomonas putida* was the first to be structurally characterized in the Fe(III) camphor-bound state at 2.6 Å resolution [80]. The structure has been refined to a resolution of 1.63 Å, which revealed three *cis*-proline residues and allowed a more detailed picture of substrate–protein interactions (PDB ID 2CPP) [81]. This enzyme, originally described as camphor 5-*exo*-monooxygenase (EC 1.14.15.1), catalyzes the regio- and stereo-selective hydroxylation of the *exo* C5-H bond of camphor to 5-*exo*-hydroxycamphor which is the first step in the utilization of camphor as its sole source of carbon and energy by the common soil microorganism *P. putida*. This terpene is first hydroxylated and then metabolized to isobutyrate and acetate (Figure 3) [81].

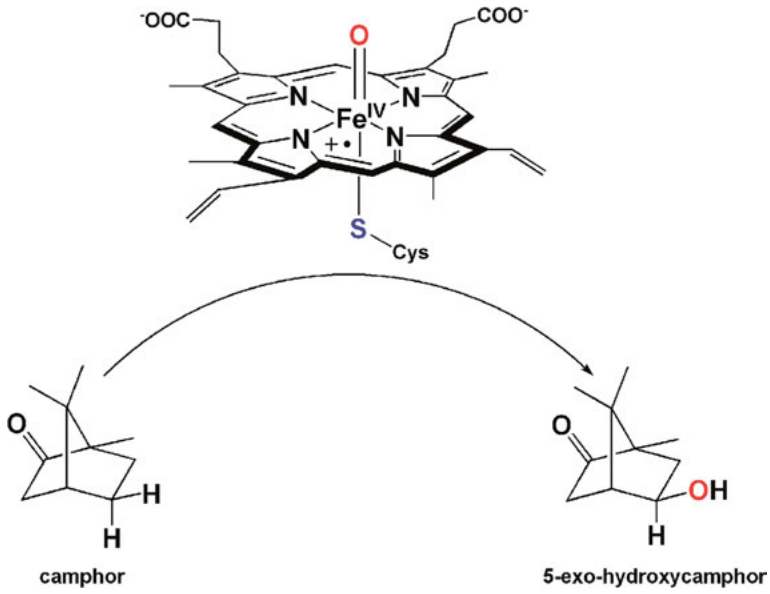


Figure 3. Schematic representation of the regio- and stereo-selective hydroxylation of camphor by $P450_{cam}$ compound I.

Early on, in view of the homology in active site properties, the X-ray structure of $P450_{cam}$ was expected to provide a robust molecular basis for understanding the structure and function of eukaryotic CYPs. Currently there are 119 $P450_{cam}$ structures deposited at the Protein Data Bank, amongst them 114 structures for the enzyme of *P. putida*. Two recent publications focus on the effect of redox partner binding on P450 conformational dynamics [82] and the identification of productive and futile encounters in an electron transfer protein complex [83, 84].

The results of structural investigations on the microsomal, membrane-associated, CYP3A4 bound to metyrapone and progesterone (PDB ID 1W0G, 1W0F) [85] and to ketoconazole and erythromycin (PDB ID 2J0C, 2J0D) [86] have been of major pharmacological interest as this enzyme plays the major role in drug detoxification in humans.

In view of the immense literature on structural studies on both membrane-associated and soluble CYPs and their interaction with substrates, inhibitors, and redox partners, we will mainly focus on results obtained for the soluble protein of *P. putida*, $P450_{cam}$ (CYP101). For further information and deeper insights the following reviews are recommended as prime references [1–3, 9, 15, 21, 60, 68, 87–94].

3.1. Overall Protein Structure and Flexibility

CYPs are single polypeptides (40–55 kDa), and the overall fold in all known structures is basically identical as illustrated for $P450_{cam}$, which was not too surprising, based on sequence alignments. The conserved P450 structural core is formed by a four-helix bundle composed of three parallel helices (D, L, I) and one antiparallel helix E (Figure 4) [89–92]. The ubiquity of the P450 fold suggests that it has been co-opted by evolution many times, and likely presents a useful compromise between structural stability and conformational flexibility. Recent progress in metagenomics revealed much about organisms capable of living under extreme conditions, such as high temperature (thermophiles) and low temperature (psychrophiles), and P450s have been isolated from organisms living at each extreme. Both types have the potential to inform about the structural adaptations to CYPs necessitated by changes in ambient temperature. Numerous CYPs were identified in thermophiles and detailed analysis of their three-dimensional structures provided valuable information to stabilize the P450 fold; the first crystal structure was reported for the enzyme from the acidothermophilic microorganism *Sulfolobus acidocaldarius* (CYP119A1; PDB ID 1IO7). Protein engineering, particularly by directed or artificial evolution, led to mutations that served to stabilize particular mesophilic enzymes of interest for industrial applications [94, 95].

A significant difference between membrane-bound and soluble proteins is a nonpolar N-terminal extension that anchors particulate CYPs to the membrane but this appears to have little effect on the core structure. Common features of membrane P450 structures have been recently reviewed [91, 92]. The catalytic domain (~ 460 amino acids) folds into a triangular prism shape similar to that of soluble prokaryotic CYPs, with 12 α-helices first identified for the structure of

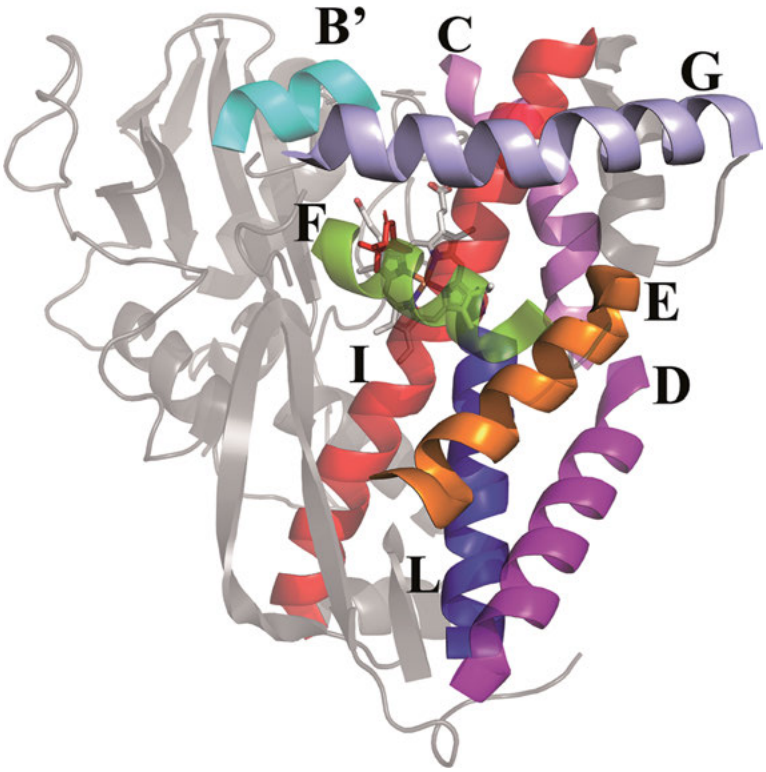


Figure 4. Crystal structure of $P450_{cam}$ (PDB ID 2CPP). Overall structure viewed looking along the heme normal with the substrate binding site facing the viewer. The helix I (red) runs directly over the heme and provides part of the O_2 binding pocket. Adapted with permission from [15] (T. L. Poulos, *Chem. Rev.* **2014**, *114*, 3919−3962); copyright (2014) American Chemical Society.

soluble $P450_{cam}$ [80, 81]. Additionally, there is a highly conserved β-sheet domain near the N-terminus of the protein. The number of helices is typically larger, but these helices are less conserved. Spatial conservation is highest for the structural core of the protein and diverges most for the substrate-binding site [91, 92].

A major challenge is to understand how P450s adapt to accommodate that many different substrates. The first crystal structure ($P450_{cam}$; PDB ID 2CPP) presented a puzzle as to how the substrate enters the active site; clearly, the substrate-binding pocket of the soluble enzyme of *P. putida* was inaccessible to the outside world [80, 81]. The prevailing view developed since then is that the F and G helices and the loop connecting these helices are flexible and have to undergo an open/close motion, which allows substrates to enter and products to leave (Figure 4). Progress came from the 1.6 Å crystal structure of mammalian enzyme P450 2B4 (PDB ID 1PO5), which frequently served as an experimental model for membrane-associated CYPs. It revealed a large open cleft that extended from the protein surface directly to the heme iron center between the α-

helical and β-sheet domains without perturbing the overall P450 fold. The enzyme had been trapped in an open conformation, with the cleft primarily formed by helices B′ to C and F to G, in contrast to other structurally defined mammalian P450s, in which the F to G and B′ to C regions encapsulated one side of the active site to produce a closed form of the enzyme. These structural differences suggest that defined regions of xenobiotic metabolizing CYPs can adopt energetically accessible conformations without perturbing the overall fold, to facilitate substrate access, metabolic versatility, and product egress [96]. Along these lines, crystal structures of $P450_{cam}$ complexed with two different molecular wires were reported, with both wires bound in similar positions at the camphor-binding site. However, each wire induced a distinct conformational response in the protein that differed from the camphor-bound structure. The results of these structural studies suggested that the coupling of substrate-induced conformational changes to active-site residues may be different in $P450_{cam}$ compared to those observed for mammalian CYPs [97, 98]. More recently, $P450_{cam}$ was crystallized in the absence of substrate and at high and low concentration of potassium ion (K^+). The data clearly illustrate that $P450_{cam}$, contrary to the general belief that the active site of $P450_{cam}$ is more static and shielded from solvent than other CYPs, can dynamically visit a closed (PDB ID 3L63) and an open (PDB ID 3L62) conformation that allows access to the deeply buried active site without being induced by substrate or ligand [99]. Double electron–electron resonance experiments with spin-labelled *P. putida* $P450_{cam}$, performed in solution, demonstrated conclusively that, in the physiologically relevant solution state, the substrate-free enzyme exists in the open conformation, binding of the substrate camphor caused the conversion to the closed form [100].

3.2. Active Site Cavity and the Axial Cysteine Ligand

Although the overall protein fold of CYPs is maintained, the precise positioning of various structural elements can differ substantially. The closer one moves towards the catalytic heme center, the more conserved is the structure, especially helices I and L, which directly contact the heme. On the other hand, those regions controlling substrate specificity differ the most, especially the B′ helix. For example, in CYP107 ($P450_{eryF}$), the B′ helix is oriented about 90° from the orientation observed in CYP101 ($P450_{cam}$). The effect is a substantial change in local environment that is required for substrate selectivity [89]. The prosthetic heme group is held by the pincer action of two helices. It is confined between the distal helix I and the proximal helix L and bound to the adjacent cysteine heme-ligand loop containing the P450 signature amino acid sequence Phe-XX-Gly-X_bXX-Cys-X-Gly. The very long helix I runs over the surface of the heme and contributes groups that interact with both the substrate and O_2. There are some variations on the exact positioning of helix I and local helical distortions near the O_2 binding site in various CYPs. Helix I forms a wall of the heme pocket and contains the signature amino acid sequence (Ala/Gly)GlyX(Glu/Asp)Thr which is centered at a kink in the middle of the helix. It provides a threonine residue believed to be involved in catalysis. This threonine preceded by an acidic residue

is positioned in the active site and is required for efficient activation of O_2 to the reactive high-valent iron species (Figure 4) [15]. Thr252 (numbering $P450_{cam}$) is involved in a local helical distortion such that the threonine side chain OH donates an H-bond to a peptide carbonyl oxygen that would normally be involved in an α-helical H-bond. However, this threonine is not strictly conserved, for example, in $P450_{eryF}$ it is replaced by alanine instead. Even so, $P450_{eryF}$ also exhibits a similar distortion in helix I. Furthermore, a water molecule takes the place of the threonine side chain OH thus maintaining a very similar H-bonding pattern. This arrangement is thought to be quite important in the proper delivery of protons to the iron-linked oxygen required for cleavage of the O–O bond thus generating the active Fe-O hydroxylating species. In $P450_{cam}$ the proximal heme face lies closest to the surface, at a distance of about 8 Å, although no part of the heme is directly exposed to bulk solvent. The heme propionate groups are surrounded by hydrogen-bond donors (Arg112, Arg299, His355, Gln108, Asp297 and a water molecule); in addition, one of the heme propionate groups forms an unusual interaction with the side-chain of Asp297 [81].

As the $P450_{cam}$ crystal structure suggested that Thr252 in helix I constitutes an important part of the dioxygen activation machinery, it was an early target for mutagenesis. Replacement of Thr252 with an aliphatic residue led to a loss in activity although the rate of NADH utilization did not decrease significantly. However, exchange of the second conserved residue in helix I, Asp251 by Asn caused a large drop in NADH consumption rate. Detailed kinetic isotope effect and proton inventory studies with the Asp251Asn variant showed that the Asp251 exchange decelerated the electron transfer step and associated proton transfer events. These investigations clearly implicate Thr252 and Asp251 as critical residues in the dioxygen activation process and especially in the proper delivery of protons to the iron-linked dioxygen, and were later confirmed by crystallographic studies [101–103].

The most conserved structural elements in CYPs center on the heme thiolate dioxygen activation chemistry, with the most noteworthy feature, the β-bulge segment housing the proximal thiolate ligand, just prior to the L helix (Figure 5). Such a rigid architecture is required to both protect the sulfur ligand and hold it in place in order to accept hydrogen bonds from peptide NH groups [48]. This arrangement is not only found in all CYPs but also in the closely related enzymes NOS and CPO, which are both heme thiolate enzymes and catalyze monooxygenase reactions. The proximal thiolate ligand is located at the C-terminal end of helix L where it accepts an H-bond from a peptide NH group. This type of H-bonding is very common in iron-thiolate proteins and is required for modulating the iron redox potential. Without such H-bonds the redox potential would be too low for physiological reductants. Note that the arrangement is not unique to heme thiolate proteins. It is a characteristic structural feature of proteins containing Cys-S^--Fe ligation and was first observed in the iron-sulfur protein ferredoxin [104]. Apparently, the protein must provide a suitable electrostatic environment around the sulfur ligand in order to maintain the redox potential in a physiologically accessible range [105, 106].

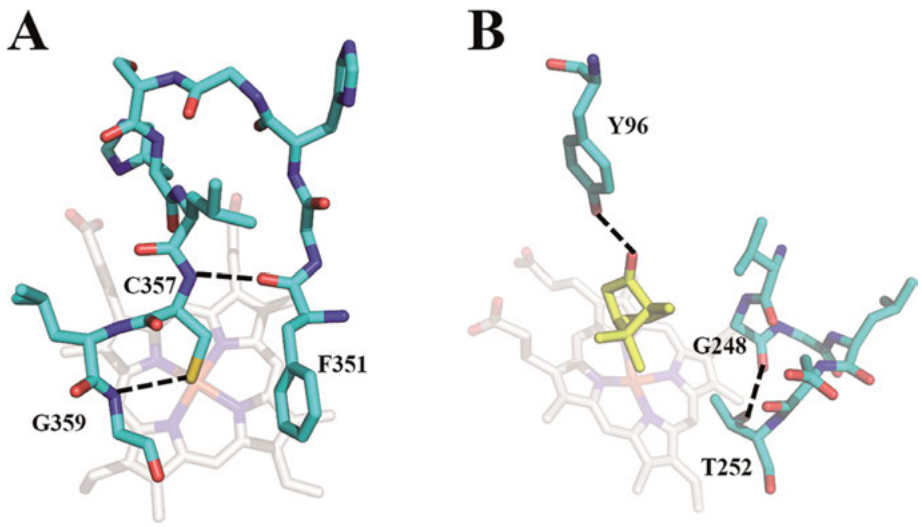


Figure 5. Crystal structure of $P450_{cam}$ (PDB ID 2CPP). (**A**) Detailed structure around the cysteine C357 heme ligand of $P450_{cam}$; this structure is highly conserved in all CYPs and forms a tight β-bulge structure. Note the H-bond between the C357 sulfur atom and a peptide NH group which helps to modulate the heme iron redox potential. (**B**) Close up of the helix I near the O_2 binding site; the local helical H-bonding pattern is disrupted and glycine G248 no longer participates in an α-helical H-bond but instead accepts an H-bond from the highly conserved threonine T252. Adapted with permission from [15] (T. L. Poulos, *Chem. Rev.* **2014**, *114*, 3919−3962); copyright (2014) American Chemical Society.

The question how structural factors influence functional properties of different heme enzymes – these included CYP, CPO, horseradish peroxidase, and secondary amine monooxygenase – was addressed in detail by pioneer researcher John Dawson in 1988 [107]. These four enzymes have several common metal coordination features but different functional properties. Clearly, the identity of the axial ligand, the nature of the heme environment, and the steric accessibility of the heme iron and heme edge combined to play major roles in determining the reactivity of each enzyme. The remarkable reactivity of cytochrome P_{450} is mostly attributed to the presence of the proximal thiolate ligand in the active site, which is proposed to exert a "push effect" and has been proposed to influence the Fe(III)/Fe(II) reduction potential, weaken the axial ligand binding, and affect the heterolytic O−O bond cleavage [108–112]. In its resting form the iron center is best described as a six-coordinate, low-spin Fe(III), with a H_2O molecule *trans* to the sulfur ligand as shown by ^{17}O-pulsed EPR and crystallographic techniques [8, 9, 14, 102, 113–116]. The Fe(III) state of the resting enzyme can, in principle, equilibrate between the low-spin (S = 1/2) and the high-spin state (S = 5/2). Results obtained from first principles molecular simulations indicate that the Fe(III) low-spin state may be stabilized by shortening of the proximal Fe–S bond; a bond length of less than ∼2.05 Å between the heme Fe(III) and cysteine sulfur

would stabilize the low-spin state, thus inhibiting the progress of the P450 catalytic cycle [117]. The low-spin state is favored in the absence of substrate, in the presence of substrate (e.g., camphor) the axial H_2O molecule and other water molecules are displaced from the hydrophobic pocket of the $P450_{cam}$ enzyme (discussed in Section 4). The spin equilibrium of the Fe(III) center is shifted toward the high-spin form. This modification of the coordination sphere of the metal center induces a significant change in the redox potential of the iron center after substrate binding and will facilitate the reduction of the penta-coordinated Fe(III) center by the reductase [8].

Further important insight into the role of the proximal sulfur ligand came from variable temperature Mössbauer and X-ray absorption spectroscopies of CYP and CPO. Both enzymes catalyze the activation of C–H bonds employing the principal intermediate Cpd I (discussed in Section 4) [52–55]. Hereby the CYP Cpd I is known to be significantly more reactive than CPO Cpd I, which only cleaves activated C–H bonds. The length of the Fe–S bond of CYP Cpd I obtained by these experiments was 2.27 Å compared to 2.36 Å for Cpd I of CPO. This significant difference in Fe–S bond lengths can be understood in terms of variations in hydrogen bonding patterns within the cysteine pocket, a portion of the proximal helix that encircles the thiolate ligand as discussed above. Weaker hydrogen bonding in CYP Cpd I results in a shorter Fe–S bond, which enables greater electron donation from the axial thiolate ligand which may in part explain CYP's greater propensity for C–H bond activation [53, 54].

Protonation or deprotonation of the proximal sulfur ligand often accompanies redox transformations of heme thiolate centers, such as the transformation of active cytochrome P450 to inactive cytochrome P420 [118–126]. To get further insight into these phenomena, and to establish spectroscopic signatures for such systems, five-coordinate adducts of the Fe(II) myoglobin His94Gly cavity variant with neutral thiol (R-SH) and thioether (R-S-R′) ligands were investigated [127]. Furthermore, a series of Thr78Cys/Lys79Gly/Met80X variants of redox protein cytochrome *c* have been engineered in which Cys78 became one of the axial heme ligands. At neutral pH, the protonation state of the coordinated cysteine differed for the Fe(III) and Fe(II) states, with cysteine binding as a thiolate ($R\text{-}S^-$) and a thiol (R-SH), respectively.

Analysis of redox-dependent stability and alkaline transitions of the cytochrome *c* variants demonstrated that the protein scaffold and solvent interactions take important roles in stabilizing a particular Cys–heme coordination [128]. Earlier, the heme protein nitrophorin was used as a template for replacing the native axial histidine ligand (His60) with cysteine. Crystal structures of the His60Cys variant showed that the ligand exchange was well tolerated by the protein scaffold, binding experiments with exogenous ligands indicated that heme–thiolate coordination required some movement of the heme within its binding cavity [129]. Density functional theory (DFT) calculations were used to investigate the influence of the axial ligand on the geometry, electronic structure, and spin states of the active heme site, comparing models with His, His + Asp, Cys, Tyr, and Tyr + Arg as found in myoglobin and hemoglobin, heme peroxidases, CYP, and heme catalases, respectively [130]. The results of the

theoretical study show, that it is harder to reduce and protonate the O_2 complex with His than with the other ligands, in accordance with the O_2 carrier function of globins and the oxidative chemistry of the other proteins. For most properties, the trend Cys < Tyr < Tyr + Arg < His + Asp < His is found, reflecting the donor capacity of cysteine. It is easier to reduce Cpd I with a His + Asp ligand than with a Cys ligand, in accordance with the one-electron chemistry of peroxidases and the hydroxylation reactions of CYPs [130]. In general, mutation of the key proximal cysteine residue of CYP abolishes monooxygenase activity, recent work by Arnold and associates [131, 132] with CYP102 from *B. megaterium* and the thermostable CYP119 from *S. acidocaldarius* has shown that mutation to either serine or histidine unlocks non-natural carbine and nitrene-transfer activities. The first crystal structure of a His-ligated CYP (Cys317His) variant revealed an intact heme iron coordination sphere through the introduced His ligand, accompanied by large changes in the overall protein structure. Notably the axial cysteine C317 could be substituted with any other amino acid without abrogating folding and heme cofactor incorporation, producing CYP variants with unusual spectral features. As mentioned earlier, these novel, highly stable enzyme active sites will be fruitful starting points for investigations of non-natural CYP catalysis and mechanisms [131, 132]. Earlier, replacement of the proximal thiolate (Cys29) in related CPO from the fungus *Caldariomyces fumago* with a histidine residue led to a variant, which retained most of their chlorination, peroxidation, epoxidation, and catalase activities. This result downplayed the importance of the proximal thiolate ligand in CPO to some extent and suggests that the distal environment of the heme active site plays the major role in maintaining the diverse activities of this enzyme [124, 133].

The proximal thiolate ligand exerts a "push effect" and has been proposed to influence the Fe(III)/Fe(II) reduction potential. To explore the effect of an increased level of donation of electrons by the proximal ligand in the CYP catalytic cycle, seleno-cysteine (SeCys) was incorporated *in silico* [134] into $P450_{cam}$ and *in vivo* into the active site of CYP119, a cytochrome P_{450} isolated from thermophilic organism *S. acidocaldarius* [95, 135, 136]. Interestingly, in the reaction of Se-CYP119 with *m*-chloroperbenzoic acid it appears that the seleno-cysteinate ligand ($Cys\text{-}Se^-$) reduced rather than stabilized the reaction intermediate Cpd I. It also protected the heme from oxidative destruction, leading to the generation of a new stable species with an UV/Vis absorption maximum at 406 nm. Thus, in thermostable Se-CYP119 the oxidative damage shifted from the heme to the proximal ligand, presumably because (a) an increased level of donation of electrons more efficiently quenches reactive species such as Cpd I and (b) the protection of the thiolate ligand provided by the protein active site structure is insufficient to shield the more oxidizable selenolate ligand [136]. Hilvert and associates [137] inserted seleno-cysteine at position 357 into $P450_{cam}$ by bacterial expression of an engineered gene containing the requisite UGA codon for seleno-cysteine. The sulfur-to-selenium substitution subtly modulated the structural, electronic, and catalytic properties of the enzyme. Catalytic activity decreased only 2-fold, whereas substrate oxidation became partially uncoupled from elec-

tron transfer. Although seleno-cysteine insertion required two additional amino acid changes, the structural integrity of the active site was maintained in the structure [137]. The Fe–Se bond length is somewhat longer than the Fe–S bond length in the wild-type $P450_{cam}$ structure and falls within the range of experimentally determined Fe–Se bond distances (2.39–2.49 Å) for synthetic Fe–selenolate complexes and non-heme enzymes. Despite the larger size of selenium, the network of hydrogen bonding interactions with the amide bonds, which regulate the redox potential of the Fe center, is preserved in the seleno-enzyme. The redox potential of the Fe(III)/Fe(II) couple of the Se-CYP had a 48 mV more negative peak potential than the control protein, consistent with the stronger electron-donating effect of selenium compared with sulfur [137].

3.3. Electron Transfer Complexes

Most of the P450 structural work has been done on single proteins and their variants. However, almost all CYPs operate with redox partners containing flavin or Fe-S cofactors to initiate and complete monooxygenase activity. Class I CYP systems (predominately found in prokaryotes) consist of multiple redox partners where NAD(P)H-derived electrons are transferred to the P450 heme center through a flavoprotein and an iron-sulfur protein. Most mammalian microsomal CYPs belong to class II, with highest concentrations in the liver and adrenal glands; they are membrane-bound and accept electrons from a membrane-bound cytochrome P450 reductase which contains FAD and FMN redox cofactors as discussed in Section 2.1 [21, 65, 138]. The crystal structure (2.03 Å resolution) of the electron transfer complex system of P450BM-3 from *B. megaterium* (CYP102) was the first structure reported (PDB ID 1BVY) [69]. This self-sufficient fatty acid monooxygenase carries heme- and FMN/FAD-containing reductase domains linked together on one single polypeptide, and represents an excellent model system for studying the mechanism of electron transfer, with the following key features:

(i) the flavodoxin-like flavin domain is positioned at the proximal face of the heme domain, the flavin and heme planes are nearly perpendicular, the 7-methyl group of the FMN is 4.1 and 4.0 Å away from the carbonyl oxygen of Ile385 and amide nitrogen of Gln387, respectively, and 18.4 Å away from the heme iron,

(ii) the indole ring of Trp574, shown to be critical for electron transfer from the FMN to the heme, is coplanar to the isoalloxazine ring of the FMN and shields a significant part of it from solvent,

(iii) the proximity of the isoalloxazine ring and the indole ring of Trp574, with their conjugated π-orbitals, to the heme-binding peptide implicates the latter as the through-bond electron-transfer pathway from FMN to the heme; the heme binding peptide represents the most efficient and coupled through-bond electron transfer pathway to the Fe center (Figure 6) [69].

Since then, much progress in our understanding of redox partner dynamics and electron transfer has been achieved by X-ray crystallography, NMR and

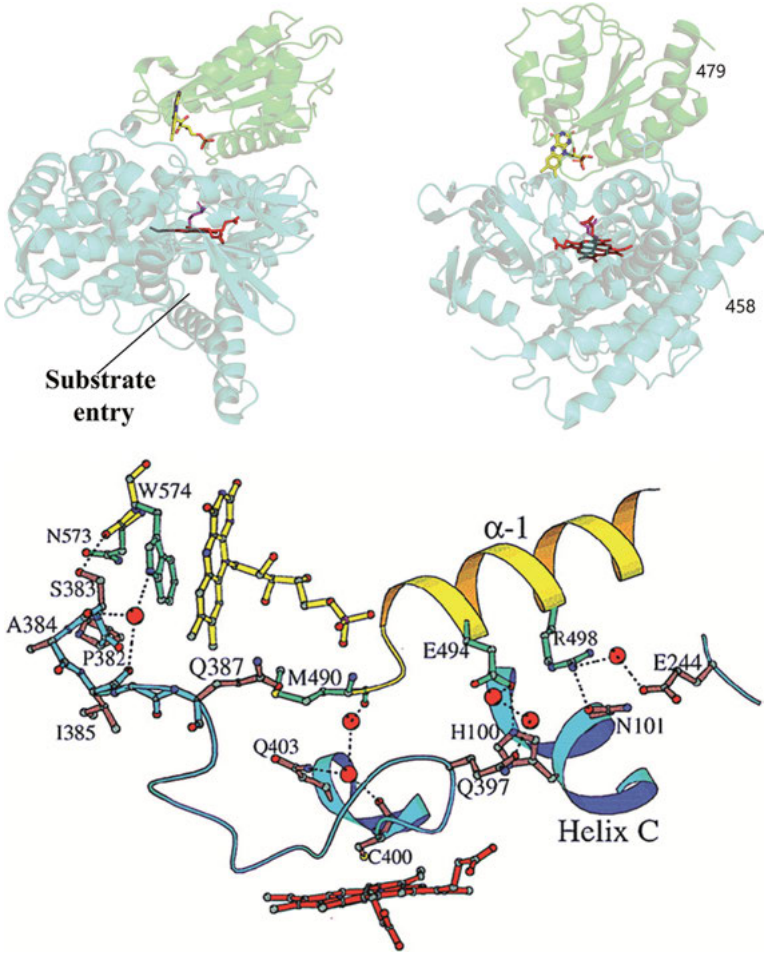


Figure 6. Crystal structure of the electron transfer complex system of P450BM-3 from *B. megaterium* (PDB ID 1BVY); taken from reference [69]. (**Top**) Two views of the heme-FMN domain complex of cytochrome P450BM-3 with the heme domain(blue), flavin domain (green), the FMN (yellow), and the heme (red). The proximal side of the heme domain is defined as the surface of the protein closest to the heme ligand cysteine C400 is shown in magenta (left). The distal surface forms part of the substrate-binding pocket. The last visible residue in the heme domain and the first visible residue in the flavin domain are labeled 458 and 479, respectively (right). The length of the missing linker, which was cleaved by proteolysis during crystallization, allows connection between the two domains. (**Bottom**) The heme-FMN domain interface. Blue and yellow colors are assigned to the backbones for the heme and flavin domains, respectively. Some of the solvent molecules mediating domain–domain interaction are shown in red. The dotted lines represent hydrogen bonds and salt bridges. The proline P382-cysteine C400 part of the heme-binding peptide provides a direct through-bond electron transfer pathway from the FMN to the heme; adapted with permission from [69] (I. F. Sevrioukova, H. Li, H. Zhang, J. A. Peterson, T. L. Poulos, *Proc. Natl Acad. Sci. USA* **1999**, *96*, 1863–1868); copyright (1999) National Academy of Sciences, U. S.A.

2D infrared spectroscopy, isothermal titration calorimetry, and increasingly by computational methods [82–84, 139–159].

In humans, the precursor to all steroid hormones, pregnenolone, is synthesized from cholesterol by an enzyme complex comprising adrenodoxin reductase, adrenodoxin, a [2Fe-2S] redox protein, and $P450_{scc}$ (CYP11A1). This complex plays a key role in steroidogenesis, its crystal structure was the first of a complex between a eukaryotic CYP and its redox partner (PDB ID 1AYF) [146, 147]. The structures with substrate and a series of reaction intermediates allowed to define the mechanism underlying sequential hydroxylations of cholesterol and to propose the mechanism of C–C bond cleavage. In the complex the [2Fe-2S] cluster of adrenodoxin is positioned 17.4 Å away from the heme iron center. After the initial protein–protein association driven by electrostatic forces, the complex adopts an optimized geometry between the redox centers. Conservation of the interaction interface suggests that this mechanism is common for all mitochondrial P450s. The distance between the iron atom closest to the surface in the [2Fe-2S] cluster and the heme iron is 17.4 Å and is similar to that observed between the heme- and FMN-binding domains of the bacterial electron transfer complex system of P450BM-3 from *B. megaterium* (CYP102) [69]. Notably, a major structural change was observed in the putidaredoxin-$P450_{cam}$ X-ray (PDB ID 4JWS, 4JWU) [141] and nuclear magnetic resonance (PDB ID 5GXG) structures [83, 84]. Originally it was anticipated that putidaredoxin would favor the closed substrate-bound $P450_{cam}$ state, which differs substantially from the open conformer. However, the crystal structure of oxidized and reduced $P450_{cam}$ complexed with its redox partner putidaredoxin (2.2 and 2.09 Å resolution) revealed that the iron-sulfur redox partner favored the open state.

These new structural data indicate that the effector role of putidaredoxin is to shift $P450_{cam}$ toward the open conformation, which enables the establishment of a water-mediated H-bonded network, which is required for proton-coupled electron transfer [83, 84, 141]. More recently, molecular dynamics docking was used to produce a model of the FMN-containing protein cindoxin with its redox partner, $P450_{cin}$ (CYP176A) from the soil bacterium *Citrobacter brakki*. The $P450_{cin}$ system has a FMN-containing flavodoxin redox partner instead of a [2Fe-2S] ferredoxin. This CYP catalyzes the regio- and stereo-selective oxidation of 1,8-cineole to (1R)-6β-hydroxy-cineole [144, 145, 147, 148]. The model highlights the potential importance of the cindoxin residue Tyr96 in bridging the FMN and heme cofactors as well as $P450_{cin}$ residues Arg102 and Arg346 in the electron transfer process [147, 148]. In the well-studied $P450_{cam}$ system [83, 84, 151, 154, 156–158], redox partner binding stabilizes the open low-spin conformation of $P450_{cam}$ and greatly decreases the stability of the Fe-O_2 complex. Cindoxin does not shift $P450_{cin}$ to a low-spin state, although the stability of the Fe-O_2 complex of $P450_{cin}$ is decreased 10-fold, suggesting that cindoxin has a modest effect on the open−closed equilibrium in $P450_{cin}$ compared to that in $P450_{cam}$ [148]. Models of complexes between P450 (CYP2B4; CYP17A1) and redox partners (cytochrome b_5; reductase FMN domain) were constructed from NMR and mutagenesis data helped to identify key domains and residues for catalytic regulation [143, 153].

4. DIOXYGEN ACTIVATION TO BREAK STRONG C–H BONDS

Nature has evolved a great variety of catalysts that contain open shell metal atoms, preferentially iron or copper, for the adaptation of life to aerobic environments for oxidative metabolism and protection against reactive oxygen species, overcoming the high kinetic barrier imposed by the O_2 molecule ($^3\Sigma^+$ ground state, S = 1) by undergoing single-electron transfer reactions. The products of these electron transfer reactions are reactive metal-oxygen intermediates, such as metal-superoxo, metal-peroxo, and high-valent metal-oxo species, in a process called dioxygen activation. Much of the understanding of the reactivity of these catalytic intermediates has been derived from the study of heme proteins and biomimetic porphyrin complexes [160, 161].

How does Nature cleave strong C–H bonds? By making stronger O–H bonds. This formalism is the central paradigm in our understanding of cytochrome P_{450} catalyzed hydroxylation of saturated hydrocarbons (Figure 7) [110]. The P450 heme iron center − able to oxidize inert C–H bonds – is assisted by the proximal sulfur ligand, a cysteine thiolate (Figure 1). Chemically, this feature is difficult to understand, thiolate sulfur is a strong electron donor to iron. Furthermore, it is easily oxidized, thus represents no obvious candidate to generate an enzyme-bound oxidant, Cpd I, strong enough to break the 410 kJ mol^{-1} C–H bond of a typical hydrocarbon. To cite leading P450 researcher John (Jay) Groves, it seemed counterintuitive to most coordination chemists and heme protein biochemists that thiolate sulfur coordination could be advantageous for an enzyme designed to oxidize even aliphatic hydrocarbons that have very high oxidation potentials and very strong C–H bonds, and we recommend his review entitled "Cytochrome P450 enzymes: understanding the biochemical hieroglyphs" to get an expert introduction to P450 catalysis [162].

4.1. The P450 Catalytic Cycle

Despite the large number of different CYPs which catalyze distinct oxidative transformations on a wide variety of molecules, such as alkanes, fatty acids, steroids, and exogenous pollutants, they share a sophisticated catalytic cycle that involves a series of consecutive transformations of the heme thiolate active site. Each of these steps has its characteristic structural and electronic features (Figure 7). In brief, dioxygen gets activated at a heme iron center, strong C–H bonds are cleaved, and stronger O–H bonds are formed.

The cycle starts with the octahedral Fe(III) heme center (**1**) (resting state), coordinated axially by a water molecule and a cysteine thiolate, residing in the low-spin, S = 1/2, state [114, 115]. As the substrate is moving towards the iron center, the bound water is displaced followed by a geometrical change producing a high-spin (S = 5/2) penta-coordinate Fe(III) heme, (**2**). This state has a more positive reduction potential, rising by 80–130 mV, the low to high-spin transition results in

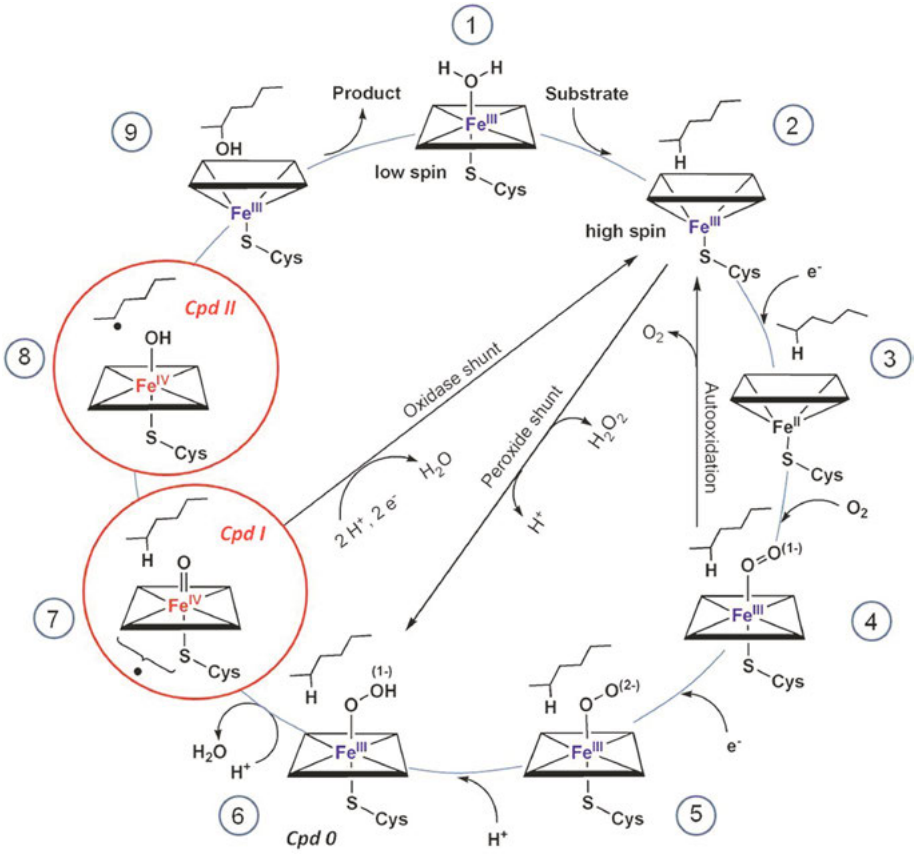


Figure 7. The catalytic P450 cycle. In the first step the substrate binds to the resting low-spin Fe(III) enzyme (**1**). This process induces structural changes, which often, but not always, manifest themselves in the dissociation of the distally coordinated water and the conversion of the heme from low to high-spin (**2**).These structural changes facilitate reduction of the Fe(III) enzyme, allowing delivery of the first electron to generate the Fe(II) substrate-bound form of the enzyme (**3**). Dioxygen then binds to the Fe(II) heme, forming a species that is best described as a Fe(III) superoxide complex (**4**). The subsequent reduction of this species forms a Fe(III) peroxo species (**5**), which is protonated at the distal oxygen to generate a Fe(III) hydroperoxo complex (**6**). The delivery of an additional proton to the distal oxygen cleaves the O–O bond, yielding Cpd I (**7**) and a water molecule. Cpd I then abstracts hydrogen from substrate to yield Cpd II (**8**) and a substrate radical, which rapidly recombine to yield the hydroxylated product and the Fe(III) enzyme (**9**). Hydroxylated product then dissociates, and water coordinates to the heme to regenerate the resting Fe(III) enzyme (**1**).

a much easier reduction to a Fe(II) state by the reductase protein (Section 3.3), to give a high-spin Fe(II) heme (S = 2) (**3**) [163–165]. It is this penta-coordinate Fe(II) species which binds O_2; note that the binding energy of H_2O in the Fe(III) complex is 10 times higher than in the corresponding Fe(II) complex, consistent

with the reduction process and displacement of H_2O [166]. Dioxygen binds to the Fe(II) heme, stabilizing a diamagnetic hexa-coordinate Fe(III)-superoxide complex (low-spin Fe(III)) (S = 1/2) plus a $O_2^{\cdot-}$ radical (S = 1/2) (**4**) [102, 167]. In the following electron transfer step the Fe(III)-$O_2^{\cdot-}$ species becomes reduced to the paramagnetic Fe(III)-peroxo compound (S = 1/2) (**5**) [168]. Hereafter, the dioxygen moiety of complex (**5**) is protonated rapidly to give the Fe(III)-hydroperoxo complex (**6**) (Cpd 0, S = 1/2) [169]. This step of the reaction cycle is still not fully understood, both from an experimental and from a computational point of view. For instance, the Fe(III)-hydroperoxo intermediate is scarcely observed, and extreme experimental conditions are needed [50]. Furthermore, a second proton attacks the distal dioxygen in complex (**6**), breaking the O–O bond, producing a water molecule and the Fe(IV) species Cpd I, a Fe(IV)=O porphyrin π radical (**7**) [50]. Since Cpd I was elusive for its experimental observation, its properties were inferred from comparison with other related species and from theoretical studies [170]. Note that C–H cleavage was only observed in heme iron centers carrying the axial thiolate ligand (discussed in Section 3.2). Experimental characterization of Cpd I was achieved recently, confirming its key role in the oxidation of the substrate [170]. The highly reactive [Fe(IV)=O] porphyrin radical cation, Cpd I, attacks the substrate directly, activates the C–H bond, followed by its homolytic cleavage via hydrogen atom transfer (HAT) and production of the Fe(IV)-hydroxo species (Cpd II) and a radical substrate (**8**) [52–55]. This becomes further stabilized by means of a re-bound mechanism to yield the final product, R-OH, which is subsequently released into the medium. Finally, the penta-coordinate Fe(III) compound (**9**) will bind a water molecule to give the resting state (**1**), closing the catalytic cycle.

Cpd I carries altogether three unpaired electrons, thus both doublet and quartet spin states have been considered in formulating the hydroxylation process (so-called two-state reactivity) [171–178]. In fact, in the case of camphor, hydroxylation occurred with both doublet and quartet states of Cpd I [173]. Experimental and theoretical studies for different substrates confirmed the overall features of the re-bound mechanism [166, 169], documenting that the first hydrogen abstraction is the rate-limiting step for this part of the P_{450} catalytic cycle.

CYPs also react with alkenes. Surprisingly, both hydroxylation and epoxidation reactions are observed with these substrates containing double C=C bonds. Hereby the hydroxylation of the C–H bonds of alkenes follows the re-bound mechanism described for the hydroxylation of alkanes [179, 180]. The epoxidation of alkenes proceeds in several steps. In the first step, which is the rate determining step, the C–O bond between the Cpd I Fe(IV) oxygen atom and the C=C sp^2 carbon atom is formed [180]. In the reaction of several CYPs with cyclohexene significant amounts of both hydroxylation and epoxidation products were found, implying roughly equal activation barriers for these two oxidation reactions [181–183].

Finally, the hydroxylation reaction catalysed by P450s and biomimetic iron porphyrin complexes usually show quite large primary kinetic isotope effects ($k_{C\text{-}H/C\text{-}D}$, 4–22) [18–186], thus are consistent with the breakage of the C–H bond of the substrate being the rate-determining step, generating the intermediate

Fe(IV)-OH (Cpd II) and alkyl-radical species via a HAT mechanism. Clearly, further experimental and computational data will be needed to fully understand all the mechanistic details of P450 reactivity [162, 187, 188].

4.2. Watching P450 at Work

The identification and characterization of the powerful oxidants Cpd I and Cpd II of the P450 catalytic cycle (Figure 7) represented a major challenge to spectroscopists and crystallographers which was crucial for our understanding of the overall catalytic process [15, 102, 114, 115, 189–197]. Although CYPs have been studied for over five decades, key questions remained with respect to the nature of Cpd I and Cpd II, their ability to oxidize substrates, and the role of the proximal thiolate ligand for catalysis. Progress in this regard has been stalled by the difficulty in trapping the reaction intermediates, which are usually short-lived and generated only in low yield. Recently, thanks to the efforts by Green and associates, both Cpd I and Cpd II could be obtained in remarkably high yield for the first time by using two unique P450s (CYP119, thermophilic, and CYP158, with a large, solvent-exposed active site) [55, 170]. All members of the CYP family exhibit similar unique spectral features, most prominent the split Soret absorption maximum at 450 nm of the Fe(II)-CO adduct. Each catalytic step of the cycle is characterized by specific changes in oxidation and spin state of heme iron, planarity of the porphyrin ring or disposition of the dioxygen moiety, providing a specific spectral signature that can be easily detected and analyzed by spectroscopic techniques and serve as a source of indispensable information about structure, dynamic and functional properties. In addition to time-resolved crystallography [102] spectroscopic techniques, electronic absorption (UV/Vis), magnetic circular dichroism (MCD), magnetic resonance (EPR, ENDOR, NMR), Mössbauer, X-ray absorption (XAS), infrared (IR) and resonance Raman (rR), provided multifaceted and valuable structural and mechanistic insights into the functioning of these potent enzymes [197].

Density functional theory also contributed important information on the electronic and geometrical properties of the high-valent heme iron-thiolate center. One is that the thiolate ligand acquired radical character during formation of Cpd I [52–55, 170], which provided a satisfactory explanation for the strong antiferromagnetic coupling in CPO Cpd I between Fe(IV)=O ($S = 1$) and the radical center ($S = 1/2$) [187]. These calculations also address the issue of differences in Fe–S bond lengths observed in the different high-valent iron species [55, 170, 183]. Furthermore, the Fe–O bond length of 1.82 Å reported by Green and colleagues [52] for CPO Cpd II is significantly longer than the typical length of ≈ 1.65 Å observed in related iron-oxo species [171, 172], suggesting protonation of the oxygen in CPO Cpd II, in agreement with DFT calculations [187].

5. FUTURE DIRECTIONS

A recent survey of human oxidoreductases and P450 enzymes involved in the metabolism of xenobiotic and natural chemicals revealed that > 90 % of enzymatic reactions are catalyzed by CYPs, underlining their dominant role in the oxidative activation of these compounds. They generally utilize mixed-function oxidase stoichiometry, and use pyridine nucleotides as electron donors (Equation 1). The catalysis of oxidations is largely understood in the context of the formation of ferryl intermediates (Fe^{IV}=O), such as Cpd I, named because of its history in peroxidase chemistry (Figure 7) [198–200]. However, despite six decades of intensive research since the description of a new CO-binding pigment by Martin Klingenberg [201], there still remain many unresolved issues concerning Nature's blowtorch P450 according to Peter Guengerich [202–204].

Clearly, P450 has become a mature field. The basic catalytic mechanism is understood at the atomic level in the context of electronic changes leading to O_2 activation and the oxidation of a substrate (Figure 7) [205]. The catalytic versatility makes CYPs particularly valuable for numerous applications. Nanodisc technology combined with various biophysical methods was recently introduced by Denisov and Sligar [206]. Nanodiscs provide a new and powerful tool for a broad spectrum of biochemical and biophysical studies of membrane proteins. They are acknowledged as an optimal membrane mimetic system that provides control over size, composition, and specific functional modifications on the nanometer scale. Important structural and mechanistic information on membrane-bound proteins including CYPs can be obtained by this technology, ranging from quaternary information, protein−protein and protein−lipid recognition to highly resolved molecular pictures [197, 206]. As CYPs are key enzymes in the metabolism of drugs, electrochemical CYP-based biosensors are of great interest for pharmaceutical industry to develop sensitive and cost-effective systems for screening new drugs and investigation of drug–drug interactions, as well as for screening potential new substrates and inhibitors of CYPs [207]. Another interesting application of CYPs is the light-driven biosynthesis of important natural products. By fusion of the iron-sulfur protein ferredoxin with CYP79A1 from the plant *Sorghum bicolor*, CYP79A1 can receive electrons for catalysis by interacting directly with photosystem I, thus will outperform non-fused CYP79A1 *in vivo* [152]. Sequential layer-by-layer assembly was used to design nanostructured films made of recombinant bacterial membrane fractions, which overexpress P450 and cytochrome P450 reductase. Hereby, the membrane fractions preserve a remarkable CYP catalytic property in the adsorbed phase. Atomic force microscopy images reveal that the membrane fractions mostly adopt a flattened conformation in the adsorbed phase with an extensive tendency to aggregate within the multilayered films, which is more pronounced when increasing the number of bilayers. Interestingly, this behavior seems to enhance the ability of embedded membrane fractions to remain active after repeated uses. The proposed strategy constitutes a practical alternative for the immobilization of active CYP enzymes. Membrane fraction-based multilayers represent useful nano-objects for the creation of new biomimetic reactors for the assessment of CYP xenobiotic metabolism [208].

The catalytic proficiency of CYPs largely owes to the highly oxidizing power of the intermediate formed at the iron active site, Cpd I, generated from atmospheric O_2. Its activation involves the orchestrated delivery of reducing equivalents from NAD(P)H through an associated redox chain, and finely coordinated active-site proton delivery. The dual requirement for accessory enzymes and pyridine nucleotide co-substrate imposes significant limitations on the successful deployment of P450 enzymes for large-scale biotechnological applications. One promising alternative could come from a shunt pathway in which hydrogen peroxide is used to generate Cpd I [209]. Unfortunately, the efficiency of "peroxygenase" turnover by O_2 activating P450s is rather low, due to inefficient O-O heterolysis and subsequent enzyme inactivation. The CYP enzyme OleT uses H_2O_2 as a co-substrate, it removes the carboxylate from C_n chain length fatty acid substrates to produce C_{n-1} terminal alkenes and CO_2, a decarboxylation reaction of considerable commercial interest, which may serve as a useful biocatalytic platform for the synthesis of important commodity chemicals. The OleT-catalyzed $C-C_\alpha$ scission reaction is a significant departure from the *classical* function of most CYPs, where an oxygen atom derived from atmospheric dioxygen is inserted into chemically unreactive substrates, generating a metabolite that is more water-soluble than the parent compound [210–212].

In summary, CYP research will continue to be a rich, vibrant, and important field of research. Determining and understanding the reaction mechanisms of CYP substrate oxidations have greatly advanced a variety of fields and our understanding of dioxygen activation and C-H bond cleavage. Molecular dynamics simulations and quantum-mechanical/molecular-mechanical calculations will complement experiments and elucidate the choreography by which the CYP protein regulates the catalytic cycle. New P450 enzymes and unusual P450 reactions are being discovered, and P450 enzymes can be engineered and evolved for particular applications [15, 17, 77, 162, 213–218].

ACKNOWLEDGMENTS

The authors are grateful for continuous financial support by DGAPA-UNAM, Facultad de Química-UNAM and CONACYT (FMCM, RDPL, PDSP, MEST), and by Deutsche Forschungsgemeinschaft and Universität Konstanz (PK).

ABBREVIATIONS AND DEFINITIONS

Cpd I	compound I; Fe(IV)=O porphyrin π radical
Cpd II	compound II; oxyferryl [Fe=O] intermediate
CPO	chloroperoxidase
CYP	cytochrome P_{450} (P450)
DFT	density functional theory
EPR	electron paramagnetic resonance

FAD	flavin adenine dinucleotide
FMN	flavin mononucleotide
HAT	hydrogen atom transfer
NADH	nicotinamide adenosine dinucleotide, reduced
NADPH	nicotinamide adenosine dinucleotide phosphate, reduced
NOS	nitric oxide synthase
PDB	Protein Data Bank; https://www.rcsb.org/
UV/Vis	ultraviolet/visible

REFERENCES

1. *The Ubiquitous Roles of Cytochrome P450 Proteins*. Vol. 3 of *Metal Ions in Life Sciences*, Eds A. Sigel, H. Sigel, R. K. O. Sigel, John Wiley & Sons Ltd, Chichester, UK, 2007.
2. *Cytochrome P450: Structure, Mechanism, and Biochemistry*. 4th edn., Ed. P. R. Ortiz de Montellano, Springer International Publishing, Switzerland, 2015.
3. *Monooxygenase, Peroxidase and Peroxygenase Properties and Mechanisms of Cytochrome P450*, Eds E. G. Hrycay, S. M. Bandiera, Springer International Publishing, Switzerland, 2015.
4. V. Ullrich, *Angew. Chem. Inter. Ed.* **1972**, *11*, 701–712.
5. M. Sono, M. P. Roach, E. D. Coulter, J. H. Dawson, *Chem. Rev.* **1996**, *96*, 2841–2887.
6. J. H. Capdevila, J. R. Falck, R. C. Harris, *J. Lipid Res.* **2000**, *41*, 163–181.
7. J. T. Groves, *Proc. Natl Acad. Sci. USA* **2003**, *100*, 3569–3574.
8. B. Meunier, S. P. de Visser, S. Shaik, *Chem. Rev.* **2004**, *104*, 3947–3980.
9. I. G. Denisov, T. M. Makris, S. G. Sligar, I. Schlichting, *Chem. Rev.* **2005**, *105*, 2253–2277.
10. M. Hofrichter, R. Ullrich, *Appl. Microbiol. Biotechnol.* **2006**, *71*, 276–288.
11. R. Bernhardt, *J. Biotechnol*. **2006**, *124*, 128–145.
12. V. B. Urlacher, M. Girhard, *Trends in Biotechnology* **2012**, *30*, 26–36.
13. D. P. Collins, J. H. Dawson in *Comprehensive Inorganic Chemistry II*, 2nd edn., *From Elements to Applications* **2013**, *3*, 65–102.
14. T. Omura, *Biotechnol. Appl. Biochem.* **2013**, *60*, 4–8.
15. T. L. Poulos, *Chem. Rev.* **2014**, *114*, 3919–3962.
16. G. Gilardi, G. Di Nardo, *Rend. Fis. Acc. Lincei* **2017**, *28* (Suppl 1), S159–S167.
17. A. W. Munro, K. J. McLean, J. L. Grant, T. M. Makris, *Biochem. Soc. Trans.* **2018**, https://doi.org/10.1042/BST20170218.
18. S. C. Hammer, G. Kubik, E. Watkins, S. Huang, H. Minges, F. H. Arnold, *Science* **2017**, *358*, 215–218.
19. A. M. Knight, S. B. J. Kan, R. D. Lewis, O. F. Brandenberg, K. Chen, F. H. Arnold, *ACS Cent. Sci*. **2018**, *4*, 372–377.
20. O. F. Brandenberg, C. K. Prier, K. Chen, A. M. Knight, Z. Wu, F. H. Arnold, *ACS Catal*. **2018**, *8*, 2629–2634.
21. A. W. Munro, H. M. Girvan, A. E. Mason, A. J. Dunford, K. J. McLean, *Trends Biochem. Sci.* **2013**, *38*, 140–150.
22. H. Renault, J.-E. Bassard, B. Hamberger, D. Werck-Reichhart, *Curr. Op. Plant Biol.* **2014**, *19*, 27–34.
23. H. M. Girvan, A. W. Munro, *Curr. Op. Chem. Biol.* **2016**, *31*, 136–145.
24. D. Schmitz, S. Janocha, F. M. Kiss, R. Bernhardt, *Biochim. Biophys. Acta* **2018**, *1866*, 11–22.

25. K. Yasuda, H. Sugimoto, K. Hayashi, T. Takita, K. Yasukawa, M. Ohta, M. Kamakura, S. Ikushiro, Y. Shiro, T. Sakaki, *Biochim. Biophys. Acta* **2018**, *1866*, 23–31.
26. H. Shalan, M. Kato, L. Cheruzel, *Biochim. Biophys. Acta* **2018**, *1866*, 80–87.
27. D. A. Sychev, G. Md. Ashraf, A. A. Svistunov, M. L. Maksimov, V. V. Tarasov, V. N. Chubarev, V. A. Otdelenov, N. P. Denisenko, G. E. Barreto, G. Aliev, *Drug Design, Development and Therapy* **2018**, *12*, 1147–1156.
28. D. R. Nelson, *Biochim. Biophys. Acta* **2018**, *1866*, 141–154.
29. D. S. Goodsell, *The Oncologist* **2001**, *6*, 205–206.
30. F. P. Guengerich, *Chem. Res. Toxicol.* **2008**, *21*, 70–8370.
31. A. Veith, B. Moorthy, *Curr. Op. Toxicol.* **2018**, *7*, 44–51.
32. J. Martin, M. Fay, *Aust. Prescr.* **2001**, *24*, 10–12.
33. P. Anzenbacher, E. Anzenbacherová, *Cell. Mol. Life Sci.* **2001**, *58*, 737–747.
34. H. S. Mason, W. L. Fowlks, E. Peterson, *J. Am. Chem. Soc*. **1955**, *77*, 2914–2915.
35. H. S. Mason, *Adv. Enzymol*. **1957**, *19*, 79–233.
36. H. Shoun, S. Fushinobu, L. Jiang, S.-W. Kim, T. Wakagi, *Phil. Trans. R. Soc. B* **2012**, *367*, 1186–1194.
37. F. P. Guengerich, *Am. Sci.* **1993**, *81*, 440–447.
38. T. Omura, R. Sato, *J. Biol. Chem.* **1962**, *237*, 1375–1376.
39. A. Y. H. Lu, K. W. Junk, M. J. Coon, *J. Biol. Chem.* **1969**, *244*, 3714–3721.
40. M. J. Coon, *J. Biol. Chem.* **2002**, *277*, 28351–28363.
41. M. J. Coon, *Annu. Rev. Pharmacol. Toxicol.* **2005**, *45*, 1–25.
42. D. C. Lamb, L. Lei, A. G. S. Warrilow, G. I. Lepesheva, J. G. L. Mullins, M. R. Waterman, S. L. Kelly, *J. Virol.* **2009**, *83*, 8266–8269.
43. R. H. Wickramasinghe, C. A. Villee, *Nature* **1975**, *256*, 509–511.
44. S. L. Kelly, D. E. Kelly, *Phil. Trans. R. Soc. B* **2013**, *368*, 20120476, http://dx.doi.org/10.1098/rstb.2012.0476.
45. Y. Miyake, H. S. Mason, W. Landgraf, *J. Biol. Chem.* **1967**, *242*, 393–397.
46. K. Murakami, H. S. Mason, *J. Biol. Chem.* **1967**, *242*, 1102–1110.
47. D. W. Smith, R. J. P. Williams, *Structure and Bonding* **1970**, *7*, 1–45.
48. S. Yoshioka, S. Takahashi, H. Hori, K. Ishimori, I. Morishima, *Eur. J. Biochem.* **2001**, *268*, 252–259.
49. M. T. Green, *J. Am. Chem. Soc.* **1998**, *120*, 10772–10773.
50. M. T. Green, *J. Am. Chem. Soc.* **1999**, *121*, 7939–7940.
51. M. T. Green, *Curr. Op. Chem. Biol.* **2009**, *13*, 84–88.
52. M. T. Green, J. H. Dawson, H. B. Gray, *Science* **2004**, *304*, 1653–1656.
53. C. M. Krest, E. L. Onderko, T. H.Yosca, J. C. Calixto, R. F. Karp, J. Livada, J. Rittle, M. T. Green, *J. Biol. Chem.* **2013**, *288*, 17074–17081.
54. C. M. Krest, A. Silakov, J. Rittle, T. H. Yosca, E. L. Onderko, J. C. Calixto, M. T. Green, *Nat. Chem.* **2015**, *7*, 696–702.
55. T. H. Yosca, A. P. Ledray, J. Ngo, M. T. Green, *J. Biol. Inorg. Chem.* **2017**, *22*, 209–220.
56. R. W. Estabrook, *Drug Metabolism and Disposition* **2003**, *31*, 1461–1473.
57. D. W. Nebert, D. R. Nelson, M. J. Coon, R. W. Estabrook, R. Feyereisen, Y. Fujii-Kuriyama, F. J. Gonzalez, F. P. Guengerich, I. C. Gunsalus, E. F. Johnson,J. C. Loper, R. Sato, M. R. Waterman, D. J. Waxman, *DNA and Cell Biol.* **1991**, *10*, 1–14.
58. F. J. Gonzalez, H. V. Gelboin, *Environ. Health Persp.* **1992**, *98*, 81–85.
59. D. R. Nelson, *Cytochrome P450 Nomenclature*, in *Cytochrome P450 Protocols.* 2nd edn., Eds I. R. Phillips, E. A. Shephard, Humana Press, New York City, 2006, pp. 1–10.
60. P. Urban, T. Lautier, D. Pompon, G. Truan, *Int. J. Mol. Sci.* **2018**, *19*, 1617; doi:10.3390/ijms19061617.
61. M. A. Schuler, S. G. Sligar, *Met. Ions Life Sci.* **2007**, *3*, 1–26.

62. T. L. Poulos, *Proc. Natl. Acad. Sci. USA* **2003**, *100*, 13121–13122.
63. S. Yoshikawa, A. Shimada, K. Shinzawa-Itoh, *Met. Ions Life Sci.* **2015**, *15*, 89–130.
64. S. Govindaraj, T. L. Poulos, *J. Biol. Chem.* **1997**, *272*, 7915–7921.
65. F. Hannemann, A. Bichet, K. M. Ewen, R. Bernhardt, *Biochim. Biophys. Acta* **2007**, *1770*, 330–344.
66. D. R. Nelson, *Phil. Trans. R. Soc. B* **2013**, *368*, 20120430, http://dx.doi.org/10.1098/rstb.2012.0430.
67. D. C. Lamb, M. R. Waterman, *Phil. Trans. R. Soc. B* **2013**, *368*, 20120434, http://dx.doi.org/10.1098/rstb.2012.0434..
68. D. J. Cook, J. D. Finnigan, K. Cook, G. W. Black, S. J. Charnock, *Adv. Prot. Chem. Struct. Biol.* **2016**, *105*, 105–126.
69. I. F. Sevrioukova, H. Li, H. Zhang, J. A. Peterson, T. L. Poulos, *Proc. Natl Acad. Sci. USA* **1999**, *96*, 1863–1868.
70. O. Einsle, G. B. Seifert, M. E. Sosa Torres, P. M. H. Kroneck, *Biospektrum* **2006**, 346–348.
71. R. Fasan, M. M. Chen, N. C. Crook, F. H. Arnold, *Angew. Chem. Int. Ed.* **2007**, *46*, 8414–8418.
72. R. Fasan, Y. T. Meharenna, C. D. Snow, T. L. Poulos, F. H. Arnold, *J. Mol. Biol.* **2008**, *383*, 1069–1080.
73. C. A. Tracewell, F. H Arnold, *Curr. Opin. Chem. Biol.* **2009**, *13*, 3–9.
74. S. T. Jung, R. Lauchli, F. H. Arnold, *Curr. Op. Biotechnol.* **2011**, *22*, 809–817.
75. O. F. Brandenberg, R. Fasan, F. H. Arnold, *Curr. Op. Biotechnol.* **2017**, *47*, 102–111.
76. F. H. Arnold, *Angew. Chem. Int. Ed.* **2018**, *57*, 4143–414.
77. A. M. Knight, S. B. J. Kan, R. D. Lewis, O. F. Brandenberg, K. Chen, F. H. Arnold, *ACS Cent. Sci.* **2018**, *4*, 372–377.
78. J. Ducharme, K. Auclair, *Biochim. Biophys. Acta* **2018**, *1866*, 32–51.
79. D. Holtmann, F. Hollmann, *ChemBioChem* **2016**, *17*, 1391–1398.
80. T. L. Poulos, B. C. Finzel, I. C. Gunsalus, G. C. Wagner, J. Kraut, *J. Biol. Chem*. **1985**, *260*, 16122–16130.
81. T. L. Poulos, B. C. Finzel, A. J. Howard, *J. Mol. Biol.* **1987**, *195*, 687–700.
82. D. Batabyal, L. S. Richards, T. L. Poulos, *J. Am. Chem. Soc.* **2017**, *139*, 13193–13199.
83. Y. Hiruma, M. A. S. Hass, Y. Kikui, W.-M. Liu, B. Ölmez, S. P. Skinner, A. Blok, A. Kloosterman, H. Koteishi, F. Löhr, H. Schwalbe, M. Nojiri, M. Ubbink, *J. Mol. Biol.* **2013**, *425*, 4353–4365.
84. W. Andrałoj, Y. Hiruma, W.-M. Liu, E. Ravera, M. Nojiri, G. Parigi, C. Luchinat, M. Ubbink, *Proc. Natl Acad. Sci. USA* **2017**, *114*, E1840–E1847.
85. P. A. Williams, J. Cosme, D. Matak Vinković, A. Ward, H. C. Angove, P. J. Day, C. Vonrhein, I. J. Tickle, H. Jhoti, *Science* **2004**, *305*, 683–686.
86. M. Ekroos, T. Sjögren, *Proc. Natl Acad. Sci. USA* **2006**, *103*, 13682–13687.
87. H. Li, *Cytochrome P450*, in *Handbook of Metalloproteins*, Eds A. Messerschmidt, R. Huber, K. Wieghardt, T. Poulos, John Wiley & Sons Ltd, Chichester, UK, 2001, pp. 1–16.
88. O. Pylypenko, I. Schlichting, *Annu. Rev. Biochem.* **2004**, *73*, 991–1018.
89. T. L. Poulos, E. F. Johnson, *Structures of Cytochrome P450 Enzymes*, in *Cytochrome P450: Structure, Mechanism, and Biochemistry.* 3rd edn., Ed. P. R. Ortiz de Montellano, Kluwer Academic/Plenum Publishers, New York, 2005, pp. 87–114.
90. T. L. Poulos, *Biochem. Biophys. Res. Comm*. **2005**, *338*, 337–345.
91. E. F. Johnson, C. D. Stout, *Biochem. Biophys. Res. Comm.* **2005**, *338*, 331–336.
92. E. F. Johnson, C. D. Stout, *J. Biol. Chem*. **2013**, *288*, 17082–17090.
93. F. P. Guengerich, M. R. Waterman, M. Egli, *Trends Pharmacol. Sci.* **2016**, *37*, 629–640.

94. K. L. Harris, R. E. S. Thomson, S. J. Strohmaier, Y. Gumulya, E. M. J. Gillam, *Biochim. Biophys. Acta* **2018**, *1866*, 97–115.
95. S.-Y. Park, K.Yamane, S.-I. Adachi, Y. Shiro, K. E. Weiss, S. A. Maves, S. G. Sligar, *J. Inorg. Biochem.* **2002**, *91*, 491–501.
96. E. E. Scott, Y. Ai He, M. R. Wester, M. A. White, C. C. Chin, J. R. Halpert, E. F. Johnson, C. D. Stout, *Proc. Natl Acad. Sci. USA* **2003**, *100*, 13682–13687.
97. I. J. Dmochowski, B. R. Crane, J. J. Wilker, J. R. Winkler, H. B. Gray, *Proc. Natl. Acad. Sci. USA* **1999**, *96*, 12987–12990.
98. M. A. Hays, A. R. Dunn, R. Chiu, H. B. Gray, C. D. Stout, D. B. Goodin, *J. Mol. Biol.* **2004**, *344*, 455–469.
99. Y.-T. Lee, R. F. Wilson, I. Rupniewski, D. B. Goodin, *Biochemistry* **2010**, *49*, 3412–3419.
100. S. Stoll, Y.-T. Lee, M. Zhang, R. F. Wilson, R. D. Britt, D. B. Goodin, *Proc. Natl. Acad. Sci. USA* **2012**, *109*, 12888–12893.
101. M. Vidakovic, S. G. Sligar, H. Li, T. L. Poulos, *Biochemistry* **1998**, *37*, 9211–9219.
102. I. Schlichting, J. Berendzen, K. Chu, A. M. Stock, S. A. Maves, D. E. Benson, R. M. Sweet, D. Ringe, G. A. Petsko, S. G. Sligar, *Science* **2000**, *287*, 1615–1622.
103. S. Nagano, T. L. Poulos, *J. Biol. Chem.* **2005**, *280*, 31659–31663.
104. E. Adman, K. D. Watenpaugh, L. H. Jensen, *Proc. Natl. Acad. Sci. USA* **1975**, *12*, 4854–4858.
105. N. Ueyama, T. Terakawa, M. Nakata, A. Nakamura, *J. Am. Chem. Soc.* **1983**, *105*, 7098–7102.
106. N. Ueyama, N. Nishikawa, Y. Yamada, T. Okamura, A. Nakamura, *J. Am. Chem. Soc.* **1996**, *118*, 1286–1287.
107. J. H. Dawson, *Science* **1988**, *240*, 433–439.
108. P. Anzenbacher, J. H. Dawson, T. Kitagawa, *J. Mol. Struct.* **1989**, *214*, 149–158.
109. P. K. Das, S. Chatterjee, S. Samanta, A. Dey, *Inorg. Chem*. **2012**, *51*, 10704–10714.
110. J. T. Groves, *Nat. Chem.* **2014**, *6*, 89–91.
111. S. Samanta, P. K. Das, S. Chatterjee, A. Dey, *J. Porphyrins Phthalocyanines* **2015**, *19*, 92–108.
112. X. Huang, J. T. Groves, *J. Biol. Inorg. Chem.* **2017**, *22*, 185–207.
113. R. Raag, T. L. Poulos, *Biochemistry* **1989**, *28*, 917–922.
114. H. Thomann, M. Bernardo, D. Goldfarb, P. M. H. Kroneck, V. Ullrich, *J. Am. Chem. Soc.* **1995**, *117*, 8243–8251.
115. D. Goldfarb, M. Bernardo, H. Thomann, P. M. H. Kroneck, V. Ullrich, *J. Am. Chem. Soc.* **1996**, *118*, 2686–2693.
116. K. P. Conner, A. M. Schimpf, A. A. Cruce, K. J. McLean, A. W. Munro, D. J. Frank, M. D. Krzyaniak, P. Ortiz de Montellano, M. K. Bowman, W. M. Atkins, *Biochemistry* **2014**, *53*, 1428–1434.
117. M. D. Segall, M. C. Payne, W. Ellis, G. T. Tucker, N. Boyes, *Chem. Res. Toxicol.* **1998**, *11*, 962–966.
118. Y. Imai, R. Sato, *Biochem. Biophys. Res. Commun.* **1966**, *22*, 620–626.
119. Y. Imai, R. Sato, *Eur. J. Biochem.* **1967**, *1*, 419–426.
120. J. P. Collmann, T. N. Sorrell, J. H. Dawson, J. R. Trudell, E. Bunnenberg, C. Djerassi, *Proc. Nat. Acad. Sci. USA* **1976**, *73*, 6–10.
121. S. P. Cramer, J. H. Dawson, K. O. Hodgson, L. P. Hager, *J. Am. Chem. Soc.* **1978**, *100*, 7282–7290.
122. M. Sono, L. A. Andersson, J. H. Dawson, *J. Biol. Chem.* **1982**, *257*, 8308–8320.
123. H. I. Liu, M. Sono, S. Kadkhodayan, L. P. Hager, B. Hedman, K. O. Hodgson, J. H. Dawson, *J. Biol. Chem.* **1995**, *270*, 10544–10550.
124. R. Perera, M. Sono, J. A. Sigman, T. D. Pfister, Y. Lu, J. H. Dawson, *Proc. Natl. Acad. Sci. USA* **2003**, *100*, 3641–3646.

125. R. Perera, M. Sono, H. L. Voegtle, J. H. Dawson, *Arch. Biochem. Biophys.* **2011**, *507*, 119–125.
126. Y. Sun, W. Zeng, A. Benabbas, X. Ye, I. Denisov, S. G. Sligar, J. Du, J. H. Dawson, P. M. Champion, *Biochemistry* **2013**, *52*, 5941–5951.
127. M. P. Roach, A. E. Pond, M. R. Thomas, S. G. Boxer, J. H. Dawson, *J. Am. Chem. Soc.* **1999**, *121*, 12088–12093.
128. F. Zhong, G. P. Lisi, D. P. Collins, J. H. Dawson, E. V. Pletneva, *Proc. Natl. Acad. Sci. USA* **2014**, *111*, E306–E315.
129. S. W. Vetter, A. C. Terentis, R. L. Osborne, J. H. Dawson, D. B. Goodin, *J. Biol. Inorg. Chem.* **2009**, *14*,179–191.
130. P. Rydberg, E. Sigfridsson, U. Ryde, *J. Biol. Inorg. Chem.* **2004**, *9*, 203–223.
131. P. S. Coelho, Z. J. Wang, M. E. Ener, S. A. Baril, A. Kannan, F. H. Arnold, E. M. Brustad, *Nat. Chem. Biol.* **2013**, *9*, 485–487.
132. J. A. McIntosh, T. Heel, A. R. Buller, L. Chio, F. H. Arnold, *J. Am. Chem. Soc.* **2015**, *137*, 13861–13865.
133. X. Yi, M. Mroczko, K. M. Manoj, X. Wang, L. P. Hager, *Proc. Natl. Acad. Sci. USA* **1999**, *96*, 12412–12417.111.
134. S. Cohen, D. Kumar, S. Shaik, *J. Am. Chem. Soc.* **2006**, *128*, 2649–2653.
135. Y. Jiang, S. Sivaramakrishnan, T. Hayashi, S. Cohen, P. Moënne-Loccoz, S. Shaik, P. R. Ortiz de Montellano, *Angew. Chem. Int. Ed.* **2009**, *48*, 7193–7195.
136. S. Sivaramakrishnan, H. Ouellet, J. Du, K. J. McLean, K. F. Medzihradszky, J. H. Dawson, A. W. Munro, P. R. Ortiz de Montellano, *Biochemistry* **2011**, *50*, 3014–3024.
137. C. Aldag, I. A. Gromov, I. García-Rubio, K. von Koenig, I. Schlichting, B. Jaun, D. Hilvert, *Proc. Natl. Acad. Sci. USA* **2009**, *106*, 5481–5486.
138. I. Hanukoglu, *Adv. Mol. Cell Biol.* **1996**, *14*, 29–56.
139. I. F. Sevrioukova, T. L. Poulos, *Proc. Natl. Acad. Sci. USA* **2010**, *107*, 18422–18427.
140. I. F. Sevrioukova, T. L. Poulos, *Arch. Biochem. Biophys.* **2011**, *507*, 66–74.
141. S. Tripathi, H. Li, T. L. Poulos, *Science* **2013**, *340*, 1227–1230.
142. S. A. Hollingsworth, D. Batabyal, B. D. Nguyen, T. L. Poulos, *Proc. Natl Acad. Sci. USA* **2016**, *113*, 8723–8728.
143. S. Ahuja, N. Jahr, S.-C. Im, S.Vivekanandan, N. Popovych, S. V. LeClair, R. Huang, R. Soong, J. Xu, K.Yamamoto, R. P. Nanga, A. Bridges, L.Waskell, A. Ramamoorthy, *J. Biol. Chem.* **2013**, *288*, 22080–22095.
144. D. B. Hawkes, G. W. Adams, A. L. Burlingame, P. R. Ortiz de Montellano, J. J. De Voss, *J. Biol. Chem.* **2002**, *277*, 27725–27732.
145. N. Kimmich, A. Das, I.Sevrioukova, Y. Meharenna, S. G. Sligar, T. L. Poulos, *J. Biol. Chem.* **2007**, *282*, 27006–27011.
146. N. Strushkevich, F. MacKenzie, T. Cherkesov, I. Grabovec, S. Usanov, H.-W. Park, *Proc. Natl Acad. Sci. USA* **2011**, *108*, 10139–10143.
147. N. Mast, A. J. Annalora, D. T. Lodowski, K. Palczewski, C. D. Stout, I. A. Pikuleva, *J. Biol. Chem.* **2011**, *286*, 5607–5613.
148. Y. Madrona, S. A. Hollingsworth, B. Khan, T. L. Poulos, *Biochemistry* **2013**, *52*, 5039–5050.
149. Y. Madrona, S. A. Hollingsworth, S. Tripathi, J. B. Fields, J.-C. N. Rwigema, D. J. Tobias, T. L. Poulos, *Biochemistry* **2014**, *53*, 1435–1446.
150. H.-M. Peng, J. Liu, S. E. Forsberg, H. T. Tran, S. M. Anderson, R. J. Auchus, *J. Biol. Chem.* **2014**, *289*, 33838–33849.
151. S. A. Hollingsworth, T. L. Poulos, *Prot. Sci.* **2015**, *24*, 49–57.
152. S. Busck Mellor, A. Zygadlo Nielsen, M. Burow, M. Saddik Motawia, D. Jakubauskas, B. Lindberg Møller, P. E. Jensen, *ACS Chem. Biol.* **2016**, *11*, 1862–1869.
153. D. Fernando Estrada, J. S. Laurence, E. E. Scott, *J. Biol. Chem.* **2016**, *291*, 3990–4003.

154. D. Batabyal, A. Lewis-Ballester, S.-R. Yeh, T. L. Poulos, *Biochemistry* **2016**, *55*, 6517–6523.
155. W. Zhang, L. Du, F. Li, X. Zhang, Z. Qu, L. Han, Z. Li, J. Sun, F. Qi, Q. Yao, Y. Sun, C. Geng, S. Li, *ACS Catal.* **2018**, *8*, 9992–10003.
156. S. Ramos, E. J. Basom, M. C. Thielges, *Front. Mol. Biosci.* **2018**, *5*, 94. doi: 10.3389/fmolb.2018.00094.
157. D. Batabyal, T. L. Poulos, *J. Inorg. Biochem*. **2018**, *183*, 179–183.
158. A. H. Follmer, M. Mahomed, D. B. Goodin, T. L. Poulos, *J. Am. Chem. Soc.* **2018**, *140*, 16222–16228.
159. M. Šrejber, V. Navrátilová, M. Paloncýová, V. Bazgier, K. Berka, P. Anzenbacher, M. Otyepka, *J. Inorg. Biochem.* **2018**, *183*, 117–136.
160. G. M. Yee, W. B. Tolman, *Met. Ions. Life Sci.* **2015**, *15*, 131–204.
161. X. Huang, J. T. Groves, *Chem. Rev.* **2018**, *118*, 2491–2553.
162. J. T. Groves, *F1000Research* **2015**, 4 (F1000 Faculty Rev.), 178. DOI: 10.12688/f1000research.6314.1.
163. S. G. Sligar, *Biochemistry* **1976**, *15*, 5399–5406.
164. K. Auclair, P. Moënne-Loccoz, P. R. Ortíz de Montellano, *J. Am. Chem. Soc.* **2001**, *123*, 4877–4885.
165. P. M. Champion, J. D. Lipscomb, E. Münck, P. Debrunner, I. C. Gunsalus, *Biochemistry* **1975**, *14*, 4151–4158.
166. S. Shaik, S. Cohen, Y. Wang, H. Chen, D. Kumar, W. Thiel, *Chem. Rev.* **2010**, *110*, 949–1017.
167. K. L. Bren, R. Eisenberg, H. B. Gray, *Proc. Natl. Acad. Sci. USA* **2015**, *112*, 13123–13127.
168. R. Davydov, T. M. Makris, V. Kofman, D. E. Werst, S. G. Sligar, B. M. Hoffman, *J. Am. Chem. Soc.* **2001**, *123*, 1403–1415.
169. R. Davydov, B. M. Hoffman, *Arch. Biochem. Biophys.* **2011**, *507*, 36–43.
170. J. Rittle, M. T. Green, *Science* **2010**, *330*, 933–937.
171. S. Shaik, D. Kumar, S. P. de Visser, A. Altun, W. Thiel, *Chem. Rev.* **2005**, *105*, 2279–2328.
172. G. H. Loew, D. L. Harris, *Chem. Rev.* **2000**, *100*, 407–419.
173. S. Shaik, M. Filatov, D. Schröder, H. Schwarz, *Chem. Eur. J.* **1998**, *4*, 193–199.
174. T. Kamachi, K. A. Yoshizawa, *J. Am. Chem. Soc.* **2003**, *125*, 4652–4661.
175. J. C. Schöneboom, S. Cohen, H. Lin, S. Shaik, W. Thiel, *J. Am. Chem. Soc.* **2004**, *126*, 4017–4034.
176. M. M. Purdy, L. S. Koo, P. R. Ortiz de Montellano, J. P. Klinman, *Biochemistry* **2004**, *43*, 271–281.
177. R. Lonsdale, J. N. Harvey, A. J. Mulholland, *J. Phys. Chem. Lett.* **2010**, *1*, 3232–3237.
178. R. Lonsdale, J. N. Harvey, A. J. Mulholland, *J. Chem. Theory Comput.* **2012**, *8*, 4637–4645.
179. S. P. de Visser, F. Ogliaro, P. K. Sharma, S. Shaik, *J. Am. Chem.Soc.* **2002**, *124*, 11809–11826.
180. S. P. de Visser, F. Ogliaro, S. Shaik, *Chem. Commun.* **2001**, 2322–2323.
181. S. Yoshioka, S. Takahashi, K. Ishimori, I. Morishima, *J. Inorg. Biochem*. **2000**, *81*, 141–151.
182. R. E. White, J. T. Groves, G. A. McClusky, *Acta Biol. Med. Ger.* **1979**, *38*, 475–482.
183. J. T. Groves, G. A. McClusky, *J. Am. Chem. Soc.* **1976**, *98*, 859–861.
184. J. T. Groves, G. E. Avarianeisser, K. M. Fish, M. Imachi, R. I. Kuczkowski, *J. Am. Chem. Soc.* **1986**, *108*, 3837–3838.
185. P. R. Ortiz de Montellano, *Chem. Rev.* **2010**, *110*, 932–948.
186. A. Sorokin, A. Robert, B. Meunier, *J. Am. Chem. Soc.* **1993**, *115*, 7293–7299.

187. M. Blomberg, T. Borowski, F. Himo, R.-Z. Liao, P. E. M. Siegbahn, *Chem. Rev.* **2014**, *114*, 3601–3658.
188. X. Huang, J. T. Groves, *J. Biol. Inorg. Chem.* **2017**, *22*, 185–207.
189. V. Schünemann, F. Lendzian, C. Jung, J. Contzen, A.-L. Barra, S. G. Sligar, A. X. Trautwein, *J. Biol. Chem.* **2004**, *279*, 10919–10930.
190. M. Newcomb, R. Zhang, R. E. P. Chandrasena, J. A. Halgrimson, J. H. Horner, T. M. Makris, S. G. Sligar, *J. Am. Chem. Soc.* **2006**, *128* , 4580–458.
191. C. Jung, *Biochim. Biophys. Acta* **2011**, *1814*, 46–57.
192. M. M. Palcic, R. Rutter, T. Araiso, L. P. Hager, H. B. Dunford, *Biochem. Biophys. Res. Commun.* **1980**, *94*, 1123–1127.
193. T. Egawa, D. A. Proshlyakov, H. Miki, R. Makino,T. Ogura,T. Kitagawa,Y. Ishimura, *J. Biol. Inorg. Chem.* **2001**, *6*, 46–54.
194. R. Rutter, L. P. Hager, H. Dhonau, M. Hendrich, M.Valentine, P. Debrunner, *Biochemistry* **1984**, *23*, 6809–6816.
195. B. M. Hoffman, *Proc. Natl. Acad. Sci. USA* **2003**, *100*, 3575–3578.
196. J. Antony, M. Grodzicki, A. X. Trautwein, *J. Phys.Chem. A* **1997**, *101*, 2692–2701.
197. P. J. Mak, I. G. Denisov, *Biochim. Biophys. Acta* **2018**, *1866*, 178–204.
198. S. Rendic, F. P. Guengerich, *Chem. Res. Toxicol.* **2015**, *28*, 38–42.
199. C. M. Casadei, A. Gumiero, C. L. Metcalfe, E. J. Murphy, J. Basran, M. G. Concilio, S. C. M. Teixeira, T. E. Schrader, A. J. Fielding, A. Ostermann, M. P. Blakeley, E. L. Raven, P. C. E. Moody, *Science* **2014**, *345*, 193–197.
200. J. T. Groves, N. C. Boaz, *Science* **2014**, *345*, 142–143.
201. M. Klingenberg, *Arch. Biochem. Biophys.* **1958**, *75*, 376–386.
202. F. P. Guengerich, *Chemistry & Biology* **2009**, *16*, 1215–1216.
203. F. P. Guengerich, *ACS Catal.* **2018**, *8*, 10964–10976.
204. F. P. Guengerich, F. K. Yoshimoto, *Chem. Rev.* **2018**, *118*, 6573–6655.
205. A. B. McQuarters, M. W. Wolf, A. P. Hunt, N. Lehnert, *Angew. Chem. Int. Ed.* **2014**, *53*, 4750–4752.
206. I. G. Denisov, S. G. Sligar, *Chem. Rev.* **2017**, *117*, 4669–4713.
207. V. V. Shumyantseva, A. V. Kuzikov, R. A. Masamrekh, T. V. Bulko, A. I. Archakov, *Biosensors and Bioelectronics* **2018**, *121*, 192–204.
208. P. Quantin, E. Colaço, K. El Kirat, C. Egles, H. Ficheux, J. Landoulsi, *ACS Omega* **2018**, *3*, 12535–12544.
209. C. H. Hsieh, T. M. Makris, *Biochem. Biophys. Res. Commun.* **2016**, *476*, 462–466.
210. C. H. Hsieh, X. Huang, J. A. Amaya, C. D. Rutland, C. L. Keys, J. T. Groves, R. N. Austin, T. M. Makris, *Biochemistry* **2017**, *56*, 3347–3357.
211. J. A. Amaya, C. D. Rutland, N. Leschinsky, T. M. Makris, *Biochemistry* **2018**, *57*, 344–353.
212. C. E. Wise, C. H. Hsieh, N. L. Poplin, T. M. Makris, *ACS Catal.* **2018**, *8*, 9342–9352.
213. F. P. Guengerich, A. W. Munro, *J. Biol. Chem.* **2013**, *288*, 17065–17073.
214. X. Wang, S. Peter, R. Ullrich, M. Hofrichter, J. T. Groves, *Angew. Chem. Int. Ed.* **2013**, *52*, 9238–9241.
215. O. Shoji, S. Yanagisawa, J. K. Stanfield, K. Suzuki, Z. Cong, H. Sugimoto,Y. Shiro, Y. Watanabe, *Angew. Chem. Int. Ed.* **2017**, *56*, 10324–10329.
216. W. Tian, C. Sun, M. Zheng, J. R. Harmer, M. Yu, Y. Zhang, H. Peng, D. Zhu, Z. Deng, S.-L. Chen, M. Mobli, X. Jia, X. Qu, *Nat. Commun.* **2018**, *9*, 4428; DOI: 10.1038/s41467-018-06528-z.
217. J. Kim, P.-g. Lee, E.-o. Jung, B.-G. Kim, *Biochim. Biophys. Acta* **2018**, *1866*, 60–67.
218. K. Dutta Dubey, S. Shaik, *Acc. Chem. Res.* **2019**, *52*, 389–399.

7

Basic Iron-Sulfur Centers

Claudia Andreini and Simone Ciofi-Baffoni

Magnetic Resonance Center, University of Florence,
Via Luigi Sacconi 6, I-50019 Sesto Fiorentino, Italy
<andreini@cerm.unifi.it>
<ciofi@cerm.unifi.it>

Metal Ions in Life Sciences, Volume 20 Guest Editors: Martha E. Sosa Torres and Peter M. H. Kroneck
Series Editors: Astrid Sigel, Eva Freisinger, and Roland K. O. Sigel

https://doi.org/10.1515/9783110589757-007

Abstract: Iron-sulfur clusters are ubiquitous protein cofactors composed of iron and inorganic sulfur. These cofactors are among the most ancient ones and may have contributed to the birth of life on Earth. Therefore, they are found even today in many enzymes central to metabolic processes like nitrogen fixation, respiration, and DNA processing and repair. Due to the toxicity associated with iron and sulfur ions, living organisms evolved dedicated machineries to synthetize and then transfer iron-sulfur clusters into client proteins. The iron-sulfur cluster (ISC) machinery is responsible for iron-sulfur cluster biogenesis in prokaryotes and in the mitochondrion of eukaryotes; the sulfur mobilization (SUF) machinery is present in prokaryotes and in the chloroplasts of plants; finally, the cytosolic iron-sulfur assembly (CIA) machinery is only present in the cytoplasm of eukaryotes. Genome analysis allowed the prediction of the proteins containing iron-sulfur clusters across a broad variety of living organisms, establishing links between the size and composition of iron-sulfur proteomes and the types of organisms that encode them. For example, the iron-sulfur proteomes of aerobes are generally smaller than those of anaerobes with similar genome size; furthermore, aerobes are enriched in [2Fe-2S] proteins compared to anaerobes, which predominantly use [4Fe-4S] proteins. This relates to the lower bioavailability of iron and the higher lability of [4Fe-4S] clusters within aerobic environments. Analogous considerations apply to humans, where the occurrence and functions of iron-sulfur proteins depend on the cellular compartment where they are localized. For example, an emerging primary role for nuclear iron-sulfur proteins is in DNA maintenance. Given their key functions in metabolism, dysfunctions of mutations in iron-sulfur proteins, or in proteins participating in iron-sulfur cluster biogenesis, are associated with serious human diseases.

Keywords: [2Fe-2S] · [3Fe-4S] · [4Fe-4S] · Fe-S biogenesis · iron-sulfur · iron-sulfur proteomes

1. INTRODUCTION

1.1. Types of Basic Iron-Sulfur Centers

Iron-sulfur (Fe-S) clusters are metal cofactors composed of iron and inorganic sulfide ions, which are ubiquitous and essential components of living cells [1, 2]. Classical and most common types of biological Fe-S clusters include from two to four iron ions (either in the $+2$ or in the $+3$ oxidation state), bridged by

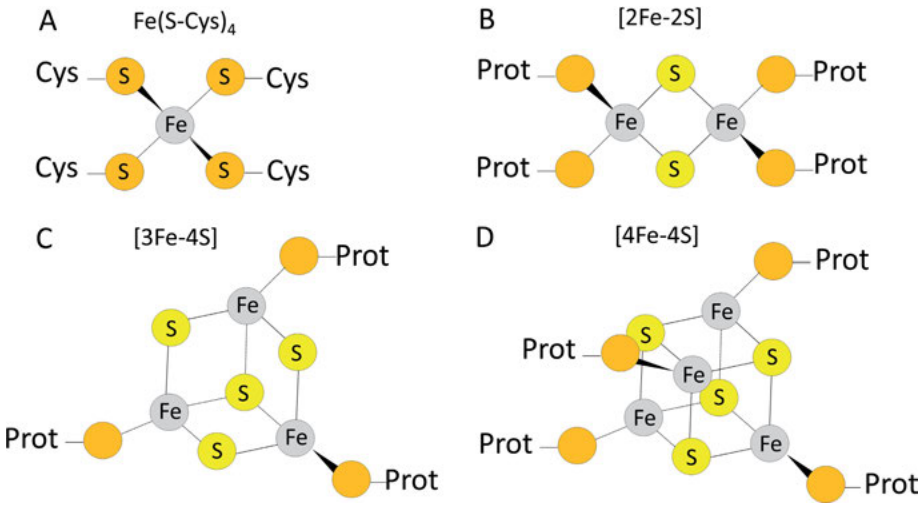


Figure 1. Basic types of Fe-S clusters in proteins.

sulfide ions (Figure 1). In all these clusters, iron ions have a tetrahedral geometry and anchor the cluster to the protein matrix by coordinating protein residues. In most cases, they are coordinated by thiolate from cysteine [3], although there are increasing examples of nitrogen coordination provided by histidine and arginine residues, and of oxygen coordination from aspartate, tyrosine, and serine [4–6]. It is customary to include in the class of Fe-S sites also the $Fe(S\text{-}Cys)_4$ sites of rubredoxins (Figure 1A), which contain just one iron ion coordinated by four cysteine residues [7]. The simplest Fe-S cluster, also called "diamond", is the [2Fe-2S] cluster and includes two iron ions and two bridged inorganic sulfides on the same plane (Figure 1B). The most common Fe-S cofactor in nature is the [4Fe-4S] cluster, which includes four iron ions bridged by four sulfide ions (Figure 1D). The cluster is structured as a distorted cube and is also called "cubane". The [3Fe-4S] cluster (Figure 1C), much less common, can be viewed as a cubane cluster missing an iron ion and the corresponding protein ligand [7].

Fe-S clusters are among the most ancient cofactors in nature and may have been one of the earliest types of protein prosthetic groups in widespread use. When early forms of life began to evolve, the atmosphere of the Earth was anaerobic and reducing: iron was probably highly abundant and soluble [8]. Geothermal vents may have led to high concentration of hydrogen sulfide in some areas. In these conditions, it is plausible that Fe-S clusters have acted as catalysts on the primordial Earth [9]; indeed, simple Fe-S clusters can spontaneously assemble in the presence of high concentrations of Fe^{2+} and S^{2-} in aqueous solutions [10], under reducing conditions. So, Fe-S clusters may have contributed to the transition from the abiotic to the biotic world, explaining why Fe-S clusters are found as cofactors in many enzymes central to the metabolism, including enzymes involved in the Krebs cycle, nitrogen fixation, ribosome assembly, and DNA repair. Nowadays, Fe-S clusters serve multiple purposes including electron transfer, catalysis, protein structural stabilization, iron storage and sensing, and

therefore many proteins use them to carry out their physiological function. Presently, known Fe-S proteins range in size from 6 kDa to over 500 kDa, and contain up to nine Fe-S clusters [6]. They are present in all kinds of cells and cellular compartments and are widespread among organisms. While Fe-S clusters are intrinsically oxygen-sensitive, their stability in proteins largely depends on the polypeptide matrix [9, 11]. The coordination environment of the cluster and the protein matrix around the cluster also serve to modulate the redox potential of different types of Fe-S clusters, which may range from –700 to +300 mV [6, 12].

1.2. Basic Iron-Sulfur Centers in Structures

As reported in Figure 2, the number of Fe-S structures released increases every year and has accelerated rapidly in recent years [13, 14]. This is because these systems are increasingly attracting the attention of scientists, due to the importance they had in evolution and the roles they play for the survival of organisms. At the beginning of 2019 the number of Fe-S binding structures reached almost 1500 units, representing about 10 % of the structures in the Protein Data Bank (PDB) [15]. By considering only those structures with less than 30 % sequence identity, this number reduces to 329 proteins that still represent about 10 % of the non-redundant PDB. About 20 % of these are enzymes and thus have an associated EC number [16]. Most of them are oxidoreductases that use the Fe-S cluster either to create an electron transfer path providing electrons to the active site [17], or to directly bind and activate molecules such as in radical *S*-adenosylmethionine (SAM) enzymes [18].

The most common cluster bound to PDB structures is the [4Fe-4S] type, which occurs in 74 % of sites [13, 14]. Among classical [4Fe-4S] folds, the first discovered [19] and probably the most widespread, is 2 × [4Fe-4S] ferredoxin [6], which possesses a βαββαβ topology [20]. This fold is associated to ancient electron carrier proteins mainly involved in energy production processes, like respiration and photosynthetic reactions. Fe-S clusters in 2[4Fe-4S] ferredoxin undergo a $[4Fe\text{-}4S]^{2+,+}$ redox transition and have a low redox potential (–150 to –700 mV) [6]. Conversely, high potential iron proteins (HiPIP) bind a [4Fe-4S] cluster which implements a $[4Fe\text{-}4S]^{3+,2+}$ transition with a redox potential ranging between +100 and +400 mV [7]. HiPIP are globular proteins nearly devoid of secondary structure, which bind the [4Fe-4S] cluster in a buried cavity [21]. These small proteins are both found in photosynthetic bacteria [22], where they behave as electron donors to a tetra-heme cytochrome, and in non-photosynthetic prokaryotes. The function of HiPIP in these last organisms remains unclear. Structures coordinate [4Fe–4S] clusters almost exclusively through thiols provided by cysteine residues, with an average coordination distance of 2.29 Å ± 0.07 Å [13, 14]. The only two relevant exceptions are 4-hydroxybutyryl-CoA dehydratase from *Clostridium aminobutyricum* (PDB code 1U8V), which uses a $[4Fe\text{-}4S]^{2+}$ cluster coordinated by three cysteine and one histidine to catalyze the reaction, and lipoyl synthase from *Thermosynechococcus elongates* and *Mycobacterium tuberculosis*, which is a member of the radical SAM enzyme superfamily, which binds two [4Fe-4S] clusters, one of this being coordinated by three cysteines and

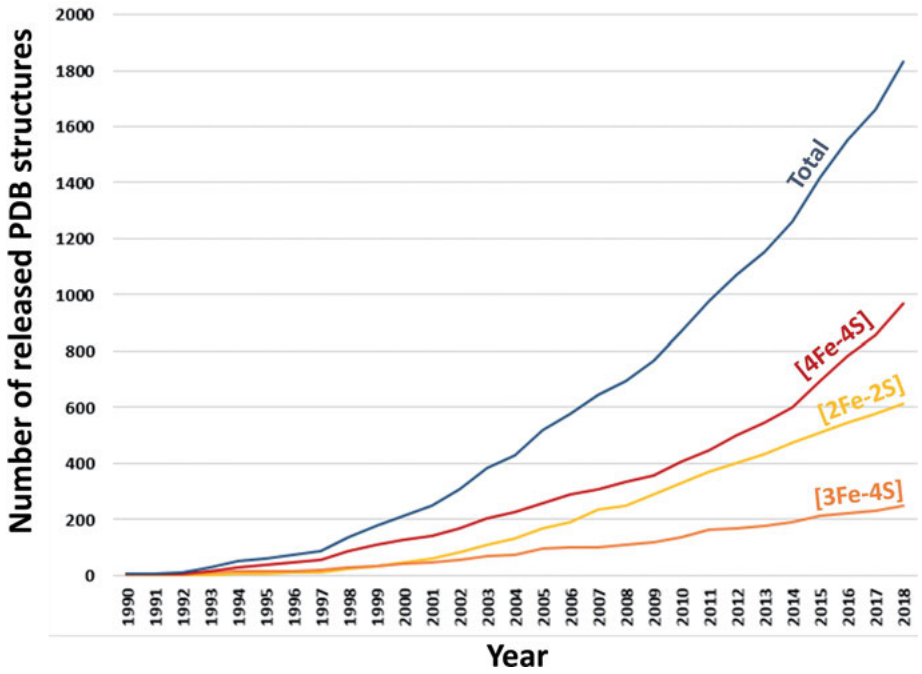


Figure 2. Number of available Fe-S structures in the PDB per year (divided per Fe-S cluster type).

one serine and used to insert two sulfur atoms into the relatively unreactive C–H bonds of an octanoyl substrate [4, 23].

The second most common Fe-S cluster type in PDB structures is the [2Fe-2S] cluster, occurring in 21 % of sites [13, 14]. Among classical [2Fe-2S] folds, the most widespread is [2Fe-2S] ferredoxin [24], whose structure is characterized by four β-sheets covered by an α-helix. This fold is associated with globular and small proteins mainly performing electron transfer [25, 26]. The Fe-S cluster undergoes a $[2Fe\text{-}2S]^{2+,+}$ redox transition and has a low redox potential (–150 to –450 mV) [10]. Other low potential [2Fe-2S] proteins (about –300 mV) are ferredoxins with a thioredoxin-like fold; these are less widespread because they are only found in bacteria. Another common [2Fe-2S] fold is that associated with Rieske proteins, which have the feature of coordinating the Fe-S cluster with two cysteines and two histidines. The redox transition in Rieske proteins is $[2Fe\text{-}2S]^{2+,+}$, like in ferredoxins, however the histidine ligation causes an upshift of the potential; as a consequence, these proteins generally have a positive redox potential (+100 to +400 mV) [10]. The fold of the cluster-containing subdomain of Rieske proteins is strikingly similar to that of rubredoxins [27].

In agreement with fold occurrence, cysteine is by far the most common ligand for [2Fe-2S] clusters (average coordination distance 2.32 Å ± 0.07 Å) followed by histidine, which almost always coordinates Fe-S clusters with the Nδ of the imidazole ring (average coordination distance 2.15 Å ± 0.09 Å [13, 14]). In a

few cases, other ligands are possible for this type of cluster: for instance, one arginine is found in the [2Fe-2S] site of biotin synthase of *Escherichia coli* (PDB code 1R30 [28]).

The [3Fe-4S] sites generally perform electron transfer functions. In the PDB, these are always anchored to the protein structure by cysteine ligation, with an average coordination distance of 2.34 Å $\pm$ 0.09 Å [13, 14].

2. IRON-SULFUR PROTEIN BIOGENESIS

Although Fe-S clusters can be synthesized *in vitro* and transferred to a client protein without enzymatic assistance, toxicity of free iron and sulfide ions has led organisms to evolve intricate mechanisms to assemble and transfer Fe-S clusters within cells. Based on biochemical evidences and organization of genes in bacterial operons, three different pathways of Fe-S proteins biosynthetic systems have been discovered in bacteria, i.e., the iron-sulfur cluster (ISC) machinery (Figure 3A), the sulfur mobilization (SUF) machinery (Figure 3B), and the nitrogen fixation (NIF) machinery [29–31]. All of them include two main steps: (i) the *de novo* Fe-S cluster biosynthesis, and (ii) the transfer and insertion of the cluster to the client protein(s) by a transfer system. In all known bacterial machineries, the *de novo* Fe-S cluster biosynthesis requires two players: first, a cysteine desulfurase (IscS, SufSE, NifS) [32–34] which uses L-cysteine as a stable and safe source of sulfide ions; second, a scaffold protein (IscU, SufB, NifU) [35–37] which combines sulfide and iron ions to assemble clusters. Genomic analysis revealed that the number and type of operons coding for these machineries vary from one bacterial organism to another.

Eukaryotic mitochondria have adopted the ISC machinery from bacteria by integrating components of the NIF pathway, and the ISC machinery works for maturing mitochondrial [2Fe-2S] and [4Fe-4S] client proteins. The eukaryotic ISC machinery also provides an essential precursor to a cytosolic pathway for Fe-S protein biogenesis, unique to eukaryotes, namely the cytosolic iron-sulfur protein assembly (CIA) machinery. Plants also conserve the SUF machinery for Fe-S protein biogenesis in chloroplasts.

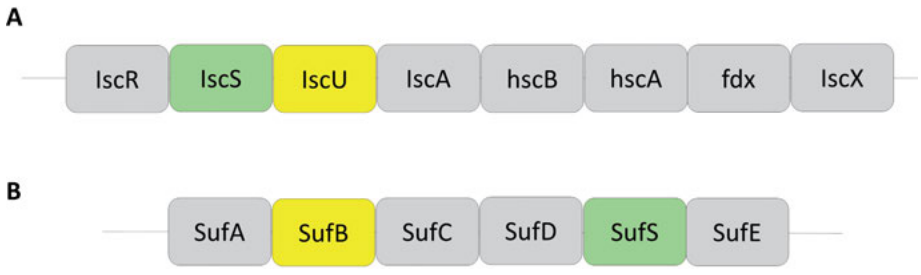


Figure 3. General genetic organization of ISC and SUF machineries into bacterial operons; scaffold proteins are in yellow while cysteine desulfurases are in green.

In the next sections we will discuss in detail the ISC, SUF, and CIA machineries while the NIF system, which is responsible for the maturation of nitrogenase in azototrophic bacteria [29–38], will not be discussed in this chapter.

2.1. The Iron-Sulfur Cluster Machinery in Bacteria

2.1.1. de novo Fe-S Cluster Biosynthesis in Bacteria

The ISC machinery is the primary system for Fe-S protein biogenesis in bacteria (Figure 3A) [35]. The scaffold IscU protein requires three ingredients to synthesize both [2Fe-2S] and [4Fe-4S] clusters, i.e., sulfide, electrons, and iron ions (Figure 4A, see below) [37, 39–40]. The cluster assembly follows a sequential process: first, a $[2Fe\text{-}2S]^{2+}$ cluster is generated by the three ingredients on a dimeric IscU, then this form of IscU is converted to a dimeric IscU containing two $[2Fe\text{-}2S]^{2+}$ clusters, and finally a reductive coupling of the two $[2Fe\text{-}2S]^{2+}$ clusters occurs on the homodimeric IscU, to form a single $[4Fe\text{-}4S]^{2+}$ cluster [39, 40].

It has been shown that the first ingredient is provided by cysteine desulfurase (IscS) [31, 41], a pyridoxal phosphate-dependent enzyme, which reduces a cysteine to alanine generating persulfide sulfur; this is then transferred to a cysteine on the IscU scaffold [42]. The second ingredient is provided by the [2Fe-2S] ferredoxin, an electron donor protein [43], which reduces the persulfide sulfur to sulfur ions [44]. Reduced ferredoxin was also suggested to provide electrons required to obtain the reductive coupling of two $[2Fe\text{-}2S]^{2+}$ clusters on IscU to form a $[4Fe\text{-}4S]^{2+}$ cluster. The final ingredient, iron ions, is toxic for cells so it may be provided by a specific iron donor. At the moment, the identity of this player is still unknown even though several proteins have been proposed for this role [31]. The first is bacterial frataxin (CyaY) which does not belong to the ISC operon. This protein was initially suggested to be an iron donor due to its *in vitro* behavior toward iron. However, there are contradictory experimental data on the possibility that the protein actually releases iron for cluster biosynthesis [45, 46]. The possibility that CyaY might act as a negative regulator of Fe-S formation, rather than as an iron donor, has recently emerged both in prokaryotes and in eukaryotes [47]. Furthermore, recent works in eukaryotes also suggested that it stimulates sulfide transfer from cysteine desulfurase to IscU [48, 49]. A second protein that was initially suggested as the putative iron donor is IscX, a small protein in the ISC operon whose function is still unclear. Like CyaY, the deletion of the IscX gene does not lead to any specific phenotype related to the Fe-S metabolism [50]. Biochemical studies indicate that IscX inhibits the Fe-S cluster formation on the IscU scaffold without affecting the cysteine desulfurase activity of IscS [51]. So, like CyaY, it is more plausible that the protein behaves as a molecular adaptor which can modulate Fe-S cluster formation.

Structural and functional analysis demonstrated that IscS and IscU form a complex where each IscS in a dimer interacts with an IscU protein [42]. Further structural studies revealed that each IscS in the complex has an exposed second binding site (an arginine-rich region) which is able to bind one among CyaY,

ferredoxin (Fdx), and IscX [51, 52]. Probably, during Fe-S cluster synthesis different ternary complexes are formed, but it is still unclear which of the complexes, and in what order, are used; in a plausible model, IscS:IscU:CyaY is the first complex generated. In this system, CyaY would regulate iron entry, whereas IscS would produce persulfide which would be transferred to IscU through a transpersulfuration reaction. Then, Fdx would displace CyaY and reduces persulfide on IscU, allowing Fe-S biosynthesis [31].

2.1.2. Transfer of the Fe-S Cluster to Client Protein(s) in Bacteria

HscA and HscB (Figure 4B), constitute the ATP-dependent chaperone system which interacts with IscU and facilitates Fe-S cluster transfer to the client protein [53–55]. Interestingly, the HscA-HscB system is able to enhance the rate of [2Fe-2S] cluster transfer from the IscU dimer but not the rate of [4Fe-4S] cluster transfer. At the moment, two possible mechanisms have been proposed for Fe-S cluster transfer, both involving the formation of a ternary complex HscA(ATP-bound)-HscB-IscU(dimer). The hydrolysis of ATP to ADP induces a conformational change in HscA, which promotes the dissociation of the HscB subunit (step 1); then, also ADP is released (step 2) and replaced by a new ATP molecule (step 3); this induces the release of the IscU-dimer (step 4); finally, the HscA-(ATP-bound) protein is able to bind a HscB-IscU-dimer complex to recreate the ternary adduct (step 5).

The two proposed mechanisms differ in the step of the [2Fe-2S] cluster release. In the first model, the transfer of the Fe-S cluster is coupled to ATP hydrolysis by HscA (step 1), whereas in the second model the cluster is transferred to its target after the release of the IscU-dimer (step 4) [31]. It is plausible that the proposed mechanisms are largely correct, however, current knowledge on the protein–protein interactions supports the idea that the Fe-S transfer process is much more complex. Furthermore, *in vitro* and *in vivo* experiments have demonstrated that Fe-S transfer from IscU to client proteins occurs either directly or via Fe-S carrier proteins (i.e., IscA in the case of ISC system) [56, 57]. IscA belongs to the A-type family [56] of ISC proteins that have been shown to be either homo- or heterodimers, which can coordinate Fe-S clusters (either [2Fe-2Fe] or [4Fe-4S]) and/or mononuclear iron via three conserved cysteine residues [58–61]. Recent studies suggested that IscA could specifically fulfill the role of [4Fe-4S] protein maturation after reductive coupling of its $[2Fe\text{-}2S]^{2+}$ cluster to form $[4Fe\text{-}4S]^{2+}$ clusters on IscA itself [62, 63]. Furthermore, under aerobic conditions IscA may act as an iron chaperone to provide iron for [4Fe-4S] cluster assembly [60]. Fdx interacts with IscA and could be involved in providing electrons for the [2Fe-2S] reduction. Other proteins not present in the ISC operon, such as monothiol glutaredoxin (Glrx), NfuA, and ErpA, have been proposed to have roles in shuttling Fe-S clusters to client proteins [56, 58, 64–68].

The last protein within the ISC operon [69] (Figure 3A) is IscR, a [2Fe-2S]-binding transcriptional regulator which controls the expression of more than 40 genes in 20 predicted operons of *E. coli*; from transcriptomic studies, IscR appears to have roles in the regulation of Fe-S protein biogenesis as well as in

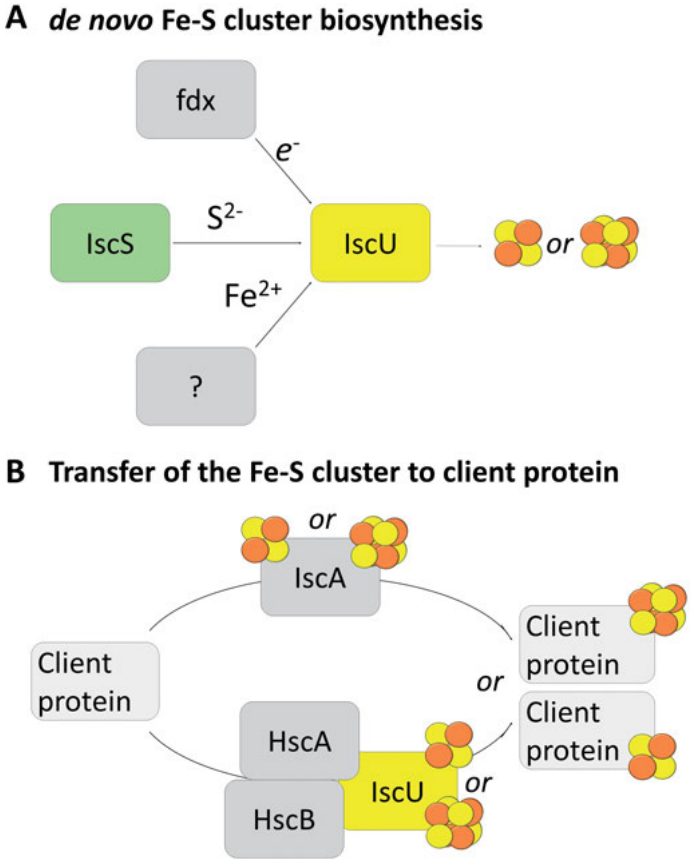


Figure 4. Schematic representation of the two major steps of the ISC machinery in prokaryotes: (**A**) the *de novo* Fe-S cluster biosynthesis and (**B**) the transfer of the cluster to the client protein. This transfer may occur directly from the IscU with the help of chaperones HscA/HscB or indirectly through carrier proteins like IscA.

the O_2 regulation of several promoters controlling the expression of anaerobic Fe-S proteins [70].

2.2. The Iron-Sulfur Cluster Machinery in Eukaryotes

In eukaryotes, the ISC machinery was essentially inherited from the bacterial endosymbiont during evolution and is thus located in the mitochondria [71]. The biogenesis process is schematized in Figure 5 and consists of four steps: (1) *de novo* [2Fe-2S] cluster synthesis by the ISU complex; (2) transfer of the [2Fe-2S] cluster from the ISU complex to the monothiol glutaredoxin-5 (Glrx5) a chaperone system. The [2Fe-2S] cluster may then follow two different paths: it may be delivered to client mitochondrial proteins, or may be transferred to the

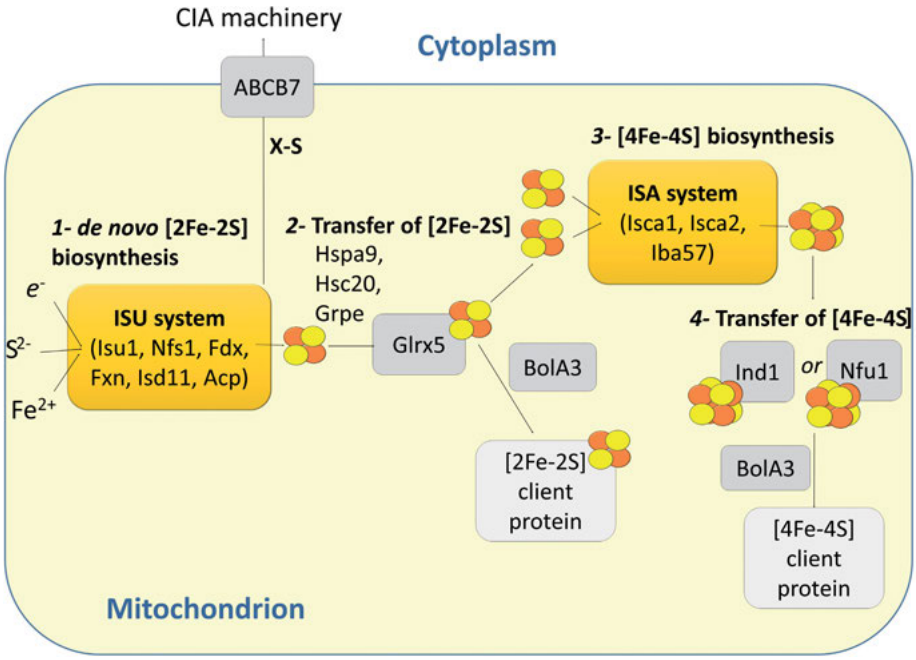


Figure 5. Schematic representation of the ISC machinery in the mitochondrion of eukaryotes.

ISA system in order to (3) synthesize a [4Fe-4S] cluster. In the latter case (4), the newly generated [4Fe-4S] cluster is then transferred and inserted to mitochondrial client proteins with the help of mitochondrial ISC targeting factors, i.e., Ind1 and Nfu1 [72, 73]. A third pathway delivers an S-containing product generated on the ISU complex (whose identity remains unknown) to the Fe-S protein biogenesis pathway taking place in the cytoplasm (CIA machinery) [74]. The export of the product from the mitochondrion to the cytoplasm is mediated by the ABCB7 protein (Atm1 in yeast), an inner mitochondrial membrane protein (Figure 5) [75, 76]. Another crucial role played by the mitochondrial ISC machinery is the regulation of cellular iron homeostasis. Eukaryotic cells apparently use the efficiency of the mitochondrial Fe-S assembly as an important sensor of the intracellular iron status.

2.2.1. de novo [2Fe-2S] Cluster Biosynthesis

This step is very similar to bacterial *de novo* Fe-S cluster biosynthesis. The scaffold protein (Isu1) [37] receives sulfide ions from cysteine desulfurase (Nfs1), iron ions from an unknown source and electrons from ferredoxin. Transfer of the persulfate sulfur appears to be activated by frataxin (Fxn, or Yfh1 in yeast) [47, 72]. The Fe-S cluster synthesis requires the formation of the ISU complex, which is similar to the bacterial system: eukaryotic Nfs1 forms a complex with

the scaffold protein Isu1 and frataxin. However, eukaryotes have two additional proteins in the adduct, i.e., acyl carrier protein (ACP) [77] and Isd11 [78, 79]. Both these proteins stably interact with Nfs1; Isd11 contributes to persulfide formation, while ACP stabilizes the sub-complex between Nfs1 and Isd11 [80].

2.2.2. Transfer of the [2Fe-2S] Cluster

A second multimeric protein complex is responsible for facilitating the release of the [2Fe-2S] cluster from the ISU complex to Glrx5 [81–83]. In this complex, which is similar to the bacterial adduct, the chaperone Hspa9 (HscA is the prokaryotic homolog), and its co-chaperone Hsc20 (HscB is the prokaryotic homolog) work together in targeting the Isu1 scaffold for specific release of the bound [2Fe-2S] cluster to Glrx5 [82–85]. The process is dependent on ATP hydrolysis and the Grpe protein facilitates the exchange of ADP to ATP in the complex [83]. The association of Glrx5 and Hspa9 with the Hsc20-Isu1 heterodimer stimulates ATP hydrolysis, triggering a conformational change in Hspa9 that releases Hsc20 and the Fe-S cluster to Glrx5. Glrx5 is then involved both in transferring the [2Fe-2S] cluster to the client proteins and in presenting the two [2Fe-2S] clusters to the ISA system for [4Fe-4S] synthesis. Recent studies support the hypothesis that also BolA1 and BolA3, members of the BOLA family, may be involved in Fe-S cluster transfer by forming a complex with Glrx5, which is able to receive a [2Fe-2S] cluster from Isu1 [86–88]. How the Glrx5-BolA1/BolA3 and Glrx5-Glrx5 dimers interplay with each other is still unresolved.

2.2.3. Synthesis of [4Fe-4S] Cluster

Three proteins are required for the synthesis of [4Fe-4S] clusters in the cell, Isca1, Isca2 (two homologs of the IscA prokaryotic protein), and Iba57 [89, 90]. It was proposed that these three proteins form a complex for the synthesis of a [4Fe-4S] cluster [90]. However, the structure of the complex as well as the stoichiometry and the Fe-S sites, are still unknown. Moreover, it has been found *in vitro* that only Isca1 and Isca2 are required to form a [4Fe-4S] cluster, without the need of Iba57 [91–93]. Isca1 and Isca2 form, indeed, a heterodimeric complex that accepts two [2Fe-2S] clusters from Glrx5 synthesizing a [4Fe-4S] cluster [92]. Depending on the oxidation states of the cluster, $[2Fe\text{-}2S]^{1+}$ or $[2Fe\text{-}2S]^{2+}$, a biological electron source may or may not be required for $[4Fe\text{-}4S]^{2+}$ cluster synthesis, respectively. The role of Iba57 in the biosynthesis of [4Fe-4S] clusters is still obscure and its clarification represents a challenge in this research area. Recently, it has been shown that human Iba57 forms a heterodimeric complex with Isca2 by bridging a [2Fe-2S] cluster, and that [2Fe-2S] cluster binding is absolutely required to promote the complex formation [94]. This complex is formed when Iba57 is either exposed to [2Fe-2S] Isca2 or in the presence of [2Fe-2S] Glrx5 and apo Isca2. The stability of the [2Fe-2S] cluster against oxida-

tive degradation, observed upon its binding to the Isca2-Iba57 complex with respect to when it is bound to Isca2 alone, supports a model where the heterocomplex can work under an oxidative metabolism. These findings delineate a new pathway determinining the formation of the [2Fe-2S] Isca2-Iba57 complex, which might work as a specific system maturing [4Fe-4S] enzymes under aerobic cellular conditions. According to this, it has been observed that this complex is able to mature aconitase, which contains an oxygen-labile [4Fe-4S] cluster [94].

2.2.4. Transfer of the [4Fe-4S] Cluster to Client Proteins

The [4Fe-4S] cluster synthesized in the ISA system is finally transferred to a client mitochondrial protein by mitochondrial ISC targeting factors. At present, two of these proteins are known, Nfu1 [95, 96] and Ind1 [97, 98], which both bind a [4Fe-4S] cluster. A [4Fe-4S] Nfu1-homodimer was proposed to mediate the transfer of a [4Fe-4S] cluster from the ISA system to client proteins such as aconitase, lipoyl synthase, and the Fe-S subunit of succinate dehydrogenase. The second mitochondrial ISC targeting factor is Ind1 which was involved in the maturation of the respiratory complex I. Two further proteins, BolA1 and BolA3, which are members of the BOLA family, were proposed to have a role in the transfer of [4Fe-4S] clusters to final client proteins. While the specific role of BolA1 in the ISC machinery is still obscure, the BolA3 protein was co-immunoprecipitated with Isca2 and [4Fe-4S] client proteins, a protein set overlapping with the Nfu1 interactome. These data as well as the similar phenotypes of Nfu1 and BolA3 and the dramatic increase of BolA3 levels upon overexpression of Nfu1 further support the role of these proteins in the late stage of the cluster transfer. Nevertheless, the specific role of BolA3 in the maturation of Fe-S proteins remains unresolved [86, 95]. The role of BolA3 in the ISC machinery is also complicated by the fact that BolA3 forms a stable [2Fe-2S] complex with Glrx5, the latter operating in the ISC machinery upstream of the Nfu1 protein [87].

2.3. The Sulfur Mobilization Machinery

2.3.1. de novo Fe-S Cluster Biosynthesis

In many bacteria, the SUF machinery is able to maturate Fe-S proteins similarly to the ISC machinery, however, it is only operative in conditions of iron limitations and oxidative stress [99]. Not surprisingly, the SUF machinery is conserved in plant chloroplasts, an organelle producing O_2 [100]. Among eukaryotes, SUF is present only in photosynthetic eukaryotes. The SUF system also appears to be the sole system for Fe-S protein biogenesis in archaea, cyanobacteria, and many Gram-positive bacteria [101]. The SUF operon contains from two (*SufB*, *SufC*) to seven genes (*SufA*, *SufB*, *SufC*, *SufD*, *SufS*, *SufE*, *SufU*) [102] demonstrating that oxygen-exposed environments require more complex Fe-S biogenesis systems, thus driving the evolution of SUF in recruiting new players.

Following the same theme of other machineries (Figure 6), the SUF system assembles new Fe-S clusters on scaffold proteins by using iron and sulfur ions

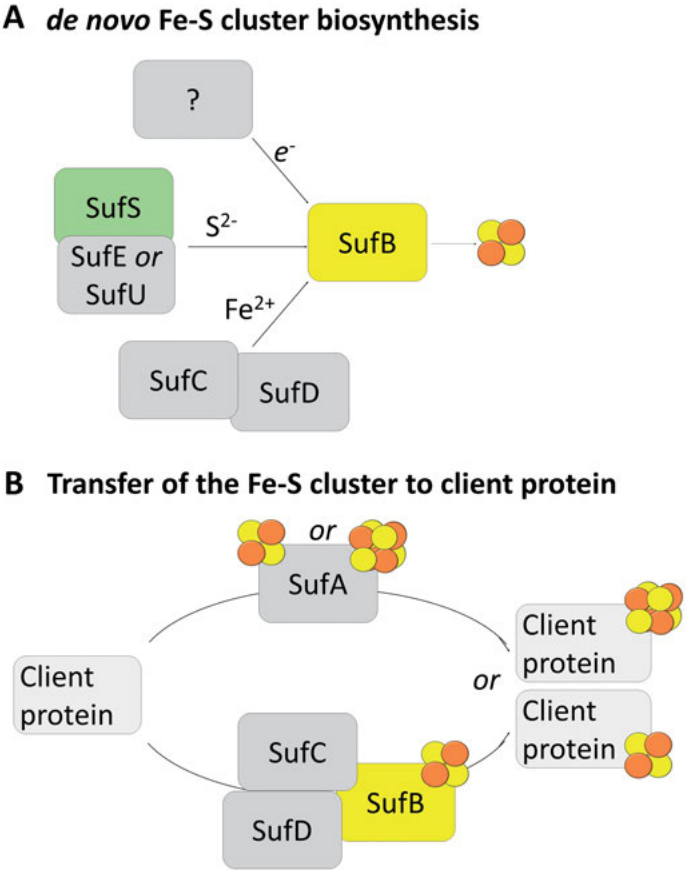


Figure 6. Schematic representation of the two major steps of the SUF machinery: (**A**) the *de novo* Fe-S cluster biosynthesis and (**B**) the transfer of the cluster to the client protein. This transfer may occur directly from a SufBCD complex or indirectly through carrier proteins like SufA. This latter protein may be also involved in the maturation of the [4Fe-4S] proteins.

released by other specialized proteins. In particular, SufB is the scaffold protein [103] of the system and it was demonstrated that a [2Fe-2S] cluster bound to SufB is more stable and resistant to oxidative agents than a cluster bound to IscU [104]. *In vivo* experiments showed that SufB binds a [2Fe-2S] cluster, reinforcing the idea that SufB might be a [2Fe-2S] protein rather than a [4Fe-4S] protein [105]. Biochemical and structural studies have demonstrated that SufB is able to interact with SufC and SufD to form $SufB_2C_2$ and $SufBC_2D$ complexes [30, 106–108]. The role of SufC and SufD in the complex is likely to provide the iron ions necessary for cluster synthesis. Indeed, it was reported that SufC possesses an ATPase activity required for *in vivo* iron acquisition, while deletion of the *SufD* gene decreases the iron content of the $SufB_2C_2$ complex. SufD is a paralog of SufB, but it is unable to behave as a scaffold for Fe-S assembly [101].

Sulfur ions are provided by a dimeric cysteine desulfurase SufS, which extracts a persulfide intermediate from a cysteine substrate. Despite its similarity with IscS, SufS has a lower basal activity, possibly due to the lower length of the loop containing the functional Cys364, which thus has a lower flexibility. The activity of SufS is largely enhanced by its association with SufE (in a 1:1 stoichiometry) [109], which induces a conformational change in SufS that facilitates the desulfuration reaction. It was shown that the SufSE complex is able to interact with the $SufBC_2D$ complex [103], further enhancing the cysteine desulfurase activity. Similarly to SufE, also SufU is able to largely increase the activity of SufS [110], and therefore cooperates with cysteine desulfurase in providing sulfur atoms to the scaffold system. SufU is largely similar to IscU in sequence, however the hypothesis that it may function as an alternative protein scaffold is still debated [110, 111].

2.3.2. Transfer of the Cluster to Client Proteins

SufA is a member of the A-type carrier family (also including IscA and ErpA) of proteins involved in Fe-S cluster transfer to client proteins [56]. Like IscA, it is able to bind both the [2Fe-2S] cluster and the [4Fe-4S] cluster. SufA has the ability of transferring its cluster to downstream apo-proteins such as biotin synthase, aconitase, and ferredoxin [102, 112]. Also the Fe-S cluster holo-form of SufB (in complex with SufC and SufD) is competent for cluster transfer to other proteins (e.g., SufA, ferredoxin, and aconitase), however SufA is much more efficient. Recent studies demonstrated that the cluster transfer to aconitase from $SufBC_2D$ or $SudB_2C_2$ proceed through a Fe-S SufA intermediate when apo-SufA is present during cluster transfer [112]. All the available data support the idea that SufA is a carrier protein important for the maturation of both [2Fe-2S] and [4Fe-4S] proteins [101].

2.4. The Cytosolic Iron-Sulfur Assembly Machinery

The cytosolic Fe-S protein assembly (CIA) machinery is uniquely found in eukaryotes, and functions downstream of mitochondria to maturate Fe-S proteins in the cytosol and in the nucleus [113]. The first and essential step for the biogenesis of cytosolic and nuclear Fe-S proteins is the mitochondrial delivery of a sulfur-containing compound to the cytosol (Figure 7). The ABC transporter ABCB7 mediates the export reaction of the S-compound [75, 114], which is dependent on glutathione. After the export of the S-compound to the cytosol the components of the CIA machinery perform the synthesis of the Fe-S cluster and conduct its insertion into client proteins. The iron source of the CIA machinery is still not well-defined. To acquire this piece of information it is fundamental to define the chemical nature of the S-compound and whether it contains only inorganic sulfur, sulfur and iron, or an iron-sulfur cluster. However, it appears that PCBP-family iron chaperones, that traffick cytosolic ferrous ions to ferritin for storage or to non-heme iron enzymes [115, 116], are linked to the CIA machinery. Indeed, HEK293 cells lacking PCBP1 or PCBP2 exhibit loss of activity

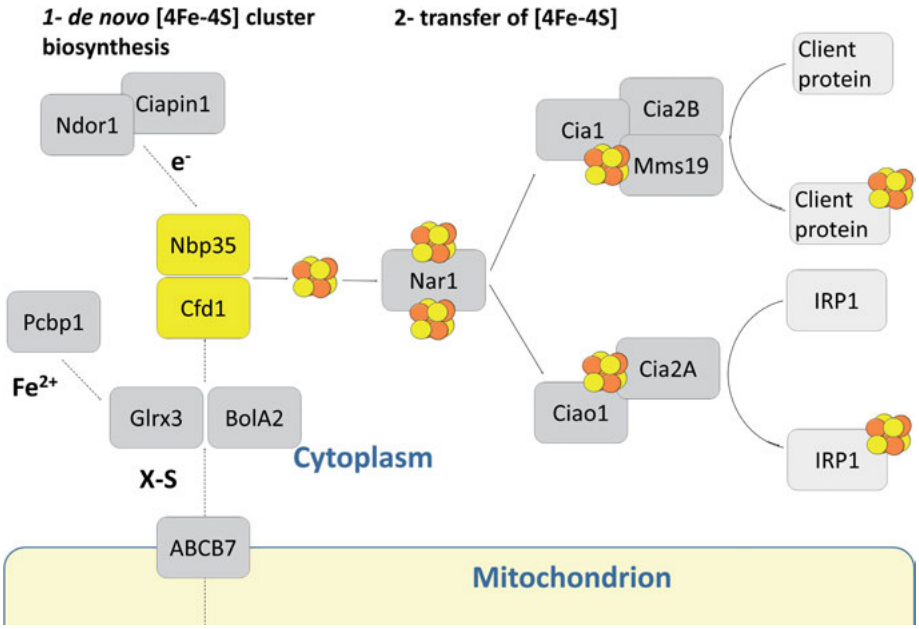


Figure 7. Schematic representation of the CIA machinery in the cytosol of eukaryotes.

in cytosolic aconitase [117], an enzyme with a labile [4Fe-4S] cluster. It might be that PCBP1 may contribute to the flow of iron ions into the cytosolic Fe-S cluster pool, but this activity has not been directly demonstrated [118].

Like ISC and SUF, the CIA-dependent assembly is divided into two main steps: (1) the Fe-S cluster biosynthesis, exclusively producing a [4Fe-4S] cluster, and (2) the transfer of the cluster to client proteins [74].

2.4.1. [4Fe-4S] Cluster Biosynthesis

The scaffold system of the CIA machinery is a hetero-complex formed by Cfd1 and Nbp35, two P-loop NTPases [119, 120]. This protein complex possesses two conserved [4Fe-4S]-binding sites: one is labile and located at the interface of the two proteins; the other binds the [4Fe-4S] cluster with a high affinity and is located at the N-terminus of the Nbp35 protein. The cluster formed in the former site is likely transferred to client proteins in the second step of biogenesis, whereas the cluster formed in the latter site remains anchored to Nbp35 and is essential for its function. Like other machineries, also the CIA machinery depends on the supply of electrons. These have been proposed to be provided by a hetero-dimer composed by Ndor1 and Ciapin1, which represent an electron transfer chain [121, 122]. Ndor1 is an essential protein containing NADPH-, FAD-, and FMN-binding sites. Ciapin1 binds a [2Fe-2S] cluster in a cluster-binding site and a [2Fe-2S] or [4Fe-4S] cluster on a second cluster-binding site depending on how the preparation of the protein sample has been conducted *in vitro* [123, 124].

Electrons are transferred from NADPH, via FAD, to the FMN cofactor. From FMN, the electrons arrive to the [2Fe-2S] cluster of Ciapin1 [122]. The precise destination of electrons in the CIA pathway is still unknown; however, in yeast, the absence of Ndor1 and Ciapin1 homologs results in the impairment of the [4Fe-4S] cluster in the N-terminus of Nbp35. Another essential player in the CIA machinery is Glrx3. A plausible hypothesis for the role of Glrx3 in the CIA pathway is that Glrx3 could provide its [2Fe-2S] cluster to the scaffold complex for its conversion to a [4Fe-4S] cluster; still, this hypothesis is experimentally not demonstrated. On the other hand, it has been shown that Glrx3 can work as a [2Fe-2S] cluster trafficking protein. Indeed, Glrx3 matures Ciapin1 by specifically transferring two [2Fe-2S] clusters to the two cluster-binding sites of Ciapin1 [125]. Ciapin1 and Glrx3 are, indeed, protein partners in the cellular context as identified by the yeast-two-hybrid assay [126]. Glrx3 also forms a complex with the cytosolic protein BolA2 [127–129], which binds two [2Fe-2S] clusters. The Glrx3-BolA2 complex was observed in mammalian cells and transfers both its two [2Fe-2S] clusters to Ciapin1 producing its mature holo state [127, 129]. In conclusion, both [2Fe-2S]-bound forms of Glrx3 and Glrx3-BolA2 can work as a [2Fe-2S] cluster transfer component in the CIA machinery.

2.4.2. Transfer of the Cluster to Client Proteins

The transfer and insertion of the newly assembled [4Fe-4S] cluster into the client apoproteins requires the combined function of Nar1 and of the CIA-targeting complex [130]. Nar1 was suggested to connect the early-acting component with the CIA-targeting complex by a yet unknown mode of action, however it coordinates two [4Fe-4S] clusters that are essential for its function [131]. The CIA targeting complex is composed by Cia1, Cia2B, and Mms19 which can form binary and tertiary adducts [132]. The CIA complex is involved in the maturation of many cytosolic and nuclear Fe-S proteins. In humans, different combinations of the CIA targeting complex components are required to maturate specific client proteins [133, 134]. The proposed model in humans assumes Nar1 as the CIA component distributing [4Fe-4S] clusters assembled on the Cfd1-Nbp35 complex to the CIA targeting complex and to a complex formed by CIAO1 (Cia1 in yeast) and Cia2A (not present in yeast), which is exclusively implicated in the [4Fe-4S] protein maturation of the cytosolic iron regulatory protein 1 (IRP1) to form aconitase [134, 135]. Recently, two newly identified CIA factors, termed Yae1 and Lto1, were shown to function as specific adaptors connecting the CIA targeting complex with a specific [4Fe-4S] client protein, a cytosolic ABCE1 protein [136, 137].

3. OCCURRENCE AND FUNCTION OF IRON-SULFUR PROTEINS IN ORGANISMS

During the first billion years of life on Earth, the environment was anaerobic, and iron and sulfur ions were abundant and available for organisms. Therefore, the interaction between Fe-S clusters and proteins was largely exploited by all ancestral forms of life to perform functions in essential processes. This remained un-

changed until the appearance and evolution of photosynthetic organisms, about two billion years ago, which led to a drastic increase in the concentration of dioxygen in the atmosphere. By oxidizing iron(II) to the iron(III) state, which rapidly precipitates as ferric hydroxide or forms insoluble salts, dioxygen reduced drastically the bioavailability of this essential metal in the environment; as a consequence, as dioxygen accumulated in the atmosphere, iron became a limiting nutrient in many aerobic habitats. Furthermore, reactive oxygen species (ROS) are able to convert exposed Fe-S clusters to unstable inorganic chemical species that quickly decompose [138]. Aerobic organisms faced the above challenges throughout their evolution by removing or substituting many Fe-S clusters in proteins with other less oxygen-sensitive metal ions, like zinc [2, 139]. As a result of this process, nowadays the Fe-S proteomes of aerobes are generally smaller than those of anaerobes with similar genome size, although all prokaryotes encode a number of Fe-S proteins which increases proportionally to the genome size [1]. Further, aerobes appear enriched in [2Fe-2S] proteins with respect to anaerobes, likely because this cluster type is less iron-demanding and more chemically stable [1]. Fe-S clusters were only conserved in those proteins in which the cluster function could not be replaced by other cofactors. Genome analysis revealed that these proteins constitute a common core of probably ancient Fe-S proteins shared by all prokaryotic organisms, and most of them bind a [4Fe-4S] cluster [1]. These proteins map to about 50 Fe-S families, which account for 57 % $\pm$ 9 % of the Fe-S genomes in anaerobes and for 71 % $\pm$ 8 % of the Fe-S genomes in aerobes (see sections below). These ancient families are mainly involved in energy production (like respiration), as well as in other essential processes such as nucleotide and amino acid metabolism, coenzyme metabolism, Fe-S biogenesis, and RNA processing or DNA maintenance. Some of these processes use the same components in both aerobes and anaerobes, such as GTP 3′,8-cyclase (MoaA), an enzyme involved in molybdopterin biosynthesis (i.e., coenzyme metabolism) [140]. This enzyme is conserved in 82 % of aerobes and in 92 % of anaerobes and performs the same catalytic mechanism in the proteins of both groups [1]. Other processes, instead, use evolutionarily related enzymes that have been recruited in different but functionally similar roles. For example, fumarate reductase and succinate dehydrogenase are evolutionarily related enzymes involved in respiration [141] (conserved in 95 % of aerobes and 63 % anaerobes) that however, catalyze the same reaction but in opposite directions [1].

Besides the above described common core, for the reasons discussed above the Fe-S proteomes of prokaryotes display specific features that depend on the oxygen requirements of the organisms (i.e., aerobes, anaerobes, and facultative anaerobes), and will be thus discussed here in the following sections.

3.1. Fe-S Proteomes of Anaerobic Prokaryotes

The large majority of anaerobes (about 77 % of the organisms analyzed in [1]) encode a large fraction of Fe-S proteins, representing on average 4.7 % $\pm$ 1.0 % of the total genome (Table 1). The organisms with the highest fraction (6.6 % of

total) are *Methanococcus maripaludis, Methanotorris igneus* and *Moorella thermoacetica*, which are thermophiles and hyperthermophiles.

Table 1. List of anaerobic organisms with a high content of Fe-S proteins.

Organism	Proteome size	# Fe-S proteins	% [4Fe-4S]	% [2Fe-2S]	% [3Fe-4S]	% Fe-Cys$_4$
Methanothermus fervidus	1283	83	95	2	1	1
Cryptobacterium curtum	1357	73	90	7	1	3
Fervidicoccus fontis	1385	57	83	9	0	8
Thermosphaera aggregans	1387	60	91	3	0	6
Ignicoccus hospitalis	1434	66	91	3	3	3
Desulfurococcus kamchatkensis	1471	61	85	11	0	4
Caldisphaera lagunensis	1478	59	84	12	0	4
Coprothermobacter proteolyticus	1482	55	79	15	1	5
Acidilobus saccharovorans	1499	60	87	8	0	5
Desulfurobacterium thermolithotrophum	1509	85	83	11	5	2
Thermoplasmatales archaeon	1523	64	87	5	3	5
Elusimicrobium minutum	1529	51	73	20	2	6
Methanosphaera stadtmanae	1535	76	90	6	2	3
Caldisericum exile	1581	69	70	19	1	9
Methanothermococcus okinawensis	1595	105	94	3	1	2
Staphylothermus hellenicus	1599	56	89	7	0	4
Hyperthermus butylicus	1603	73	94	1	0	4
Dehalogenimonas lykanthroporepellens	1659	107	84	8	3	5
Hippea maritima	1677	64	78	15	3	4
Methanopyrus kandleri	1687	99	94	4	1	2
Methanococcus maripaludis	1722	114	93	3	2	2
Nautilia profundicola	1730	69	79	17	4	0
Thermanaerovibrio acidaminovorans	1730	77	75	21	1	3

Organism	Proteome size	# Fe-S proteins	% [4Fe-4S]	% [2Fe-2S]	% [3Fe-4S]	% Fe-Cys_4
Methanocorpusculum labreanum	1741	100	91	3	0	6
Methanocaldococcus jannaschii	1771	102	94	3	1	2
Methanotorris igneus	1772	117	93	4	1	2
Pyrococcus abyssi	1780	80	87	3	0	10
Anaerococcus prevotii	1806	57	78	13	2	7
Thermodesulfobium narugense	1807	117	85	8	4	4
Finegoldia magna	1812	58	70	17	2	11
Thermovibrio ammonificans	1813	79	84	9	5	2
Veillonella parvula	1844	61	81	12	5	2
Methanothermobacter thermautotrophicus	1873	120	89	5	3	3
Thermovirga lienii	1875	88	79	15	1	6
Aminobacterium colombiense	1876	73	81	12	1	6
Thermotoga maritima	1891	88	74	17	3	6
Petrotoga mobilis	1898	81	78	14	3	5
Dictyoglomus thermophilum	1912	81	75	18	3	4
Ignisphaera aggregans	1930	58	85	8	0	7
Fervidobacterium pennivorans	1947	74	70	22	2	7
Thermosipho africanus	1954	73	67	24	2	7
Methanosalsum zhilinae	1976	107	88	4	1	7
Methanohalophilus mahii	1987	118	88	4	2	6
Anaerobaculum mobile	2017	105	74	23	1	2
Gordonibacter pamelaeae	2021	123	96	2	0	2
Wolinella succinogenes	2043	98	83	13	3	1
Marinitoga piezophila	2046	81	79	14	1	6
Thermoproteus tenax	2049	88	85	10	1	4
Ammonifex degensii	2080	118	84	12	2	2
Nitratifractor salsuginis	2088	67	76	18	4	2
Calditerrivibrio nitroreducens	2100	96	81	11	2	6

Organism	Proteome size	# Fe-S proteins	% [4Fe-4S]	% [2Fe-2S]	% [3Fe-4S]	% Fe-Cys$_4$
Pyrococcus furiosus	2122	87	84	4	2	10
Thermodesulfatator indicus	2195	120	90	6	2	2
Thermosediminibacter oceani	2197	85	78	16	1	5
Methanobrevibacter ruminantium	2217	102	91	4	2	3
Chlorobium tepidum	2245	102	77	16	3	4
Methanohalobium evestigatum	2254	105	91	3	2	4
Flexistipes sinusarabici	2261	105	79	12	3	6
Methanococcoides burtonii	2273	112	90	5	1	5
Acetohalobium arabaticum	2282	121	73	22	2	3
Thermococcus kodakarensis	2306	86	80	6	0	14
Syntrophothermus lipocalidus	2313	124	85	11	2	2
Halothermothrix orenii	2342	91	76	16	2	5
Deferribacter desulfuricans	2374	111	81	11	3	5
Acidaminococcus intestini	2401	80	82	8	2	8
Halorhodospira halophila	2407	91	78	10	5	6
Archaeoglobus fulgidus	2420	144	87	9	1	4
Moorella thermoacetica	2463	163	85	11	1	3
Halobacteroides halobius	2468	98	76	16	3	5
Ferroglobus placidus	2480	144	90	6	1	2
Vulcanisaeta distributa	2493	90	87	10	1	3
Syntrophomonas wolfei	2504	108	78	19	1	2
Desulfohalobium retbaense	2526	143	84	8	2	6
Methanomethylovorans hollandica	2556	115	90	6	1	3
Desulfurispirillum indicum	2571	99	79	16	4	2
Methanoculleus bourgensis	2580	100	86	8	1	5

Organism	Proteome size	# Fe-S proteins	% [4Fe-4S]	% [2Fe-2S]	% [3Fe-4S]	% Fe-Cys$_4$
Desulfurivibrio alkaliphilus	2620	131	81	13	1	5
Carboxydothermus hydrogenoformans	2620	154	87	8	2	3
Methanosphaerula palustris	2655	146	83	9	2	7
Caldicellulosiruptor saccharolyticus	2682	109	79	14	4	2
Pelodictyon phaeoclathratiforme	2707	105	76	15	4	5
Slackia heliotrinireducens	2766	162	92	6	0	1
Methanoplanus petrolearius	2785	139	88	4	3	5
Sulfuricurvum kujiense	2798	102	77	18	3	2
Thermacetogenium phaeum	2824	147	80	14	2	3
Melioribacter roseus	2840	88	79	18	2	1
Methanosaeta concilii	2850	123	88	7	2	3
Alkalilimnicola ehrlichii	2865	99	73	20	3	4
Desulfobacca acetoxidans	2866	145	84	11	2	2
Ilyobacter polytropus	2880	137	76	15	2	7
Natranaerobius thermophilus	2906	137	84	11	1	4
Pelotomaculum thermopropionicum	2919	149	83	14	2	1
Thermincola potens	2949	172	83	11	2	4
Denitrovibrio acetiphilus	2964	138	85	9	3	4
Coprococcus catus	2972	93	78	13	3	6
Heliobacterium modesticaldum	2999	114	80	14	3	3
Eggerthella lenta	3070	167	93	4	0	2
Methanospirillum hungatei	3131	132	88	5	2	6
Syntrophus aciditrophicus	3166	148	77	16	1	6
Allochromatium vinosum	3220	124	76	17	2	4
Desulfotalea psychrophila	3234	143	80	15	2	3
Syntrophobotulus glycolicus	3251	149	83	10	2	4

Organism	Proteome size	# Fe-S proteins	% [4Fe-4S]	% [2Fe-2S]	% [3Fe-4S]	% Fe-Cys$_4$
Desulfococcus oleovorans	3265	205	87	6	1	6
Desulfarculus baarsii	3277	163	81	14	2	3
Desulfobulbus propionicus	3283	168	81	15	2	2
Desulfomicrobium baculatum	3436	176	86	7	2	5
Acetobacterium woodii	3473	166	79	16	2	3
Desulfocapsa sulfexigens	3532	161	82	13	2	3
Desulfovibrio vulgaris	3535	154	85	8	2	5
Geobacter metallireducens	3566	189	81	13	2	4
Pelobacter propionicus	3804	163	79	15	2	4
Syntrophobacter fumaroxidans	4064	259	81	15	1	4
Spirochaeta smaragdinae	4219	128	71	24	2	4
Thiocystis violascens	4330	138	75	17	2	6
Desulfobacula toluolica	4375	269	85	10	1	4
Methanosarcina acetivorans	4540	177	90	6	1	3
Alkaliphilus metalliredigens	4625	150	75	19	1	4
Desulfobacterium autotrophicum	4943	265	85	9	1	5
Desulfitobacterium hafniense	5060	292	90	7	1	2
Desulfosporosinus orientis	5242	314	84	13	1	2
Desulfatibacillum alkenivorans	5252	299	88	9	1	2
Desulfomonile tiedjei	5494	312	77	19	1	3
	Average		**83**	**11**	**2**	**4**
	Std Dev		**6**	**5**	**1**	**2**

In anaerobic prokaryotes, the number of [4Fe-4S] sites in proteins is much larger than the other Fe-S types (Table 1). These cover from 67 % to 96 % of all Fe-S sites, and represent, on average, 83 % $\pm$ 6 % of the total. The second-most common Fe-S type is [2Fe-2S], which on average represents 11 % $\pm$ 5 % of the sites. Most of [2Fe-2S] sites perform electron transfer and are found in ferredoxin proteins, especially involved in respiratory processes. Other widespread [2Fe-2S]

proteins are those involved in Fe-S biogenesis. The Fe-$(Cys)_4$ site represents on average 4 % ± 2 %, whereas [3Fe-4S] clusters are on average only 2 % ± 1 %.

The Fe-S proteins encoded by these organisms generally map to a number of protein families that increases by increasing the genome size (Table 1). For example, anaerobes having between 2000 and 2500 genes in their genome contain on average 71 ± 10 Fe-S families, whereas those having a genome between 4500 and 5000 genes contain on average 100 ± 11 Fe-S families. The average number of distinct Fe-S families in anaerobes is generally higher than in aerobes and the difference increases with increasing genome size (see Section 3.2).

About one hundred Fe-S families are conserved in at least 30 % of anaerobes; 51 of them belong to the above discussed ancient shared core, which in anaerobes accounts for 57 % ± 9 % of the Fe-S proteomes [1]. The remaining 54 Fe-S families are not conserved in aerobes, and can be regarded as the set of Fe-S proteins that aerobes have lost during evolution [1]. The large majority of these Fe-S families are involved in energy production and conversion, and bind [4Fe-4S] clusters, whereas none of these binds a [2Fe-2S] cluster.

A minority of anaerobes (23 %) encode a low fraction of Fe-S proteins (on average 2.2 % ± 0.4 % of the genome). These are mostly symbionts and parasites that reduced their genome by exploiting the host organism for some parts of the metabolism. Interestingly, even if these organisms encode less Fe-S proteins, they still conserve almost entirely the shared core of ancient Fe-S proteins [1].

3.2. Fe-S Proteomes of Aerobic Prokaryotes

The fraction of Fe-S proteins encoded in aerobes is lower than that of anaerobes [1]. On average, this represents 1.8 % ± 0.6 % of the genome (Table 2). *Borrelia burgdorferi* appears to lack Fe-S proteins altogether, and indeed is able to survive without iron [142]. Also bacteria of the Mycoplasma genus (i.e., *Mycoplasma genitalium* and *Mycoplasma pneumoniae*) only encode a potential IscU homolog, which however does not conserve the entire Fe-S site: it conserves three cysteines, but lacks the histidine that is the fourth Fe-S ligand. *In vitro* studies revealed that both these organisms do not show a significant binding of human transferrin for iron-sequestering, and only *Mycoplasma pneumoniae* has a moderate binding affinity for lactoferrin [143], suggesting that their use of iron ions for survival is limited or completely absent.

In aerobic prokaryotes the number of [4Fe-4S] sites is larger than of the other Fe-S types, however, this prevalence is less sharp than in anaerobes (Table 2). These sites cover from 35 % to 92 % of the total sites, and represent, on average, 65 % ± 8 % of the total. The organism with the highest percentage of [4Fe-4S] sites (i.e., 92 %) is *Pyrolobus fumarii*, a hyper-thermophilic archaeon; in general, all the aerobes with a high [4Fe-4S] percentage belong to the archaeal group. In aerobes, the percentage of [2Fe-2S] sites is 29 % ± 8 %, so they appear to be enriched in [2Fe-2S] proteins with respect to anaerobes. In *Sphingomonas wittichii RW1*, the percentage of [2Fe-2S] raises to 64 %, thus being larger than that of [4Fe-4S] sites. Fe-$(Cys)_4$ sites represent on average 2 % ± 2 % whereas [3Fe-4S] clusters are on average 4 % ± 2 %.

Table 2. List of aerobic organisms and their content of Fe-S proteins.

Organism	Proteome size	# Fe-S proteins	% [4Fe-4S]	% [2Fe-2S]	% [3Fe-4S]	% Fe-Cys_4
Mycoplasma genitalium	475	0	0	0	0	0
Mycoplasma pneumoniae	648	0	0	0	0	0
Tropheryma whipplei	808	12	40	40	10	10
Rickettsia prowazekii	843	25	50	50	0	0
Neorickettsia sennetsu	932	28	52	45	3	0
Bartonella quintana	1142	31	62	32	3	3
Anaplasma phagocytophilum	1264	30	51	46	3	0
Wolbachia endosymbiont	1275	27	50	47	3	0
Rickettsia conorii	1374	25	50	50	0	0
Borrelia burgdorferi	1390	0	0	0	0	0
Picrophilus torridus	1537	45	85	13	2	0
Helicobacter pylori	1594	41	77	22	2	0
Mycobacterium leprae	1605	25	73	15	12	0
Thioalkalimicrobium cyclicum	1665	44	66	26	5	3
Streptococcus pyogenes	1696	15	71	14	0	14
Aeropyrum pernix	1700	46	84	13	0	4
Erysipelothrix rhusiopathiae	1775	14	77	23	0	0
Pediococcus claussenii	1881	7	40	40	0	20
Moraxella catarrhalis	1886	54	65	25	3	7
Neisseria meningitidis	1953	57	65	30	1	4
Caldivirga maquilingensis	1963	69	88	7	1	4
Acidimicrobium ferrooxidans	1964	56	78	17	5	0
Orientia tsutsugamushi	1967	22	48	52	0	0
Pyrolobus fumarii	1967	76	93	2	3	3
Cenarchaeum symbiosum	2014	48	83	17	0	0
Weeksella virosa	2049	37	77	21	2	0
Polynucleobacter necessarius	2077	68	64	28	5	3

Organism	Proteome size	# Fe-S proteins	% [4Fe-4S]	% [2Fe-2S]	% [3Fe-4S]	% Fe-Cys_4
Acidothermus cellulolyticus	2157	60	73	22	5	0
Sulfurimonas autotrophica	2158	74	77	18	3	2
Thermus thermophilus	2173	55	76	19	6	0
Riemerella anatipestifer	2186	33	78	22	0	0
Rothia dentocariosa	2217	27	68	29	3	0
Sulfolobus acidocaldarius	2224	71	72	24	1	4
Micrococcus luteus	2236	34	56	31	13	0
Metallosphaera sedula	2256	90	81	16	1	3
Ornithobacterium rhinotracheale	2267	38	73	18	0	8
Thermaerobacter marianensis	2327	51	63	31	6	0
Kocuria rhizophila	2357	43	64	27	9	0
Halothiobacillus neapolitanus	2357	52	59	30	6	6
Granulibacter bethesdensis	2433	67	64	32	5	0
Leifsonia xyli	2470	26	57	30	13	0
Methylacidiphilum infernorum	2472	79	68	27	4	2
Kytococcus sedentarius	2554	37	71	25	4	0
Comamonadaceae bacterium	2626	43	73	21	6	0
Idiomarina loihiensis	2628	52	55	31	8	6
Haloquadratum walsbyi	2652	55	61	35	4	0
Parvularcula bermudensis	2685	47	57	36	7	0
Croceibacter atlanticus	2701	39	70	24	2	4
Methylobacillus flagellatus	2753	71	68	26	4	2
Alcanivorax borkumensis	2755	69	53	34	6	7
Nitrosospira multiformis	2805	66	67	27	4	2
Chthonomonas calidirosea	2809	59	72	18	5	5

Organism	Proteome size	# Fe-S proteins	% [4Fe-4S]	% [2Fe-2S]	% [3Fe-4S]	% Fe-Cys$_4$
Natronomonas pharaonis	2820	69	78	19	3	0
Xylella fastidiosa	2832	46	60	31	7	2
Thermobaculum terrenum	2832	61	72	20	4	3
Thermomicrobium roseum	2854	70	65	27	5	3
Rhodothermus marinus	2863	62	68	27	5	0
Gallionella capsiferriformans	2894	77	71	21	4	4
Methylovorus glucosetrophus	2909	70	63	32	4	1
Truepera radiovictrix	2945	53	71	24	5	0
Segniliparus rotundus	3006	43	59	25	13	4
Erythrobacter litoralis	3011	54	58	34	8	0
Acetobacter pasteurianus	3049	74	61	33	4	2
Brachybacterium faecium	3068	33	71	26	3	0
Thermobifida fusca	3087	47	72	22	7	0
Halovivax ruber	3099	47	79	21	0	0
Frateuria aurantia	3101	57	60	33	6	1
Clavibacter michiganensis	3116	25	52	34	14	0
Coraliomargarita akajimensis	3120	63	77	16	5	2
Aequorivita sublithincola	3140	43	69	22	6	4
Rubrobacter xylanophilus	3140	73	67	29	4	0
Kyrpidia tusciae	3150	93	66	30	3	1
Deinococcus radiodurans	3167	50	70	23	7	0
Hirschia baltica	3187	61	63	31	6	0
Sphingopyxis alaskensis	3195	58	65	31	5	0
Brucella melitensis	3198	58	69	23	6	3
Robiginitalea biformata	3205	52	59	32	5	3
Alicyclobacillus acidocaldarius	3208	46	73	20	6	0
Salinibacter ruber	3257	39	71	22	7	0

Organism	Proteome size	# Fe-S proteins	% [4Fe-4S]	% [2Fe-2S]	% [3Fe-4S]	% Fe-Cys_4
Enterococcus faecalis	3264	30	79	17	0	4
Cyanobium gracile	3280	80	66	25	6	3
Herminiimonas arsenicoxydans	3294	80	70	25	4	2
Chromohalobacter salexigens	3298	74	60	33	5	2
Brevundimonas subvibrioides	3327	54	64	27	7	2
Acinetobacter baumannii	3367	67	52	40	7	1
Xylanimonas cellulosilytica	3443	49	69	23	6	2
Leadbetterella byssophila	3465	52	76	16	7	0
Owenweeksia hongkongensis	3485	40	76	20	0	4
Sphaerobacter thermophilus	3485	66	71	27	1	1
Oceanobacillus iheyensis	3500	34	70	20	9	0
Meiothermus silvanus	3505	59	66	28	6	0
Hyphomonas neptunium	3505	61	62	31	5	1
Psychroflexus torquis	3526	39	68	22	6	4
Morganella morganii	3526	101	73	19	3	5
Thermobispora bispora	3546	72	68	28	4	0
Halorubrum lacusprofundi	3560	64	73	24	3	0
Intrasporangium calvum	3563	74	72	24	4	0
Gramella forsetii	3584	52	65	25	7	3
Bdellovibrio bacteriovorus	3586	48	56	42	2	0
Parvibaculum lavamentivorans	3636	64	53	39	7	1
Natronobacterium gregoryi	3656	85	78	19	3	0
Listonella anguillarum	3678	66	59	28	7	6
Isosphaera pallida	3722	71	72	24	4	0
Leptospira biflexa	3725	63	73	21	5	1

Organism	Proteome size	# Fe-S proteins	% [4Fe-4S]	% [2Fe-2S]	% [3Fe-4S]	% Fe-Cys$_4$
Gluconobacter oxydans	3732	64	65	29	5	1
Caulobacter crescentus	3737	60	57	37	5	1
Cellvibrio japonicus	3750	59	61	27	6	5
Marivirga tractuosa	3757	53	65	21	11	3
Asticcacaulis excentricus	3763	66	57	38	5	0
Beijerinckia indica	3784	99	61	34	3	2
Cytophaga hutchinsonii	3785	58	72	20	6	1
Nitrosococcus halophilus	3817	80	71	22	5	3
Methylocella silvestris	3818	127	57	37	3	3
Saccharomonospora viridis	3828	67	62	33	5	0
Gluconacetobacter diazotrophicus	3849	89	59	35	4	2
Phenylobacterium zucineum	3854	68	68	27	4	1
Thermobacillus composti	3858	58	73	19	7	0
Halalkalicoccus jeotgali	3873	80	72	25	2	0
Flexibacter litoralis	3878	38	69	19	8	4
Ramlibacter tataouinensis	3879	83	58	39	3	0
Halogeometricum borinquense	3898	72	72	25	3	0
Mycobacterium tuberculosis	3906	63	63	29	6	2
Advenella kashmirensis	3933	72	52	45	2	1
Gemmatimonas aurantiaca	3935	63	66	29	5	0
Acidiphilium multivorum	3948	105	62	33	3	3
Yersinia pestis	3979	86	67	24	3	7
Saccharophagus degradans	4007	62	61	28	6	5
Haloferax volcanii	4015	79	72	25	2	0
Fluviicola taffensis	4033	45	75	23	2	0
Bacillus halodurans	4052	59	70	23	5	1

Organism	Proteome size	# Fe-S proteins	% [4Fe-4S]	% [2Fe-2S]	% [3Fe-4S]	% Fe-Cys_4
Turneriella parva	4139	67	64	28	6	2
Natronococcus occultus	4154	93	77	21	2	0
Cellulophaga algicola	4163	53	63	26	5	5
Bacillus subtilis	4175	57	68	24	6	1
Dinoroseobacter shibae	4192	94	67	27	4	2
Beutenbergia cavernae	4197	39	65	27	8	0
Natrialba magadii	4212	61	74	22	4	0
Frankia symbiont	4215	90	62	31	6	1
Halopiger xanaduensis	4221	83	83	14	2	0
Tsukamurella paurometabola	4242	52	59	33	5	3
Haloarcula marismortui	4243	76	75	23	3	0
Terriglobus roseus	4245	74	71	24	5	0
Saprospira grandis	4251	41	60	34	2	4
Pedobacter heparinus	4252	64	72	22	6	0
Emticicia oligotrophica	4267	72	70	25	5	0
Ruegeria pomeroyi	4278	101	56	38	3	3
Pseudoalteromonas atlantica	4281	74	55	36	5	4
Nitrobacter hamburgensis	4326	74	60	33	6	1
Phaeobacter gallaeciensis	4359	79	62	34	4	0
Leptothrix cholodnii	4363	131	61	35	3	1
Stenotrophomonas maltophilia	4386	58	61	29	7	3
Marinobacter adhaerens	4410	94	67	25	4	4
Starkeya novella	4431	83	63	31	3	3
Trichodesmium erythraeum	4451	96	58	36	3	3
Methylomonas methanica	4494	103	63	29	3	5
Roseobacter litoralis	4537	95	60	35	4	1
Salmonella enterica	4553	122	80	14	2	4
Arthrobacter chlorophenolicus	4590	42	59	33	7	0

Organism	Proteome size	# Fe-S proteins	% [4Fe-4S]	% [2Fe-2S]	% [3Fe-4S]	% Fe-Cys$_4$
Zunongwangia profunda	4653	51	62	26	9	3
Kineococcus radiotolerans	4681	44	66	28	7	0
Pirellula staleyi	4717	81	71	25	4	0
Zobellia galactanivorans	4732	61	64	24	8	5
Herbaspirillum seropedicae	4735	98	56	39	4	1
Acidovorax avenae	4737	80	57	38	4	1
Planctomyces brasiliensis	4750	59	76	16	5	3
Lysinibacillus sphaericus	4771	52	77	20	2	2
Ochrobactrum anthropi	4799	64	63	28	5	4
Geodermatophilus obscurus	4810	82	61	34	5	0
Granulicella mallensis	4815	88	76	20	3	2
Blastococcus saxobsidens	4817	77	61	35	5	0
Thermomonospora curvata	4890	91	69	27	4	1
Comamonas testosteroni	4894	95	50	45	4	1
Salinispora arenicola	4917	80	65	29	4	2
Xanthomonas oryzae	4988	54	62	30	6	1
Flavobacterium johnsoniae	5017	62	68	25	5	2
Bacillus anthracis	5039	62	71	21	6	1
Azotobacter vinelandii	5050	128	60	32	3	6
Paracoccus denitrificans	5077	97	66	29	3	2
Ralstonia solanacearum	5111	98	61	34	3	1
Haloterrigena turkmenica	5113	104	78	20	2	0
Nakamurella multipartita	5240	72	74	22	4	0
Herpetosiphon aurantiacus	5279	63	73	23	4	0
Microlunatus phosphovorus	5338	52	71	22	6	0
Sphingomonas wittichii	5345	155	35	61	4	0
Nocardiopsis alba	5508	57	70	24	6	0

Organism	Proteome size	# Fe-S proteins	% [4Fe-4S]	% [2Fe-2S]	% [3Fe-4S]	% Fe-Cys_4
Enterobacter cloacae	5518	99	68	24	3	6
Pseudomonas aeruginosa	5572	115	58	34	3	5
Glaciecola psychrophila	5618	74	55	38	4	3
Dyadobacter fermentans	5719	71	63	30	5	2
Chroococcidiopsis thermalis	5752	116	56	38	4	2
Runella slithyformis	5774	82	64	32	4	0
Tistrella mobilis	5783	105	60	35	4	1
Anabaena cylindrica	5838	109	55	38	3	4
Conexibacter woesei	5914	80	54	40	4	2
Brevibacillus brevis	5947	67	72	20	6	2
Delftia acidovorans	6040	100	57	39	3	1
Bacillus thuringiensis	6194	61	69	24	6	1
Sinorhizobium meliloti	6218	111	50	45	3	3
Micromonospora aurantiaca	6222	74	71	24	5	0
Cylindrospermum stagnale	6229	116	59	35	3	3
Microcystis aeruginosa	6312	85	61	32	7	1
Stackebrandtia nassauensis	6379	75	67	24	9	0
Arthrospira platensis	6630	90	66	29	4	2
Nostoc punctiforme	6689	125	57	36	3	4
Haliscomenobacter hydrossis	6752	90	61	33	2	4
Variovorax paradoxus	6753	129	53	43	3	2
Pseudonocardia dioxanivorans	6797	139	49	48	3	0
Achromobacter xylosoxidans	6815	109	54	41	3	1
Actinosynnema mirum	6916	74	69	26	5	0
Spirosoma linguale	6938	82	61	36	3	0
Kribbella flavida	6943	65	68	26	6	0

Organism	Proteome size	# Fe-S proteins	% [4Fe-4S]	% [2Fe-2S]	% [3Fe-4S]	% Fe-Cys$_4$
Rhizobium leguminosarum	7143	84	62	33	4	2
Niastella koreensis	7174	98	69	27	3	1
Chitinophaga pinensis	7192	77	68	27	5	1
Saccharopolyspora erythraea	7197	99	57	37	5	1
Mesorhizobium loti	7272	104	62	35	3	1
Myxococcus xanthus	7316	94	65	33	2	0
Rhodopirellula baltica	7325	71	74	19	5	2
Kitasatospora setae	7566	85	71	25	5	0
Cupriavidus necator	7832	173	50	46	2	2
Corallococcus coralloides	8033	75	58	38	3	1
Methylobacterium nodulans	8308	132	55	42	3	1
Stigmatella aurantiaca	8352	100	62	34	2	2
Nocardia brasiliensis	8414	90	59	34	4	3
Burkholderia xenovorans	8702	198	48	48	2	2
Bradyrhizobium japonicum	8826	134	54	42	3	1
Catenulispora acidiphila	8913	101	66	30	4	0
Streptosporangium roseum	8975	96	68	27	5	0
Rhodococcus jostii	9145	167	48	45	6	1
Amycolatopsis mediterranei	9575	116	57	36	5	1
Streptomyces bingchenggensis	10022	105	67	30	3	0
Sorangium cellulosum	10400	122	66	31	2	1
	Average		**65**	**29**	**4**	**2**
	Std Dev		**10**	**9**	**2**	**2**

The Fe-S proteins encoded by aerobic organisms map to a number of distinct protein families that levels off at about 65 families for genomes with at least 4000 genes [1]. However, the number of Fe-S proteins encoded in aerobes linear-

ly increases with the genome size, thus, the higher number of Fe-S proteins in the aerobes with larger genomes is not due to a higher number of different Fe-S families, but to a higher number of co-orthologous proteins (i.e., paralogs that were duplicated after speciation). Such proteins belong to the same Fe-S family and have a different specialization (e.g., enzymes performing the same catalysis but on different substrates).

The number of distinct Fe-S families in aerobes is generally lower than in anaerobes. This difference increases with increasing genome size. For example, aerobes having between 2000 and 2500 genes in their genome contain on average 43 $\pm$ 12 Fe-S families (versus 71 $\pm$ 10 in anaerobes, see above) whereas those having a genome between 4500 and 5000 genes contain on average 57 $\pm$ 11 Fe-S families (versus 100 $\pm$ 11 in anaerobes, see above) [1].

62 Fe-S families are conserved in at least 30 % of aerobes; as previously noted, 51 of them belong to the ancient shared core which in aerobes accounts for 71 % $\pm$ 8 % of the Fe-S genomes. The remaining 11 Fe-S families can be regarded as the set of Fe-S proteins that aerobes have evolved during evolution [1]. The majority of these Fe-S families (7 out of 11) bind a [2Fe-2S] cluster, and are involved in energy production and conversion, or constitute additional components of the Fe-S biogenesis pathways, such as the A-family of proteins (like IscA, SufA, and ErpA), which are found in the 93 % of aerobes and only in the 14 % of anaerobes, or the BolA family, which is never found in anaerobes.

3.3. Fe-S Proteomes of Facultative-Anaerobic Prokaryotes

The fraction of Fe-S proteins encoded in facultative anaerobes is generally lower than that of anaerobes but larger than that of aerobes. On average, this represents 2.3 % $\pm$ 0.6 % of the genome (Table 3). *Thiobacillus denitrificans* and *Dechloromonas aromatica* have the highest fraction of Fe-S proteins in their genome (i.e., 3.5 %) whereas *Erwinia amylovora* encodes the lowest fraction (i.e., 1.3 %) among these organisms.

Facultative anaerobes have a percentage of [4Fe-4S] sites slightly higher than aerobes, and corresponding to 67 % $\pm$ 7 % (Table 3). Depending on organisms, these sites may range from 54 % (in *Serratia plymuthica*) to 80 % (in *Edwardsiella ictaluri*). Like aerobes, this group of organisms appears enriched in [2Fe-2S] proteins with respect to anaerobes, and the corresponding percentage ranges from 13 % (in *Edwardsiella ictaluri*) to 38 % (in *Methylibium petroleiphilum*); on average, the percentage of [2Fe-2S] in facultative anaerobes is 25 % $\pm$ 6 %. Fe-$(Cys)_4$ represents on average 5 % $\pm$ 2 %, whereas [3Fe-4S] clusters are on average 3 % $\pm$ 2 % (Table 3).

The Fe-S proteins encoded by these organisms generally map to a number of protein families that increases by increasing the genome size [1]. For example, facultative anaerobes having between 2000 and 2500 genes in their genome contain on average 45 $\pm$ 4 Fe-S families, whereas those having a genome of between 4500 and 5000 genes contain on average 108 $\pm$ 23 Fe-S families.

Table 3. List of facultative-anaerobic organisms and their content of Fe-S proteins.

Organism	Proteome size	# Fe-S proteins	% [4Fe-4S]	% [2Fe-2S]	% [3Fe-4S]	% Fe-Cys_4
Haemophilus influenzae	1610	47	68	20	2	11
Pasteurella multocida	2012	58	73	18	1	9
Coxiella burnetii	2045	37	73	23	2	2
Actinobacillus suis	2248	57	73	19	1	7
Aggregatibacter actinomycetemcomitans	2255	55	66	23	1	10
Kangiella koreensis	2632	62	65	26	3	7
Thiobacillus denitrificans	2827	99	73	21	2	4
Maricaulis maris	3063	60	67	26	7	0
Tolumonas auensis	3130	84	62	27	3	8
Halomonas elongata	3473	66	57	35	5	2
Vibrio cholerae	3504	71	65	22	7	6
Erwinia amylovora	3565	48	59	28	7	7
Proteus mirabilis	3662	96	73	18	3	5
Magnetococcus MC1	3716	116	78	15	2	5
Edwardsiella ictaluri	3735	89	80	13	3	4
Ferrimonas balearica	3782	119	75	17	4	3
Simiduia agarivorans	3820	71	66	23	5	6
Aliivibrio salmonicida	3911	67	64	24	4	8
Escherichia coli	4140	124	78	16	2	4
Dechloromonas aromatica	4171	146	68	27	2	3
Shewanella oneidensis	4214	101	72	18	3	7
Chromobacterium violaceum	4405	83	71	23	4	3
Aeromonas salmonicida	4436	90	68	22	3	6
Methylibium petroleiphilum	4449	115	56	38	3	3
Cronobacter turicensis	4452	84	67	24	3	6
Xenorhabdus nematophila	4469	66	54	35	5	6
Dickeya dadantii	4549	101	67	26	3	4
Rhodobacter sphaeroides	4568	112	65	27	3	5
Aromatoleum aromaticum	4576	133	76	18	2	4

Organism	Proteome size	# Fe-S proteins	% [4Fe-4S]	% [2Fe-2S]	% [3Fe-4S]	% Fe-Cys$_4$
Photorhabdus luminescens	4683	74	61	32	3	4
Rubrivivax gelatinosus	4693	133	72	23	2	3
Alicycliphilus denitrificans	4696	87	68	29	3	0
Rhodopseudomonas palustris	4820	146	67	25	3	5
Colwellia psychrerythraea	4909	91	61	29	7	3
Serratia plymuthica	4991	93	54	35	4	7
Citrobacter koseri	5006	111	75	18	3	4
Xanthobacter autotrophicus	5035	127	56	38	3	3
Pseudovibrio FO	5467	97	59	32	5	3
Photobacterium profundum	5489	110	70	21	5	5
Klebsiella pneumoniae	5766	119	66	27	2	5
Pantoea At9b	5770	90	61	33	3	3
Shigella dysenteriae	6409	100	78	15	3	4
Azospirillum brasilense	7527	136	69	28	1	2
	Average		**67**	**25**	**3**	**5**
	Std Dev		**7**	**6**	**2**	**2**

87 Fe-S families are conserved in at least 30 % of facultative anaerobes: these organisms contain 47 out of the 51 (92 %) Fe-S families constituting the shared core, and 10 out of the 11 (91 %) Fe-S families specific to aerobes. In addition, facultative anaerobes contain 16 Fe-S families in common with anaerobes, whereas 13 of them are specific to this group of organisms [1]. We can thus generally describe facultative anaerobes as having the same frequently occurring Fe-S families as aerobic organisms, plus an additional set of 29 families. More than half of this additional set (i.e., 18 or 62 %) is involved in energy production and conversion [1].

3.4. Fe-S Proteomes of Eukaryotes

The fraction of Fe-S proteins encoded in eukaryotes is generally largely reduced with respect to prokaryotic organisms, as it represents on average 0.36 % $\pm$ 0.14 % of the genome (Table 4). Most eukaryotes share the common core of Fe-S proteins

Table 4. List of eukaryotic organisms and their content of Fe-S proteins.

Organisms	Proteome size	# Fe-S proteins	% [4Fe-4S]	% [2Fe-2S]	% [3Fe-4S]	% Fe-Cys_4
Chaetomium thermophilum	7182	52	63	36	2	0
Gallus gallus	17902	74	58	42	0	0
Drosophila melanogaster	21354	77	53	44	2	0
Bos taurus (Bovine)	24231	87	49	51	0	0
Canis familiaris	25524	79	58	42	0	0
Sus scrofa (Pig)	26149	78	54	46	0	0
Caenorhabditis elegans	27216	67	48	49	3	0
Rattus norvegicus (Rat)	29378	82	49	51	0	0
Arabidopsis thaliana	33393	175	42	54	2	2
Danio rerio (Zebrafish)	41374	110	55	45	0	0
Mus musculus (Mouse)	52654	95	44	56	0	0
	Average		**53**	**46**	**1**	**0**
	Std Dev		**6**	**6**	**1**	**0**

vertically inherited from the universal common ancestor, whereas there has been little recruitment of new Fe-S proteins in cellular processes. So, the increase of the eukaryotic proteome size was not followed by an increase in Fe-S protein usage, and their fraction was reduced [2]. *Chaetomium thermophilum*, a thermophilic filamentous fungus which has a high temperature tolerance (60 °C), has the highest fraction of Fe-S proteins in its genome (i.e., 0.7 %). Conversely, *Mus musculus* is the eukaryote with the lowest Fe-S fraction, and encodes only 0.2 % of Fe-S proteins.

In eukaryotes, the fraction of [4Fe-4S] sites is very similar to that of [2Fe-2S] sites, on average being 53 % $\pm$ 6 % and 46 % $\pm$ 6 %, respectively (Table 4). Therefore, eukaryotes have a larger content of [2Fe-2S] sites with respect to prokaryotes. In some organisms, such as *Bos taurus*, *Rattus norvegicus*, *Arabidopsis thaliana*, and *Mus musculus*, [2Fe-2S] sites are more common than [4Fe-4S] sites.

Within the cells of eukaryotes, Fe-S proteins are present in the mitochondria, the cytosol, and the nucleus. The most conserved Fe-S proteins in eukaryotes are those involved in the CIA and the ISC machinery, as well as those included in the respiratory chain. *Arabidopsis thaliana* also encodes Fe-S proteins involved in the SUF pathway, since this machinery is present in the chloroplasts of plants [100].

4. OCCURRENCE AND FUNCTION OF IRON-SULFUR PROTEINS IN HUMANS

4.1. Human Iron-Sulfur Proteins and Their Distribution within Cellular Compartments

Table 5 reports the list of all known or putative Fe-S proteins in humans; these map to 70 unique genes and correspond to about 0.35 % of the total genome [144, 145]. The most common Fe-S cluster types in humans are [4Fe-4S] and [2Fe-2S] clusters, which correspond to about 61 % (i.e. 55 sites) and 39 % (i.e., 37 sites) of the total sites. These proteins localize in the mitochondria, the cytosol, and the nucleus. [4Fe-4S] clusters are largely predominant in the nucleus, as cofactors of several proteins involved in DNA processing and maintenance, and are also preferred in the cytoplasm, where the [4Fe-4S]:[2Fe-2S] ratio is close to 2:1. On the other hand, within mitochondria the number of [2Fe-2S] is slightly larger than [4Fe-4S] sites, thus corresponding to an enrichment in [2Fe-2S] clusters in mitochondria with respect to the global average in the cell [144].

Almost 70 % of the human Fe-S proteins are involved in just six main cellular processes [144] and this corresponds to the Fe-S proteins in prokaryotes. 13 out of 70 Fe-S proteins are involved in DNA processing and maintenance. These enzymes, which bind only [4Fe-4S] clusters, are receiving an ever-increasing attention for the importance they play in physiological and pathological conditions (see the next section). Respiration includes nine Fe-S proteins, and is the second most populated process: Fe-S proteins participating in respiration mainly belong to complex I, and thus, are localized in the inner membrane of the mitochondrion; they also include some components of complex II and complex III. Proteins of the respiratory chain may bind more than one Fe-S cluster, like succinate dehydrogenase (SDHB) that binds all types (a [2Fe-2S], a [3Fe-4S], and a [4Fe-4S]) of clusters [146]. Other relevant processes are amino acid and nucleotide metabolism, which includes 6 Fe-S proteins, and ribosome function/tRNA modification, which include 5 Fe-S proteins. Both these processes mainly occur in the cytosol. The biogenesis and maturation of Fe-S proteins, which in humans involve two different yet connected mechanisms (see Section 2), cumulatively involves 14 Fe-S proteins. Eight Fe-S proteins take part in the ISC machinery, while 5 Fe-S proteins are in the CIA machinery. The human ISC machinery, located in the mitochondria has multiple links to proteins involved in cell apoptosis as well as in the regulation of response to oxidative stress. The CIA machinery has the role of maturating nuclear and cytoplasmic Fe-S proteins. The large majority of nuclear Fe-S proteins contains a LYR-type tripeptide motif [147]. This short motif has been proposed to steer substrate discrimination and guide delivery of Fe-S clusters from proteins of the assembly machinery to specific subsets of Fe-S client proteins [147, 148].

Table 5. List of human Fe-S proteins.

Gene name		Fe-S per subunit	Subcellular location	Biological process	Involvement in disease
AOX1	Aldehyde oxidase 1	2×[2Fe-2S]	cytoplasm	amino acid and nucleotide metabolism	
DPH1	2-(3-Amino-3-carboxypropyl) histidine synthase subunit 1	[4Fe-4S]	cytoplasm, nucleus	amino acid and nucleotide metabolism	Developmental delay with short stature, dysmorphic features, and sparse hair (DEDSSH)
DPH2	2-(3-Amino-3-carboxypropyl) histidine synthase subunit 2	[4Fe-4S]	cytoplasm	amino acid and nucleotide metabolism	
DPYD	Dihydropyrimidine dehydrogenase	4×[4Fe-4S]	cytoplasm	amino acid and nucleotide metabolism	Dihydropyrimidine dehydrogenase deficiency (DPYDD)
PPAT, GPAT	Amidophosphoribosyltransferase	[4Fe-4S]	cytoplasm	amino acid and nucleotide metabolism	
XDH	Xanthine dehydrogenase/oxidase	2×[2Fe-2S]	cytoplasm	amino acid and nucleotide metabolism	Xanthinuria 1
ABCB7, ATM1	ATP-binding cassette sub-family B member 7, mitochondrial	[2Fe-2S]	mitochondrion	CIA machinery	Anemia, sideroblastic, spinocerebellar ataxia (ASAT)
CIAPIN1	Anamorsin	2×[2Fe-2S]	nucleus, cytoplasm	CIA machinery	
CISD1, MitoNEET	CDGSH iron-sulfur domain-containing protein 1	[2Fe-2S]	mitochondrion, cytoplasm	CIA machinery	

NARFL, IOP1	Cytosolic iron-sulfur assembly component 3	2×[4Fe-4S]	cytoplasm	CIA machinery	
NUBP1, NBP35	Cytosolic Fe-S cluster assembly factor NUBP1	2×[4Fe-4S]	cytoplasm	CIA machinery	
NUBP2, CFD1	Cytosolic Fe-S cluster assembly factor NUBP2	[4Fe-4S]	nucleus, cytoplasm	CIA machinery	
BRIP1, FANCJ	Fanconi anemia group J protein	[4Fe-4S]	nucleus, cytoplasm	DNA maintenance	Breast cancer; Fanconi anemia complementation group J (FANCJ)
DDX11, CHLR1	ATP-dependent DNA helicase DDX11	[4Fe-4S]	nucleus	DNA maintenance	Warsaw breakage syndrome (WBRS)
DDX12P, CHLR2	Putative ATP-dependent RNA helicase DDX12	[4Fe-4S]	nucleus	DNA maintenance	
DNA2	DNA replication ATP-dependent helicase/nuclease DNA2	[4Fe-4S]	nucleus, mitochondrion	DNA maintenance	Progressive external ophthalmoplegia with mitochondrial DNA deletions, autosomal dominant, 6 (PEOA6)
ERCC2, XPD	General transcription and DNA repair factor IIH helicase subunit XPD (TFIIH subunit XPD)	[4Fe-4S]	cytoplasm, nucleus	DNA maintenance	Xeroderma pigmentosum complementation group D (XP-D); Trichothiodystrophy 1, photosensitive (TTD1); Cerebro-oculo-facio-skeletal syndrome 2 (COFS2)
MUTYH	Adenine DNA glycosylase (EC 3.2.2.31)	[4Fe-4S]	nucleus	DNA maintenance	Familial adenomatous polyposis 2 (FAP2); Gastric cancer (GASC)

Gene name		Fe-S per subunit	Subcellular location	Biological process	Involvement in disease
NTHL1	Endonuclease III-like protein 1 (hNTH1)	[4Fe-4S]	mitochondrion, nucleus	DNA maintenance	Familial adenomatous polyposis 3 (FAP3)
POLA1	DNA polymerase alpha catalytic subunit	[4Fe-4S]	nucleus	DNA maintenance	Pigmentary disorder, reticulate, with systemic manifestations, X-linked (PDR)
POLD1	DNA polymerase delta catalytic subunit	[4Fe-4S]	nucleus	DNA maintenance	Colorectal cancer 10 (CRCS10); Mandibular hypoplasia, deafness, progeroid features, and lipodystrophy syndrome (MDPL)
POLE	DNA polymerase epsilon catalytic subunit A	[4Fe-4S]	nucleus, cytoplasm	DNA maintenance	Colorectal cancer 12 (CRCS12); Facial dysmorphism, immunodeficiency, livedo, and short stature (FILS)
PRIM2	DNA primase large subunit	[4Fe-4S]	nucleoplasm	DNA maintenance	
REV3L	DNA polymerase zeta catalytic subunit	[4Fe-4S]	nucleus	DNA maintenance	
RTEL1	Regulator of telomere elongation helicase 1	[4Fe-4S]	nucleus	DNA maintenance	Dyskeratosis congenita, autosomal recessive, 5 (DKCB5); Pulmonary fibrosis, and/or bone marrow failure, telomere-related, 3 (PFBMFT3)
FDX1L, FDX2	Ferredoxin-2, mitochondrial	[2Fe-2S]	mitochondrion	ISC machinery	

GLRX5	Glutaredoxin-related protein 5, mitochondrial	[2Fe-2S]	mitochondrion	ISC machinery	Anemia, sideroblastic, 3, pyridoxine-refractory (SIDBA3); Spasticity, childhood-onset, with hyperglycinemia (SPAHGC)
ISCA1	Iron-sulfur cluster assembly 1 homolog	[2Fe-2S]/ [4Fe-4S]	mitochondrion	ISC machinery	Multiple mitochondrial dysfunctions syndrome 5 (MMDS5)
ISCA2	Iron-sulfur cluster assembly 2 homolog	[2Fe-2S]/ [4Fe-4S]	mitochondrion	ISC machinery	Multiple mitochondrial dysfunctions syndrome 4 (MMDS4)
ISCU	Iron-sulfur cluster assembly enzyme ISCU, mitochondrial	[2Fe-2S]	mitochondrion, cytoplasm, nucleus	ISC machinery	Myopathy with exercise intolerance Swedish type (MEIS)
NFU1	NFU1 iron-sulfur cluster scaffold homolog, mitochondrial	[4Fe-4S]	mitochondrion, nucleus, cytoplasm	ISC machinery	Multiple mitochondrial dysfunctions syndrome 1 (MMDS1)
NUBPL, IND1	Iron-sulfur protein NUBPL (IND1 homolog)	[2Fe-2S]/ [4Fe-4S]	mitochondrion	ISC machinery	Mitochondrial complex I deficiency (MT-C1D)
NFS1	Cysteine desulfurase, mitochondrial (EC 2.8.1.7)	[2Fe-2S]	mitochondrion	ISC machinery	
CISD2, MINER1	CDGSH iron-sulfur domain-containing protein 2	[2Fe-2S]	mitochondrion, cytoplasm	other processes	Wolfram syndrome 2 (WFS2)
ACO1, IRP1	Cytoplasmic aconitate hydratase (aconitase)	[4Fe-4S]	cytoplasm	other processes	
ACO2	Aconitate hydratase, mitochondrial	[4Fe-4S]	mitochondrion	other processes	Infantile cerebellar-retinal degeneration (ICRD)

Gene name		Fe-S per subunit	Subcellular location	Biological process	Involvement in disease
IREB2	Iron-responsive element-binding protein 2	[4Fe-4S]	cytoplasm	other processes	
ABAT	4-Aminobutyrate aminotransferase	[2Fe-2S]	mitochondrion	other processes	GABA transaminase deficiency (GABATD)
FDX1	Adrenodoxin	[2Fe-2S]	mitochondrion	other processes	
GLRX3, PICOT	Glutaredoxin-3	[2Fe-2S]	cytoplasm	other processes	
LIAS	Lipoyl synthase	2×[4Fe-4S]	mitochondrion	other processes	Hyperglycinemia, lactic acidosis, and seizures (HGCLAS)
MOCS1	Molybdenum cofactor biosynthesis protein 1	2×[4Fe-4S]	cytoplasm, nucleus	other processes	Molybdenum cofactor deficiency, complementation group A (MOCODA)
RSAD1	Radical S-adenosyl methionine domain-containing protein 1	[4Fe-4S]	mitochondrion	other processes	
RSAD2	Radical S-adenosyl methionine domain-containing protein 2	[4Fe-4S]	mitochondrion	other processes	
GLRX2	Glutaredoxin-2, mitochondrial	[2Fe-2S]	mitochondrion nucleus	other processes	
CMAHP	Inactive cytidine monophosphate-N-acetylneuraminic acid hydroxylase	[2Fe-2S]	cytoplasm	other processes	
FECH, HEMH	Ferrochelatase	[2Fe-2S]	mitochondrion	other processes	Protoporphyria, erythropoietic, 1 (EPP1)
NDUFS1	NADH-ubiquinone oxidoreductase 75 kDa subunit,	2×[4Fe-4S], [2Fe-2S]	mitochondrion	respiratory chain	Mitochondrial complex I deficiency (MT-C1D)

NDUFS2	NADH dehydrogenase [ubiquinone] Fe-S protein 2	[4Fe-4S]	mitochondrion	respiratory chain	Mitochondrial complex I deficiency (MT-C1D)
NDUFS7	NADH dehydrogenase [ubiquinone] Fe-S protein 7	[4Fe-4S]	mitochondrion	respiratory chain	Leigh syndrome (LS); Mitochondrial complex I deficiency (MT-C1D)
NDUFS8	NADH dehydrogenase [ubiquinone] Fe-S protein 8	2×[4Fe-4S]	mitochondrion	respiratory chain	Leigh syndrome (LS)
NDUFV1	NADH dehydrogenase [ubiquinone] flavoprotein 1	[4Fe-4S]	mitochondrion	respiratory chain	Leigh syndrome (LS); Mitochondrial complex I deficiency (MT-C1D)
NDUFV2	NADH dehydrogenase [ubiquinone] flavoprotein 2	[2Fe-2S]	mitochondrion	respiratory chain	
SDHB	Succinate dehydrogenase [ubiquinone] Fe-S subunit	[2Fe-2S], [3Fe-4S], [4Fe-4S]	mitochondrion	respiratory chain	Pheochromocytoma (PCC); Paragangliomas 4 (PGL4); Paraganglioma and gastric stromal sarcoma (PGGSS)
UQCRFS1	Cytochrome *b-c*1 complex subunit Rieske	[2Fe-2S]	mitochondrion	respiratory chain	
ETFDH	Electron transfer flavoprotein-ubiquinone oxidoreductase	[4Fe-4S]	mitochondrion	respiratory chain	Glutaric aciduria 2C (GA2C)
ABCE1	ATP-binding cassette sub-family E member 1	2×[4Fe-4S]	cytoplasm, mitochondrion	ribosome function and tRNA modification	
CDK5RAP1	CDK5 regulatory subunit-associated protein 1	2×[4Fe-4S]	cytoplasm	ribosome function and tRNA modification	
CDKAL1	Threonylcarbamoyladenosine tRNA methylthiotransferase	2×[4Fe-4S]	cytoplasm	ribosome function and tRNA modification	Diabetes mellitus, non-insulin-dependent (NIDDM)

Gene name		Fe-S per subunit	Subcellular location	Biological process	Involvement in disease
ELP3	Elongator complex protein 3	[4Fe-4S]	cytoplasm, nucleus	ribosome function and tRNA modification	ELP3 genetic variations may be associated with an increased risk for neurodegeneration and motor neuron diseases
TYW1, TYW1B	*S*-adenosyl-L-methionine-dependent tRNA 4-demethylwyosine synthase	[4Fe-4S]	cytoplasm	ribosome function and tRNA modification	
CISD3, MINER2	CDGSH iron-sulfur domain-containing protein 3	2×[2Fe-2S]	mitochondrion	unknown	
BOLA1	BolA-like protein 1	[2Fe-2S]	mitochondrion	unknown	
BOLA2	BolA-like protein 2	[2Fe-2S]	cytoplasm	unknown	
BOLA3	BolA-like protein 3	[2Fe-2S]	mitochondrion	unknown	Multiple mitochondrial dysfunctions syndrome 2 with hyperglycinemia (MMDS2)
GRXCR1	Glutaredoxin domain-containing cysteine-rich protein 1	[2Fe-2S]*	unknown	unknown	Deafness, autosomal recessive, 25 (DFNB25)
NARF	Nuclear prelamin A recognition factor	2×[4Fe-4S]	nucleus	unknown	
RFESD	Rieske domain-containing protein	[2Fe-2S]*	unknown	unknown	
AIFM3	Apoptosis-inducing factor 3	[2Fe-2S]*	mitochondrion	unknown	
UQCRFS1P1	Putative cytochrome *b-c*1 complex subunit Rieske	[2Fe-2S]*	unknown	unknown	

* These are putative cofactors

The 70 Fe-S human genes are associated with 47 Fe-S families. Eighteen of these families (38 %) are part of the core of proteins also shared by aerobic and anaerobic prokaryotes; 8 (17 %) of these families are mainly conserved in anaerobic organisms while 6 (13 %) are mainly found in aerobes; finally, 15 (32 %) families are rarely or never detected in prokaryotic organisms and thus can be regarded as Fe-S families specific to eukaryotes and/or to humans [1]. In the large majority of cases, human Fe-S families related to the shared core and to aerobe-specific families have mitochondrial localization. This is presumably due to the prokaryotic origin of the mitochondrion so that these proteins have a relatively short evolutionary distance. Instead, human Fe-S families related to anaerobic organisms mainly have nuclear and/or cytosolic localization. As described in the previous sections, anaerobes have a strong preference for [4Fe-4S] clusters, whereas aerobe-specific families prefer [2Fe-2S] clusters. Therefore, these relationships between human and prokaryotic families provide a rationale for the differential distribution of Fe-S cluster types in the human cell.

At the functional level, mitochondrial Fe-S proteins that belong to the families present in prokaryotes are predominantly involved in energy production and conversion or in Fe-S protein biogenesis. On the other hand, the most common process for the families inherited from anaerobes is ribosome/tRNA modification.

4.2. An Emerging Role of Iron-Sulfur Clusters: DNA Maintenance

In recent years, an ever-increasing number of Fe-S enzymes with DNA processing roles have been identified. Initially, the discovery of Fe-S cofactors in the vicinity of DNA caused perplexity within the scientific community because DNA may be damaged by iron whose redox activity can generate free radicals and other strongly oxidizing species (Fenton reactions) [149]. However, this view changed with the discovery of an Fe-S cluster present in DNA polymerase subunits (POLD1, POLA1, and POLE1), DNA helicases (XPD, FANCJ, RTEL1), DNA glycosylases (MUTYH and NTHL1), the nuclease DNA2 and the DNA primase PRIM2. As our knowledge of Fe-S containing DNA processing enzymes increases, we are discovering that these cofactors are critical to the activity of these enzymes and irreplaceable with other cofactors [150].

In 2014, Jacqueline K. Barton at the California Institute of Technology suggested for the first time that Fe-S cofactors could help enzymes involved in DNA repair to recognize damaged nucleotides, by probing DNA integrity electronically [151–154]. Indeed, DNA is like a molecular wire which conducts electrons through the stacking interactions of the aromatic rings of each base pair, and thus is able to ferry charges over long distances. Charge transfer requires the intact DNA duplex and rapidly interrupts when there is a mismatch in the pairing or damage occurs to a DNA base. Fe-S clusters in DNA processing enzymes have a unique key role because they are able to interrogate DNA integrity by donating and receiving electrons that travel through the DNA duplex [151–154].

In case the electronic signal is interrupted, DNA is damaged, thus Fe-S clusters coordinate these enzymes by inducing conformational changes, interactions, and biochemical activities that finally orchestrate replication and repair steps [150]. Even if at present this complex system of signaling events and machineries interaction is still unknown, the communication between Fe-S enzymes through the DNA duplex can be seen as a novel and alternative way to the classical protein–protein interaction signaling pathways. The efficiency of this electronic communication from the DNA to the protein strictly depends upon the distance between DNA and the Fe-S cluster, as well as on the composition of the protein matrix: some aromatic residues, like tyrosine and tryptophan, can facilitate electron transfer within the protein [152, 155]. Fe-S clusters in different DNA processing enzymes may be more or less solvent-exposed, and the architectures of these proteins may provide a means for different DNA charge transfer activity and pathway coordination.

A possible DNA charge transfer model of damage detection has been proposed for endonuclease III (NTH1 in humans) and adenine DNA glycosylase (MutY) [156, 157]. The first enzyme repairs hydroxylated pyrimidines, while the second removes the inappropriately paired adenine bases from the DNA backbone. Both of them bind a [4Fe-4S] cluster whose redox potential is highly positive and outside the range of physiological redox activity; the redox potential shifts –200 mV into the physiological range upon DNA binding, allowing the Fe-S cluster to become an active redox switch. DNA charge transfer is initiated by a guanine cation radical oxidizing a nearby MutY Fe-S cluster which increases its affinity for DNA. When a new glycosylase (endonuclease III or MutY) binds DNA in another position (also distant), it passes the electron through DNA conduction to the initial MutY, thus reducing the enzyme which detaches from the DNA. This long range charge transfer can occur only if DNA is undamaged: on the contrary, if DNA integrity is compromised the second protein is unable to reduce the first MutY and the two proteins remain attached to the double helix in proximity of the damage, thus promoting the increase of glycosylase concentration around the DNA mismatch for subsequent damage detection [150]. It was demonstrated that the ability of endonuclease III to perform DNA charge transfer is strictly connected to its capacity in localizing the DNA damage. Further, it was shown that MutY activity is aided by the charge transfer activity of endonuclease III [157].

4.3. Iron-Sulfur Proteins and Human Diseases

Numerous diseases including several neurodegenerative and hematological disorders have been associated with defects in Fe-S protein biogenesis, underlining the central importance of this process for life [72, 113, 158, 159]. At the moment, human diseases have been linked to mutations in genes coding for proteins involved in the mitochondrial Fe-S protein biogenesis. Indeed, although a number of proteins have been identified as components of the CIA machinery, many have been discovered or characterized in depth over the last 5 years. For this

reason, many of these proteins have likely not yet been linked to disease conditions.

4.3.1. Diseases Associated with Mitochondrial Fe-S Protein Biogenesis

Friedreich's ataxia (FRDA) is a rare genetic neurodegenerative disease that is by far the most common disease linked to Fe-S cluster biosynthesis [160–164]. FRDA patients carry a mutation in the first intron of the frataxin gene, which leads to transcriptional silencing of frataxin thereby resulting in severely decreased levels of the frataxin protein. Other nonsense and missense mutations or deletions also result in the same symptoms for patients. No therapy for FRDA exists today, despite much effort to increase frataxin levels or mitochondrial function, or alleviating secondary symptoms by antioxidants or iron chelator strategies [165]. In the class of the neurodegenerative disorders including, e.g., Parkinson's disease, Alzheimer's disease, and Huntington's disease, Parkinson's disease has most often been associated with iron-sulfur cluster proteins. Recent works suggested that patients with Parkinson's have a defect in ISC synthesis, partial inhibition of complex I of the respiratory chain, and increased iron levels in the substantia nigra, due to erroneous regulation of iron homeostasis by the iron regulatory protein 1 (IRP1) [166]. However, it remains to be verified if there is a direct connection between Fe-S protein biogenesis and these pathogenic conditions.

Multiple mitochondrial dysfunction syndrome (MMDS) describes another class of fatal mitochondrial diseases caused by mutations in proteins involved in mitochondrial Fe-S protein biogenesis. To date, four types of MMDS have been identified: MMDS1 due to mutations on Nfu1, MMDS2 for mutations with BOLA3, MMDS3 with mutations on Iba57, and MMDS4 with mutations on IscA2 [167]. Recently, an IscA1 variant has been described that results in MMDS-like symptoms, which may expand this class of disorders [168]. All MMDS patients exhibit similar symptoms that include metabolic acidosis with hyperglycemia, and deficiency of respiratory complexes and lipoic acid-bound enzymes, such as pyruvate dehydrogenase and α-ketogluterate dehydrogenase. Lipoate biogenesis relies on the mitochondrial [4Fe-4S] enzyme lipoate synthase.

Other mitochondrial diseases related to proteins involved in mitochondrial Fe-S protein biogenesis are myopathies, characterized by severe exercise intolerance, tachycardia, fatigue and pain in active muscles. These kind of myopathies are caused by mutations in IscU [169] and ferredoxin 2 genes [170]. Recently, Isd11 and Nfs1, which are partners of IscU and ferredoxin 2 in the ISU system, have been also linked to human disease states resulting in a decrease of the deficiencies of respiratory chain complexes [171, 172]. Pathogenic mutations or perturbations of the *Ind1* gene have resulted in a specific deficiency of complex I, while other respiratory complexes remain intact [173].

Finally, there are disorders linked to the misregulation of iron that occur specifically in Fe-S cluster proteins. Whether this link is due to a simple abundance of iron or to a more direct connection between Fe-S protein biogenesis and iron regulation, perhaps via the IRP1 pathway, remains uncertain and requires fur-

ther research. Among these disorders, there are two kinds of sideroblastic anemias linked to mutations in the *Glrx5* and *ABCB7* genes [174, 175]. The X-linked sideroblastic anemia, caused by mutations in the mitochondrial transporter protein ABCB7, is characterized by early-onset ataxia, sideroblastic anemia and iron overload in affected tissues, primarily in the mitochondria, which suggests that the transporter is unable to function properly for trafficking of clusters in the cytosol and nucleus. Pathogenic mutations in the *Glrx5* gene also result in sideroblastic-like anemia and iron overload condition. This condition leads to a decrease of Glrx5 protein levels in patients but without a complete loss of the protein function. The decreased protein levels result in impaired cluster transfer to downstream Fe-S proteins. Aside from implications in sideroblastic anemia, patients with pathogenic mutations in Glrx5 have also been found to display variant non-ketotic hyperglycinemia, with symptoms similar to those observed for patients with MMDS2 due to mutations in BolA3 [176]. This is in agreement with a strictly related functional role of BolA3 and Glrx5, which are, indeed, protein partners [87].

4.3.2. Diseases Associated with Nuclear Fe-S Cluster-Containing Proteins

Nuclear Fe-S enzymes and corresponding mutations in their coding genes are often associated to pathological conditions, due to the importance of the roles that they have in DNA processing and maintenance. Among them, FANCJ, in conjunction with 15 other genes, is implicated in causing Fanconi anemia; a condition characterized by congenital defects, progressive bone marrow failure, susceptibility to cancer with chromosome breaks, and sensitivity to DNA-damaging agents that result in intra-strand cross-links [177]. Two mutations (Met299Ile and Ala349Pro) occur in the iron-sulfur cluster binding domain, where the Met299Ile mutation has been found in breast cancer patients, and the Ala349Pro mutation causes Fanconi anemia [178]. These mutational analyses showed the importance of the Fe-S cluster in the function of FANCJ and its role in downstream DNA-processing activities in disease. A second Fe-S cluster-containing helicase linked to a variety of pathogenic human disorders is XPD. To date, only one of these disorders has been linked to a mutation in the Fe-S cluster-binding domain [179]. An Arg112His mutation is present at a highly conserved residue that removes a hydrogen bond to a cluster-ligating cysteine, resulting in helicase inactivation. The last two Fe-S cluster helicases that have been implicated in disease much more recently are ChlR1 (or DDX11) and RTEL1. The first human patient with a defect in ChlR1 resulted in a condition termed Warsaw breakage syndrome, due to mutations in ChlR1, which perturb a conserved arginine residue in the Fe-S cluster-binding domain [180, 181]. Mutations of RTEL1 have been implicated in premature aging disorders, which are characterized by multi-system bone marrow failure, due to impaired telomere maintenance [182]. One mutation (Glu275Lys) has been found in the Fe-S cluster domain, but the biochemical characterization and impact of this residue on the cluster has not yet been examined.

The human protein MUTYH, a member of the glycosylase family which binds a [4Fe-4S] cluster, has been associated with a number of disease variants that result in cancer. To date, one of the mutations (Pro281Leu) was found at a highly conserved proline in the Fe-S cluster-binding domain, which impairs the ability to bind substrate DNA [183].

A detailed understanding in examining the molecular mechanisms associated to this nuclear Fe-S DNA-binding protein family will provide assistance to clinicians in patient-screening and treatment options. In particular, biochemical characterization is required to understand to determine the role of the cluster in relation to disease phenotypes.

5. FUTURE DIRECTIONS

It is plausible that future research on Fe-S proteins will focus on two major topics. The first will concern the clarification of some important parts of the Fe-S protein biogenesis and their relationship with pathological conditions. All Fe-S machineries have still many unanswered questions regarding the interactions between the individual proteins and the formation of complexes as well as the delivery of clusters into the recipient apo-proteins. More specific questions regard, for example, the source of iron and electrons for Fe-S cluster synthesis on protein scaffolds and the function of the A-type family of proteins (e.g., IscA, SufA) in prokaryotic ISC and SUF systems, as well as how the [4Fe-4S] cluster assembled in the mitochondrial ISC machinery is inserted into the final client proteins has to be understood. Also the CIA machinery represents an important subject in the field of eukaryotic Fe-S biogenesis: several players of this system as well as their mechanism of action remain to be elucidated. Furthermore, the relationship between the ISC and the CIA machinery remains to be clarified, as well as the identity of the X-S compound that is exported from the mitochondrion to the cytoplasm for the synthesis of cytosolic Fe-S clusters. The role of the BolA protein family in eukaryotic CIA and ISC machineries represents another issue which needs further investigation. While it has been shown that, in the cytosol, BolA2 plays a role in [2Fe-2S] cluster delivery, in mitochondria the role of BolA1 and BolA3 is still not well understood; potentially they can also work in [2Fe-2S] cluster delivery, but this has not been determined yet. We expect that future investigations will lead to a more detailed mechanistic picture of cellular Fe-S protein biogenesis allowing to create the basis for a molecular understanding of Fe-S associated diseases. Finally, another under-explored area concerns the regulation of the assembly processes by environmental factors and the repair mechanisms of damaged Fe-S clusters.

The second topic regards the analysis of the roles that Fe-S clusters play in proteins involved in DNA repair, replication, and translation. Resolving questions centered around the molecular function of Fe-S enzymes involved in the maintenance of genome integrity has become a major focus of research. This area will require structural, biochemical, and genetic studies in order to define which molecular machines are involved in DNA processing and how they com-

municate with each other. In particular, we expect that ongoing structural and biophysical elucidations will take our understanding of nanoscale DNA assemblies and their control to the next level of molecular circuitry coordinating DNA replication and repair. This knowledge will lead us not only to clarify a critical part of cell biology but above all, to understand the roles that these enzymes and their cofactors play in the development of pathologies to facilitate the development of new therapies.

ABBREVIATIONS AND DEFINITIONS

ACP	acyl carrier protein
ADP	adenosine 5′-diphosphate
ATP	adenosine 5′-triphosphate
CIA	cytosolic iron-sulfur assembly
FAD	flavin adenine dinucleotide
Fdx	ferredoxin
FMN	flavin mononucleotide
FRDA	Friedreich's ataxia
Fxn	frataxin
Glrx	monothiol glutaredoxin
GTP	guanosine 5′-triphosphate
HiPIP	high potential iron protein
IRP1	cytosolic iron regulatory protein 1
ISC	iron-sulfur cluster machinery
MMDS	multiple mitochondrial dysfunction syndrome
NADPH	nicotinamide adenine dinucleotide reduced
NIF	nitrogen fixation
PDB	Protein Data Bank
ROS	reactive oxygen sepcies
SAM	*S*-adenosylmethionine
SUF	sulfur mobilization
tRNA	transfer RNA

REFERENCES

1. C. Andreini, A. Rosato, L. Banci, *PLoS ONE*, **2017**, *12*, e0171279.
2. C. Andreini, L. Banci, I. Bertini, S. Elmi, A. Rosato, *Proteins, Struct. Funct. Bioinf.* **2007**, *67*, 317–324.
3. R. Venkateswara, R. H. Holm, *Chem. Rev.* **2004**, 104, 527–559.
4. J. E. Harmer, M. J. Hiscox, P. C. Dinis, S. J. Fox, A. Iliopoulos, J. E. Hussey, J. Sandy, F. T. Van Beek, J. W. Essex, P. L. Roach, *Biochem. J.* **2014**, *464*, 123–133.
5. H. Beinert, *J. Biol. Inorg. Chem.* **2000**, *5*, 2–15.
6. J. Meyer, *J. Biol. Inorg. Chem.* **2008**, *13*, 157–170.
7. *Handbook on Metalloproteins*, Eds I. Bertini, A. Sigel, H. Sigel, Marcel Dekker, New York, 2001, pp 1–1182.

8. R. J. P. Williams, *Coord. Chem. Rev.* **2001**, *216–217*, 583–595.
9. J. A. Imlay, *Mol. Microbiol*. **2006**, *59*, 1073–82.
10. G. Wächtershauser, *Chem. Biodivers.* **2007**, *4*, 584–602.
11. ztG. Mitou, C. Higgins, P. Wittung-Stafshede, R. C. Conover, A. D. Smith, M. K. Johnson, J. Gaillard, A. Stubna, E. Münck, J. Meyer, *Biochemistry* **2003**, *42*, 1354–64.
12. D. W. Bak, S. J. Elliott, *Curr. Opin. Chem. Biol.* **2014**, *19*, 50–58.
13. V. Putignano, A. Rosato, L. Banci, C. Andreini, *Nucleic Acids Res.* **2018**, *46*(D1), D459–D464, doi: 10.1093/nar/gkx989.
14. C. Andreini, G. Cavallaro, S. Lorenzini, A. Rosato, *Nucleic Acids Res.* **2013**, *41* (Database issue), D312–D319.
15. P. W. Rose, A. Prlic, A. Altunkaya, C. Bi, A. R. Bradley, C. H. Christie, L. D. Costanzo, J. M. Duarte, S. Dutta, Z. Feng, R. K. Green, D. S. Goodsell, B. Hudson, T. Kalro, R. Lowe, E. Peisach, C. Randle, A. S. Rose, C. Shao, Y. P. Tao, Y. Valasatava, M. Voigt, J. D. Westbrook, J. Woo, H. Yang, J. Y. Young, C. Zardecki, H. M. Berman, S. K. Burley, *Nucleic Acids Res.* **2017**, *45*, D271–D281.
16. C. Andreini, I. Bertini, G. Cavallaro, G. L. Holliday, J. M. Thornton, *Bioinformatics* **2009**, *25*, 2088–2089.
17. C. Andreini, I. Bertini, G. Cavallaro, G. L. Holliday, J. M. Thornton, *J. Biol. Inorg. Chem.* **2008**, *13*, 1205–1218.
18. Y. Wu, R. Wu, D. Mandalapu, X. Ji, T. Chen, W. Ding, Q. Zhang, *Org. Biomol. Chem.* **2019**, *17*, 1809–1812.
19. L. E. Mortenson, R. C. Valentine, J. E. Carnahan, *Biochem. Biophys. Res. Commun.* **1962**, *7*, 448–452.
20. E. T. Adman, L. C. Sieker, L. H. Jensen, *J. Biol. Chem.* **1973**, *248*, 3987–3996.
21. L. Liu, T. Nogi, M. Kobayashi, T. Nozawa, K. Miki, *Acta Cryst. D Biol.* **2002**, *58*, 1085–91.
22. R. G. Bartsch, *Methods Enzymol.* **1978**, *53*, 329–340.
23. M. I. McLaughlin, N. D. Lanz, P. J. Goldman, K. H. Lee, S. J. Booker, C. L. Drennan, *Proc. Natl. Acad. Sci. USA* **2016**, *113*, 9446–9450.
24. I. Bertini, C. Luchinat, A. Provenzani, A. Rosato, P. R. Vasos, *Proteins: Struct. Funct. Genet.* **2002**, *46*, 110–127.
25. K. Tagawa, D. I. Arnon, *Nature* **1962**, *195*, 537–543.
26. T. Tsukihara, K. Fukuyama, M. Nakamura, Y. Katsube, N. Tanaka, M. Kakudo, K. Wasa, T. Hase, H. Matsubara, *J. Biochem.* **1981**, 90, 1763–1773.
27. S. Iwata, M. Saynovitis, T. A. Link, H. Michel, *Structure* **1996**, *4*, 567–579.
28. F. Berkovitch, Y. Nicolet, J. T. Wan, J. T. Jarrett, C. L. Drennan, *Science* **2004**, *303*, 76–79.
29. J. Frazzon, D. R. Dean, *Curr. Opin. Chem. Biol.* **2003**, *7*, 166–73.
30. M. Fontecave, S. Ollagnier-de Choudens, *Arch. Biochem. Biophys* **2008**, *474*, 226–237.
31. B. Blanc, C. Gerez, S. Ollagnier-de Choudens, *Biochim. Biophys. Acta* **2015**, *1853*, 1436–1447.
32. L. Loiseau, S. Ollagnier-de Choudens, L. Nachin, M. Fontecave, F. Barras, *J. Biol. Chem.* **2003**, *278*, 38352–38359.
33. L. Zheng, R. H. White, V. L. Cash, D. R. Dean, *Biochemistry* **1994**, *33*, 4714–4720.
34. L. Zheng, V. L. Cash, D. H. Flint, D. R. Dean, *J. Biol. Chem.* **1998**, *273*, 13264–13272.
35. H. D. Urbina, J. J. Silberg, K. G. Hoff, L. E. Vickery, *J. Biol. Chem* **2001**, *276*, 44521–44526.
36. W. H. Tong, G. N. Jameson, B. H. Huynh, T. A. Rouault, *Proc. Natl. Acad. Sci. USA* **2003**, *100*, 9762–9767.
37. W. H. Tong, T. A. Rouault, *Cell Metab.* **2006**, *3*, 199–210.

38. M. R. Jacobson,K. E. Brigle, L. T. Bennett, R. A. Setterquist, M. S. Wilson, V. L. Cash, J. Beynon, W. E. Newton, D. R. Dean, *J. Bacteriol.* **1989**, *171*, 1017–1027.
39. J. N. Agar, C. Krebs, J. Frazzon, B. H. Huynh, D. R. Dean, M. K. Johnson, *Biochemistry* **2000**, *39*, 7856–7862.
40. K. Chandramouli, M. C. Unciuleac, S. Naik, D. R. Dean, B. H. Huynh, M. K. Johnson, *Biochemistry* **2007**, *46*, 6804–6811.
41. C. J. Schwartz, O. Djaman, J. A. Imlay, P. J. Kiley, *Proc. Natl. Acad. Sci. USA* **2000**, *97*, 9009–9014.
42. E. N. Marinoni, J. S. de Oliveira, Y. Nicolet, E. C. Raulfs, P. Amara, D. R. Dean, J. C. Fontecilla-Camps, *Angew. Chem. Int. Ed.* **2012**, *51*, 5439–5442.
43. Y. Kakuta, T. Horio, Y. Takahashi, K. Fukuyama, *Biochemistry* **2001**, *40*, 11007–11012.
44. H. Lange, A. Kaut, G. Kispal, R. Lill, *Proc. Natl. Acad. Sci. USA* **2000**, *97*, 1050–1055.
45. D. S. Li, K. Ohshima, S. Jiralerspong, M. W. Bojanowski, M. Pandolfo, *FEBS Lett.* **1999**, *456*, 13–16.
46. E. Vivas, E. Skovran, D. M. Downs, *J. Bacteriol.* **2006**, *188*, 1175–1179.
47. C. L. Tsai, D. P. Barondeau, *Biochemistry* **2010**, *49*, 9132–9139.
48. S. Adinolfi, C. Iannuzzi, F. Prischi, C. Pastore, S. Iametti, S. R. Martin, F. Bonomi, A. Pastore, *Nature Struct. Mol. Biol.* **2009**, *16*, 390–396.
49. J. Bridwell-Rabb, N. G. Fox, C. L. Tsai, A. M. Winn, D. P. Barondeau, *Biochemistry* **2014**, *53*, 4904–4913.
50. F. Prischi, P. Konarev, C. Iannuzzi, C. Pastore, S. Adinolfi, S. R. Martin, D. I. Svergun, A. Pastore, *Nature Commun.* **2010**, *1*, 95. doi: 10.1038/ncomms1097.
51. J. H. Kim, J. R. Bothe, R. O. Frederick, J. C. Holder, J. L. Markley, *J. Am. Chem. Soc.* **2014**, *136*, 7933–7942.
52. R. Yan, P. V. Konarev, C. Iannuzzi, S. Adinolfi, B. Roche, G. Kelly, L. Simon, S. R. Martin, B. Py, F. Barras, D. I. Svergun, A. Pastore, *J. Biol. Chem.* **2013**, *288*, 24777–24787.
53. K. Cai, R. O. Frederick, J. H. Kim, N. M. Reinen, M. Tonelli, J. L. Markley, *J. Biol. Chem.* **2013**, *288*, 28755–28770.
54. F. Bonomi, S. Iametti, A. Morleo, D. Ta, L. E. Vickery, *Biochemistry* **2008**, *47*, 12795–12801.
55. F. Bonomi, S. Iametti, A. Morleo, D. Ta, L. E. Vickery, *Biochemistry* **2011**, *50*, 9641–9650.
56. D. Vinella, C. Brochier-Armanet, L. Loiseau, E. Talla, F. Barras, *PLoS Genet.* **2009**, *5*, e1000497.
57. D. Vinella, L. Loiseau, S. Ollagnier-de Choudens, M. Fontecave, F. Barras, *Mol. Microbiol.* **2013**, *87*, 493–508.
58. B. Py, C. Gerez, S. Angelini, R. Planel, D. Vinella, L. Loiseau, E. Talla, C. Brochier-Armanet, S. R. Garcia, J. M. Latour, S. Ollagnier-de Choudens, M. Fontecave, F. Barras, *Mol. Microbiol.* **2012**, *86*, 155–171.
59. G. Tan, Z. Cheng, Y. Pang, A. P. Landry, J. Li, J. Lu, H. Ding, *Mol. Microbiol.* **2014**, *93*, 629–644.
60. G. Tan, J. Lu, J. P. Bitoun, H. Huang, H. Ding, *Biochem. J.* **2009**, *420*, 463–472.
61. C. Krebs, J. N. Agar, A. D. Smith, J. Frazzon, D. R. Dean, B. H. Huynh, M. K. Johnson, *Biochemistry* **2001**, *40*, 14069–14080.
62. S. Ollagnier-de Choudens, L. Nachin, Y. Sanakis, L. Loiseau, F. Barras, M. Fontecave, *J. Biol. Chem* **2003**, *278*, 17993–18001.
63. S. Ollagnier-de-Choudens, Y. Sanakis, M. Fontecave, *J. Biol. Inorg. Chem.* **2004**, *9*, 828–838.

64. F. Vilella, R. Alves, M. T. Rodriguez-Manzaneque, G. Belli, S. Swaminathan, P. Sunnerhagen, E. Herrero, *Comp. Funct. Genomics* **2004**, *5*, 328–341.
65. K. D. Kim, W. H. Chung, H. J. Kim, K. C. Lee, J. H. Roe, *Biochem. Biophys. Res. Commun.* **2010**, *392*, 467–472.
66. D. T. Mapolelo, B. Zhang, S. Randeniya, A. N. Albetel, H. Li, J. Couturier, C. E. Outten, N. Rouhier, M. K. Johnson, *Dalton Trans.* **2013**, *42*, 3107–3115.
67. M. T. Rodriguez-Manzaneque, J. Tamarit, G. Belli, J. Ros, E. Herrero, *Mol. Biol. Cell* **2002**, *13*, 1109–1121.
68. S. Angelini, C. Gerez, S. Ollagnier-de Choudens, Y. Sanakis, M. Fontecave, F. Barras, B. Py, *J. Biol. Chem.* **2008**, *283*, 14084–14091.
69. Y. Takahashi, M. Nakamura, *J. Biochem.* **1999**, *126*, 917–926.
70. J. L. Giel, D. Rodionov, M. Liu, F. R. Blattner, P. J. Kiley, *Mol. Microbiol.* **2006**, *60*, 1058–1075.
71. M. Müller, W. Martin, *Bioessays* **1999**, *21*, 377–381.
72. O. Stehling, C. Wilbrecht, R. Lill, *Biochimie* **2014**, *100*, 61–77.
73. O. Stehling, R. Lill, *Cold Spring Harb. Perspect. Biol.* **2013**, *5*, a011312.
74. D. J. Netz, J. Mascarenhas, O. Stehling, A. J. Pierik, R. Lill, *Trends Cell. Biol.* **2014**, *24*, 303–312.
75. G. Kispal, P. Csere, C. Prohl, R. Lill, *EMBO J.* **1999**, *18*, 3981–3989.
76. C. A. Chen, J. A. Cowan, *Biochim. Biophys. Acta* **2006**, *1760*, 1857–1865.
77. M. T. Boniecki, S. A. Freibert, U. Mühlenhoff, R. Lill, M. Cygler, *Nature Commun.* **2017**, *8*, 1287. doi: 10.1038/s41467–017–01497–1.
78. N. Wiedemann, E. Urzica, B. Guiard, H. Müller, C. Lohaus, H. E. Meyer, M. T. Ryan, C. Meisinger, U. Mühlenhoff, R. Lill, N. Pfanner, *EMBO J.* **2006**, *25*, 184–195.
79. Y. Shan, E. Napoli, G. Cortopassi, *Hum. Mol. Genet.* **2007**, *16*, 929–941.
80. A. C. Adam, C. Bornhovd, H. Prokisch, W. Neupert, K. Hell, *EMBO J.* **2006**, *25*, 174–183.
81. C. Johansson, A. K. Roos, S. J. Montano, R. Sengupta, P. Filippakopoulos, K. Guo, F. von Delft, A. Holmgren, U. Oppermann, K. L. Kavanagh, *Biochem. J.* **2011**, *433*, 303–311.
82. U. Mühlenhoff, J. Gerber, N. Richhardt, R. Lill, *EMBO J.* **2003**, *22*, 4815–4825.
83. R. Dutkiewicz, J. Marszalek, B. Schilke, E. A. Craig, R. Lill, U. Mühlenhoff, *J. Biol. Chem.* **2006**, *281*, 7801–7808.
84. S. Bandyopadhyay, K. Chandramouli, M. K. Johnson, *Biochem. Soc. Trans.* **2008**, *36*, 1112–1119.
85. M. A. Uzarska, R. Dutkiewicz, S. A. Freibert, R. Lill, U. Mühlenhoff, *Mol. Biol. Cell* **2013**, *24*, 1830–1841.
86. M. A. Uzarska, V. Nasta, B. D. Weiler, F. Spantgar, S. Ciofi-Baffoni, M. Saviello, L. Gonnelli, U. Mühlenhoff, L. Banci, R. Lill, *Elife* **2016**, *5*, e16673.
87. V. Nasta, A. Giachetti, S. Ciofi-Baffoni, L. Banci, *Biochim. Biophys. Acta* **2017**, *1861*, 2119–2131.
88. S. Sen, B. Rao, C. Wachnowsky, J. A. Cowan, *Metallomics* **2018**, *10*, 1282–1290.
89. U. Mühlenhoff, N. Richter, O. Pines, A. J. Pierik, R. Lill, *J. Biol. Chem.* **2011**, *286*, 41205–41216.
90. A. D. Sheftel, C. Wilbrecht, O. Stehling, B. Niggemeyer, H. P. Elsasser, U. Mühlenhoff, R. Lill, *Mol. Biol. Cell* **2012**, *23*, 1157–1166.
91. L. Banci, D. Brancaccio, S. Ciofi-Baffoni, R. Del Conte, R. Gadepalli, M. Mikolajczyk, S. Neri, M. Piccioli, J. Winkelmann, *Proc. Natl. Acad. Sci. USA* **2014**, *111*, 6203–6208.
92. D. Brancaccio, A. Gallo, M. Mikolajczyk, K. Zovo, P. Palumaa, E. Novellino, M. Piccioli, S. Ciofi–Baffoni, L. Banci, *J. Am. Chem. Soc.* **2014**, *136*, 16240–16250.

93. D. Brancaccio, A. Gallo, M. Piccioli, E. Novellino, S. Ciofi-Baffoni, L. Banci, *J. Am. Chem. Soc.* **2017**, *139*, 719–730.
94. S. Gourdoupis, V. Nasta, V. Calderone, S. Ciofi-Baffoni, L. Banci, *J. Am. Chem. Soc.* **2018**, *140*, 14401–14412.
95. A. Melber, U. Na, A. Vashisht, B. D. Weiler, R. Lill, J. A. Wohlschlegel, D. R. Winge, *Elife* **2016**, *5*, e15991.
96. K. Cai, G. Liu, R. O. Frederick, R. Xiao, G. T. Montelione, J. L. Markley, *Structure* **2016**, *24*, 2080–2091.
97. K. Bych, S. Kerscher, D. J. Netz, A. J. Pierik, K. Zwicker, M. A. Huynen, R. Lill, U. Brandt, J. Balk, *EMBO J.* **2008**, *27*, 1736–1746.
98. A. D. Sheftel, O. Stehling, A. J. Pierik, D. J. Netz, S. Kerscher, H. P. Elsasser, I. Wittig, J. Balk, U. Brandt, R. Lill, *Mol. Cell Biol.* **2009**, *29*, 6059–6073.
99. Y. Takahashi, U. Tokumoto, *J. Biol. Chem.* **2002**, *277*, 28380–28383.
100. J. Balk, M. Pilon, *Trends Plant Sci.* **2011**, *16*, 218–226.
101. J. Perard, S. Ollagnier-de Choudens, *J. Biol. Inorg. Chem.* **2018**, *23*, 597.
102. V. Gupta, M. Sendra, S. G. Naik, H. K. Chahal, B. H. Huynh, F. W. Outten, M. Fontecave, S. Ollagnier-de Choudens, *J. Am. Chem. Soc.* **2009**, *131*, 6149–6153.
103. G. Layer, S. A. Gaddam, C. N. Ayala-Castro, S. Ollagnier-de Choudens, D. Lascoux, M. Fontecave, F. W. Outten, *J. Biol. Chem.* **2007**, *282*, 13342–13350.
104. E. S. Boyd, K. M. Thomas, Y. Dai, J. M. Boyd, F. W. Outten, *Biochemistry* **2014**, *53*, 5834–5847.
105. K. Hirabayashi, E. Yuda, N. Tanaka, S. Katayama, K. Iwasaki, T. Matsumoto, G. Kurisu, F. W. Outten, K. Fukuyama, Y. Takahashi, K. Wada, *J. Biol. Chem.* **2015**, *290*, 29717–29731.
106. B. Roche, L. Aussel, B. Ezraty, P. Mandin, B. Py, F. Barras, *Biochim. Biophys. Acta* **2013**, *1827*, 455–469.
107. F. W. Outten, *Biochim. Biophys. Acta* **2015**, *1853*, 1464–1469.
108. A. Saini, D. T. Mapolelo, H. K. Chahal, M. K. Johnson, F. W. Outten, *Biochemistry* **2010**, *49*, 9402–9412.
109. F. W. Outten, M. J. Wood, F. M. Munoz, G. Storz, *J. Biol. Chem.* **2003**, *278*, 45713–45719.
110. G. P. Riboldi, H. Verli, J. Frazzon, *BMC Biochem.* **2009**, *10*, 3. doi: 10.1186/1471–2091–10–3.
111. A. G. Albrecht, D. J. Netz, M. Miethke, A. J. Pierik, O. Burghaus, F. Peuckert, R. Lill, M. A. Marahiel, *J. Bacteriol.* **2010**, *192*, 1643–1651.
112. H. K. Chahal, Y. Dai, A. Saini, C. Ayala-Castro, F. W. Outten, *Biochemistry* **2009**, *48*, 10644–10653.
113. V. D. Paul, R. Lill, *Biochim. Biophys. Acta* **2015**, *1853*, 1528–1539.
114. J. Li, J. A. Cowan, *Chem. Commun.* **2015**, *51*, 2253–2255.
115. H. Shi, K. Z. Bencze, T. L. Stemmler, C. C. Philpott, *Science* **2008**, *320*, 1207–1210.
116. A. Nandal, J. C. Ruiz, P. Subramanian, S. Ghimire-Rijal, R. A. Sinnamon, T. L. Stemmler, R. K. Bruick, C. C. Philpott, *Cell Metab.* **2011**, *14*, 647–657.
117. A. G. Frey, A. Nandal, J. H. Park, P. M. Smith, T. Yabe, M. S. Ryu, M. C. Ghosh, J. Lee, T. A. Rouault, M. H. Park, C. C. Philpott, *Proc. Natl. Acad. Sci. USA* **2014**, *111*, 8031–8036.
118. C. C. Philpott, M. S. Ryu, A. Frey, S. Patel, *J. Biol. Chem.* **2017**, *292*, 12764–12771.
119. E. J. Camire, J. D. Grossman, G. J. Thole, N. M. Fleischman, D. L. Perlstein, *J. Biol. Chem.* **2015**, *290*, 23793–23802.
120. D. J. Netz, A. J. Pierik, M. Stumpfig, U. Mühlenhoff, R. Lill, *Nature Chem. Biol.* **2007**, *3*, 278–286.
121. D. J. Netz, M. Stumpfig, C. Dore, U. Mühlenhoff, A. J. Pierik, R. Lill, *Nature Chem. Biol.* **2010**, *6*, 758–765.

122. L. Banci, I. Bertini, V. Calderone, S. Ciofi-Baffoni, A. Giachetti, D. Jaiswal, M. Mikolajczyk, M. Piccioli, J. Winkelmann, *Proc. Natl. Acad. Sci. USA* **2013**, *110*, 7136–7141.
123. L. Banci, S. Ciofi-Baffoni, M. Mikolajczyk, J. Winkelmann, E. Bill, M. E. Pandelia, *J. Biol. Inorg. Chem.* **2013**, *18*, 883–893.
124. D. J. Netz, H. M. Genau, B. D. Weiler, E. Bill, A. J. Pierik, R. Lill, *Biochem. J.* **2016**, *473*, 2073–2085.
125. L. Banci, S. Ciofi-Baffoni, K. Gajda, R. Muzzioli, R. Peruzzini, J. Winkelmann, *Nature Chem. Biol.* **2015**, *11*, 772–778.
126. Y. Saito, H. Shibayama, H. Tanaka, A. Tanimura, I. Matsumura, Y. Kanakura, *Biochem. Biophys. Res. Commun.* **2011**, *408*, 329–333.
127. A. G. Frey, D. J. Palenchar, J. D. Wildemann, C. C. Philpott, *J. Biol. Chem.* **2016**, *291*, 22344–22356.
128. H. Li, D. T. Mapolelo, S. Randeniya, M. K. Johnson, C. E. Outten, *Biochemistry* **2012**, *51*, 1687–1696.
129. L. Banci, F. Camponeschi, S. Ciofi-Baffoni, R. Muzzioli, *J. Am. Chem. Soc.* **2015**, *137*, 16133–16134.
130. E. Urzica, A. J. Pierik, U. Mühlenhoff, R. Lill, *Biochemistry* **2009**, *48*, 4946–4958.
131. D. Song, F. S. Lee, *J. Biol. Chem.* **2011**, *286*, 15797–15805.
132. D. C. Odermatt, K. Gari, *Cell Rep.* **2017**, *18*, 1434–1443.
133. O. Stehling, A. A. Vashisht, J. Mascarenhas, Z. O. Jonsson, T. Sharma, D. Netz, A. J. Pierik, J. A. Wohlschlegel, R. Lill, *Science* **2012**, *337*, 195–199.
134. O. Stehling, J. Mascarenhas, A. A. Vashisht, A. D. Sheftel, B. Niggemeyer, R. Rosser, A. J. Pierik, J. A. Wohlschlegel, R. Lill, *Cell Metab.* **2013**, *18*, 187–198.
135. V. Maione, F. Cantini, M. Severi, L. Banci, *Biochim. Biophys. Acta* **2018**, *1862*, 1980–1987.
136. V. D. Paul, U. Mühlenhoff, M. Stumpfig, J. Seebacher, K. G. Kugler, C. Renicke, C. Taxis, A. C. Gavin, A. J. Pierik, R. Lill, *Elife* **2015**, *4*, e08231.
137. C. Zhai, Y. Li, C. Mascarenhas, Q. Lin, K. Li, I. Vyrides, C. Grant, B. Panaretou, *Oncogene* **2014**, *33*, 484–494.
138. M. K. Bruska, M. T. Stiebritz, M. Reiher, *Chemistry* **2015**, *21*, 19081–19089.
139. C. Andreini, L. Banci, I. Bertini, A. Rosato, *J. Proteome Res.* **2006**, *5*, 3173–3178.
140. B. M. Hover, K. Yokoyama, *J. Am. Chem. Soc.* **2015**, *137*, 3352–3359.
141. D. Jardim-Messeder, C. Cabreira-Cagliari, R. Rauber, A. C. Turchetto-Zolet, R. Margis, M. Margis-Pinheiro, *Mitochondrion* **2017**, *34*, 56–66.
142. B. Troxell, H. Xu, X. F. Yang, *J. Biol. Chem.* **2012**, *287*, 19284–19293.
143. V. V. Tryon, J. B. Baseman, *Microb. Pathog.* **1987**, *3*, 437–443.
144. C. Andreini, L. Banci, A. Rosato, *J. Proteome Res.* **2016**, *15*, 1308–1322.
145. C. Andreini, V. Putignano, A. Rosato, L. Banci, *Metallomics* **2018**, *10*, 1223–1231.
146. V. Yankovskaya, R. Horsefield, S. Tornroth, C. Luna-Chavez, H. Miyoshi, C. Leger, B. Byrne, G. Cecchini, S. Iwata, *Science* **2003**, *299*, 700–704.
147. N. Maio, A. Singh, H. Uhrigshardt, N. Saxena, W. H. Tong, T. A. Rouault, *Cell Metab.* **2014**, *19*, 445–457.
148. T. A. Rouault, *Nature Rev. Mol. Cell Biol.* **2015**, *16*, 45–55.
149. C. C. Winterbourn, *Toxicol. Lett.* **1995**, *82–83*, 969–974.
150. J. O. Fuss, C. L. Tsai, J. P. Ishida, J. A. Tainer, *Biochim. Biophys. Acta* **2015**, *1853*, 1253–1271.
151. E. C. M. Tse, T. J. Zwang, S. Bedoya, J. K. Barton, *ACS Central Science* **2019**, *5*, 65–72.
152. E. O'Brien, L. E. Salay, E. A. Epum, K. L. Friedman, W. J. Chazin, J. K. Barton, *Proc. Natl. Acad. Sci. USA* **2018**, *115*, 13186–13191.
153. T. J. Zwang, E. C. M. Tse, J. K. Barton, *ACS Chem. Biol.* **2018**, *13*, 1799–1809.

154. R. F. Service, *Science* **2014**, *346*, 1284–1287.
155. C. Shih, A. K. Museth, M. Abrahamsson, A. M. Blanco-Rodriguez, A. J. Di Bilio, J. Sudhamsu, B. R. Crane, K. L. Ronayne, M. Towrie, A. Vlcek, Jr., J. H. Richards, J. R. Winkler, H. B. Gray, *Science* **2008**, *320*, 1760–1762.
156. A. K. Boal, E. Yavin, O. A. Lukianova, V. L. O'Shea, S. S. David, J. K. Barton, *Biochemistry* **2005**, *44*, 8397–8407.
157. A. K. Boal, J. C. Genereux, P. A. Sontz, J. A. Gralnick, D. K. Newman, J. K. Barton, *Proc. Natl. Acad. Sci. USA* **2009**, *106*, 15237–15242.
158. A. Sheftel, O. Stehling, R. Lill, *Trends Endocrinol. Metab.* **2010**, *21*, 302–314.
159. T. A. Rouault, W. H. Tong, *Trends Genet.* **2008**, *24*, 398–407.
160. N. E. Babady, N. Carelle, R. D. Wells, T. A. Rouault, M. Hirano, D. R. Lynch, M. B. Delatycki, R. B. Wilson, G. Isaya, H. Puccio, *Mol. Genet. Metab.* **2007**, *92*, 23–35.
161. H. Puccio, M. Koenig, *Curr. Opin. Genet. Dev.* **2002**, *12*, 272–277.
162. V. Campuzano, L. Montermini, M. D. Molto, L. Pianese, M. Cossee, F. Cavalcanti, E. Monros, F. Rodius, F. Duclos, A. Monticelli, F. Zara, J. Canizares, H. Koutnikova, S. I. Bidichandani, C. Gellera, A. Brice, P. Trouillas, M. G. De, A. Filla, F. R. De, F. Palau, P. I. Patel, D. S. Di, J. L. Mandel, S. Cocozza, M. Koenig, M. Pandolfo, *Science* **1996**, *271*, 1423–1427.
163. R. A. Vaubel, G. Isaya, *Mol. Cell Neurosci.* **2013**, *55*, 50–61.
164. H. Li, O. Gakh, D. Y. Smith, W. K. Ranatunga, G. Isaya, *J. Biol. Chem.* **2013**, *288*, 4116–4127.
165. M. G. Cotticelli, L. Rasmussen, N. L. Kushner, S. McKellip, M. I. Sosa, A. Manouvakhova, S. Feng, E. L. White, J. A. Maddry, J. Heemskerk, R. J. Oldt, L. F. Surrey, R. Ochs, R. B. Wilson, *J. Biomol. Screen.* **2012**, *17*, 303–313.
166. Y. Munoz, C. M. Carrasco, J. D. Campos, P. Aguirre, M. T. Nunez, *J. Parkinsons Dis.* **2016**, *2016*, Article ID 7049108, doi: 10.1155/2016/7049108.
167. C. Wachnowsky, I. Fidai, J. A. Cowan, *Metallomics* **2018**, *10*, 9–29.
168. A. Shukla, M. Hebbar, A. Srivastava, R. Kadavigere, P. Upadhyai, A. Kanthi, O. Brandau, S. Bielas, K. M. Girisha, *J. Hum. Genet.* **2017**, *62*, 723–727.
169. F. Mochel, M. A. Knight, W. H. Tong, D. Hernandez, K. Ayyad, T. Taivassalo, P. M. Andersen, A. Singleton, T. A. Rouault, K. H. Fischbeck, R. G. Haller, *Am. J. Hum. Genet.* **2008**, *82*, 652–660.
170. R. Spiegel, A. Saada, J. Halvardson, D. Soiferman, A. Shaag, S. Edvardson, Y. Horovitz, M. Khayat, S. A. Shalev, L. Feuk, O. Elpeleg, *Eur. J. Hum. Genet.* **2014**, *22*, 902–906.
171. S. M. Farhan, J. Wang, J. F. Robinson, P. Lahiry, V. M. Siu, C. Prasad, J. B. Kronick, D. A. Ramsay, C. A. Rupar, R. A. Hegele, *Mol. Genet. Genom. Med.* **2014**, *2*, 73–80.
172. S. C. Lim, M. Friemel, J. E. Marum, E. J. Tucker, D. L. Bruno, L. G. Riley, J. Christodoulou, E. P. Kirk, A. Boneh, C. M. DeGennaro, M. Springer, V. K. Mootha, T. A. Rouault, S. Leimkühler, D. R. Thorburn, A. G. Compton, *Hum. Mol. Genet.* **2013**, *22*, 4460–4473.
173. S. H. Kevelam, R. J. Rodenburg, N. I. Wolf, P. Ferreira, R. J. Lunsing, L. G. Nijtmans, A. Mitchell, H. A. Arroyo, D. Rating, A. Vanderver, C. G. van Berkel, T. E. Abbink, P. Heutink, M. S. van der Knaap, *Neurology* **2013**, *80*, 1577–1583.
174. T. Fujiwara, H. Harigae, *Pediatr. Int.* **2013**, 55, 675–9.
175. H. Ye, S. Y. Jeong, M. C. Ghosh, G. Kovtunovych, L. Silvestri, D. Ortillo, N. Uchida, J. Tisdale, C. Camaschella, T. A. Rouault, *J. Clin. Invest.* **2010**, *120*, 1749–1761.
176. P. R. Baker, M. W. Friederich, M. A. Swanson, T. Shaikh, K. Bhattacharya, G. H. Scharer, J. Aicher, G. Creadon–Swindell, E. Geiger, K. N. MacLean, W. T. Lee, C. Deshpande, M. L. Freckmann, L. Y. Shih, M. Wasserstein, M. B. Rasmussen, A. M. Lund, P. Procopis, J. M. Cameron, B. H. Robinson, G. K. Brown, R. M. Brown, A.

G. Compton, C. L. Dieckmann, R. Collard, C. R. Coughlin, E. Spector, M. F. Wempe, J. L. Van Hove, *Brain* **2014**, *137*, 366–379.
177. H. Walden, A. J. Deans, *Annu. Rev. Biophys.* **2014**, *43*, 257–278.
178. A. N. Suhasini, R. M. Brosh, Jr., *Mutat. Res.* **2013**, *752*, 138–152.
179. L. Fan, J. O. Fuss, Q. J. Cheng, A. S. Arvai, M. Hammel, V. A. Roberts, P. K. Cooper, J. A. Tainer, *Cell* **2008**, *133*, 789–800.
180. P. van der Lelij, K. H. Chrzanowska, B. C. Godthelp, M. A. Rooimans, A. B. Oostra, M. Stumm, M. Z. Zdzienicka, H. Joenje, J. P. de Winter, *Am. J. Hum. Genet.* **2010**, *86*, 262–266.
181. J. M. Capo-Chichi, S. K. Bharti, J. A. Sommers, T. Yammine, E. Chouery, L. Patry, G. A. Rouleau, M. E. Samuels, F. F. Hamdan, J. L. Michaud, R. M. Brosh, Jr., A. Megarbane, Z. Kibar, *Hum. Mutat.* **2013**, *34*, 103–107.
182. J. B. Vannier, G. Sarek, S. J. Boulton, *Trends Cell Biol.* **2014**, *24*, 416–425.
183. M. K. Brinkmeyer, S. S. David, *DNA Repair* **2015**, *34*, 39–51.

8

The Cofactors of Nitrogenases

Ivana Djurdjevic, Christian Trncik, Michael Rohde, Jakob Gies, Katharina Grunau, Florian Schneider, Susana L. A. Andrade, and Oliver Einsle

Institute for Biochemistry, Albert-Ludwigs-Universität Freiburg,
D-79104 Freiburg im Breisgau, Germany
<einsle@biochemie.uni-freiburg.de>

Metal Ions in Life Sciences, Volume 20 Guest Editors: Martha E. Sosa Torres and Peter M. H. Kroneck
Series Editors: Astrid Sigel, Eva Freisinger, and Roland K. O. Sigel

https://doi.org/10.1515/9783110589757-008

Abstract: In biological nitrogen fixation, the enzyme nitrogenase mediates the reductive cleavage of the stable triple bond of gaseous N_2 at ambient conditions, driven by the hydrolysis of ATP, to yield bioavailable ammonium (NH_4^+). At the core of nitrogenase is a complex, iron-sulfur based cofactor that in most variants of the enzyme contains an additional, apical heterometal (Mo or V), an organic homocitrate ligand coordinated to this heterometal, and a unique, interstitial carbide. Recent years have witnessed fundamental advances in our understanding of the atomic and electronic structure of the nitrogenase cofactor. Spectroscopic studies have succeeded in trapping and identifying reaction intermediates and several inhibitor- or intermediate-bound structures of the cofactors were characterized by high-resolution X-ray crystallography. Here we summarize the current state of understanding of the cofactors of the nitrogenase enzymes, their interplay in electron transfer and in the six-electron reduction of nitrogen to ammonium and the actual theoretical and experimental conclusion on how this challenging chemistry is achieved.

Keywords: biological nitrogen fixation · FeMo cofactor · FeV cofactor · iron-sulfur cluster · nitrogenase · structural biology

1. INTRODUCTION: BIOLOGICAL NITROGEN FIXATION

Of all the natural elements required for the assembly of biological macromolecules, nitrogen poses a particular problem due to its limited bioavailability in many habitats [1]. As a building block of amino acids and sugars, certain lipids and nucleotides, nitrogen is a bulk element of the living world. In closed ecosystems,

most organisms are able to salvage nitrogen-containing biomolecules from the environment to meet their metabolic requirements. Animals and humans even depend on such nitrogen sources, as they are unable to metabolize basic, inorganic modifications of nitrogen (as observed for carbon, sulfur, and other key elements) [2]. Microorganisms and plants readily assimilate reduced nitrogen in the form of ammonium, NH_4^+, into glutamate or glutamine, and many also catalyze oxidation and reduction reactions that provide access to a series of other modifications of nitrogen in the environment, giving rise to a biogeochemical metabolic network commonly referred to as the nitrogen cycle [3, 4]. In contrast to analogous metabolic cycles that can be drawn for any element relevant to organismic growth, the nitrogen cycle includes one modification that dwarves all others by quantity: elemental nitrogen, N_2, has an extraordinarily high bond enthalpy that conveys such chemical inertness that at any given point in time approximately 99 % of all nitrogen available to the biosphere is present in this form, almost exclusively as N_2 gas in the Earth's atmosphere [4, 5]. Nitrogen gas is arguably the most difficult substrate for any biochemical reaction, and this is highlighted by the fact that in spite of its crucial role for organismic growth, evolution has brought forth only a single class of enzymes, the nitrogenases, that are able to reductively cleave the dinitrogen triple bond, forming the easily accessible ammonium cation [6].

Contrary to the well-known industrial Haber-Bosch process [7], nitrogenases cannot utilize high temperature and pressure for the activation of N_2, but rely instead on the ubiquitous source of free enthalpy in biological processes, released by hydrolysis of the phosphodiester bonds of adenosine-5′-triphosphate (ATP). In Haber-Bosch catalysis, an N_2/H_2 mixture is used as a syngas, with H_2 as a source both of low-potential electrons and protons required for the reduction of N_2. In the reaction of the nitrogenases, protons are readily available from solvent water, while an incredibly precise modification and control of the free enthalpy (or reduction potential) of the electrons is key to achieving efficient N_2 triple bond cleavage under environmental conditions. Moreover, rather than consuming H_2, nitrogenases divert a substantial part of the flow of electrons through their system towards protons, resulting in the formation of H_2 concomitant with N_2 fixation. While this side reaction was known early on, it was long regarded as the inevitable price to be paid for the formation of the low potential required for the reaction in an aqueous milieu. It therefore was no little surprise when in a now classical experiment in 1984, Simpson and Burris showed that even under 50 atm of N_2 gas, the production of H_2 by nitrogenase did not cede, but rather approached a stoichiometric ratio of 1 H_2 produced per N_2 consumed [8]. This translated to the now accepted equation for biological nitrogen fixation (Equation 1),

$$N_2 + 10\,H^+ + 8\,e^- + 16\,MgATP \rightarrow 2\,NH_4^+ + H_2 + 16\,MgADP + 16\,P_i \quad (1)$$

Nitrogenase thus hydrolyzes 2 molecules of ATP for each electron that is transferred, making the process highly costly for nitrogen-fixing – *diazotrophic* – organisms. Interestingly, all known diazotrophs are either bacteria or methanogenic

archaea, with marine cyanobacteria making up for the bulk of nitrogen fixers on the planet [9]. No eukaryotic organism has ever adopted this particular metabolic pathway, which is an intriguing finding in itself, in particular as for instance, leguminous plants enter tight symbioses with diazotrophic bacteria and supply them with up to 25 % of the energy gained through photosynthesis [10].

2. NITROGENASE ENZYMES: STRUCTURE AND REACTIVITY

In the above description of biological nitrogen fixation, the term *nitrogenases* was consistently used in plural, although the monophyletic origin of all such enzymes was emphasized. Indeed, three different types of the enzyme are known to date [11], but all share mechanistic and structural features to a very high degree and are assumed to achieve the catalytic reduction of dinitrogen along the very same lines. The three groups of nitrogenases are metalloenzymes based on complex iron-sulfur clusters that will be the main topic of this contribution, and they are classified by the occurrence of an additional heterometal in their active site cofactor. Here, the most active class of nitrogenases incorporates a molybdenum ion, while under limiting conditions for this transition metal, a second class of enzyme can be produced (from a distinct set of genes) that then employs vanadium. Only if both are not readily available, some organisms contain a third machinery that works with iron only, at the price of a reduced nitrogen-fixing activity. Hence, the three enzyme classes within the family are referred to as Mo-, V-, and Fe-nitrogenases [12].

Nitrogenases are intricate enzyme systems, consisting of two component proteins that achieve a defined and highly regulated distribution of tasks during

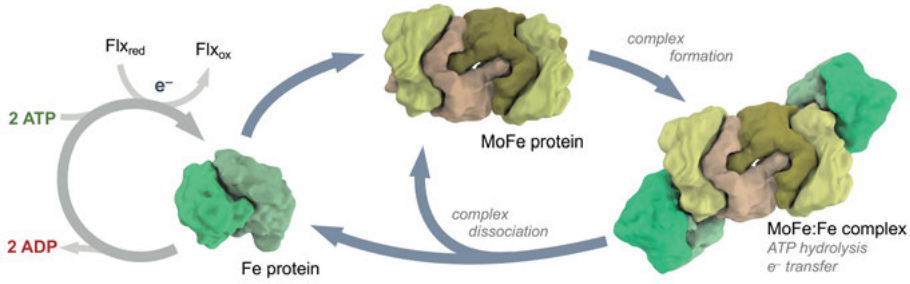


Figure 1. Components of the nitrogenase system and their interactions. The Fe protein is the homodimeric reductase that binds two molecules of ATP and receives an electron to its bridging [4Fe-4S] cluster from flavodoxin (Flx). The reduced Fe protein then forms a complex with the MoFe protein, the catalytic nitrogenase, and in this complex the electron is transferred to the P-cluster and ATP is hydrolyzed to ADP + P_i. This is followed by complex dissociation, priming Fe protein for nucleotide exchange in order to initiate the next electron transfer cycle. For the conversion of a single molecule of N_2, this entire sequence must be executed at least eight times.

their action on the inert substrate [6]. The primary component of the nitrogenase system is the catalytically active dinitrogenase (component I) that holds two large metal clusters for electron transfer and catalysis, respectively, and is named according to the metal ions contained in its active site cofactor: the MoFe protein for Mo-nitrogenase, the VFe protein for V-nitrogenase, and the FeFe protein for the iron-only enzyme. Dinitrogenase is a soluble protein located in the cytoplasm of diazotrophs and receives electrons exclusively from the second component of the system that accordingly is called dinitrogenase reductase, or – as it contains a canonical [4Fe-4S] cluster – the Fe protein (Figure 1). As all three nitrogenase variants are encoded in distinct gene clusters, *nif*, *vnf*, and *anf*, each features its dedicated Fe protein, NifH, VnfH and AnfH. All three are highly similar in structure, but only the highly homologous NifH and VnfH can functionally complement each other *in vitro*.

2.1. The Fe Protein: ATP Hydrolysis and Electron Transfer

The Fe proteins of the three nitrogenase systems are homodimeric single-domain proteins of approximately 60 kDa per dimer, and the aforementioned [4Fe-4S] is coordinated precisely on the internal twofold symmetry axis of the dimer, linked to two cysteine residues in each monomer (Figure 2A). The Fe protein is also the site of ATP binding and hydrolysis [13], and its βαβ-type fold shows clear structural homologies to the family of P-loop NTPases that characteristically switch between two conformational states in dependence on bound NTP versus NDP (Figure 2B) [14]. G-proteins, actin, and myosin are among the most prominent members of this family, and similar to these, the Fe proteins contain a P-loop sequence (residues 10–16 in NifH) that tightly binds the β- and γ-phosphate of ATP and mediates phosphodiester hydrolysis with the help of a coordinated Mg^{2+} ion. The release of inorganic phosphate (P_i) then leads to conformational changes in two loop regions of the Fe protein, switch I (aa 39–44) and switch II (aa 125–132). The terminal residue of the switch II motif, cysteine 132, is a ligand to the [4Fe-4S] cluster, and its rearrangement substantially alters the exposure of the cluster at the protein surface, likely with profound implications for electron transfer to the nitrogenase component (Figure 2C). Interestingly, the activation of the phosphatase activity of the P-loop directly involves residue Lys10 within this sequence, but this lysine cannot interact with the nucleotide in the same monomer of the Fe protein, as it is directed away from the P-loop motif, facing the dimer interface. The Fe protein binds 2 ATP molecules, likely after reduction of its metal cluster, and does so by exchanging bound ADP from a previous round of electron transfer. ATP binding then seems to make the dimer more flexible, allowing for the two protomers to swing around a hinge region that is centered precisely on the metal site. In this state, the Fe protein dimer can probe the surface of the dinitrogenase enzyme and dock on a specific binding site on its surface, but it does not yet attain a conformation that allows for the interaction of Lys10 with the respective other monomer in order to trigger ATP hydrolysis [15, 16]. Once the

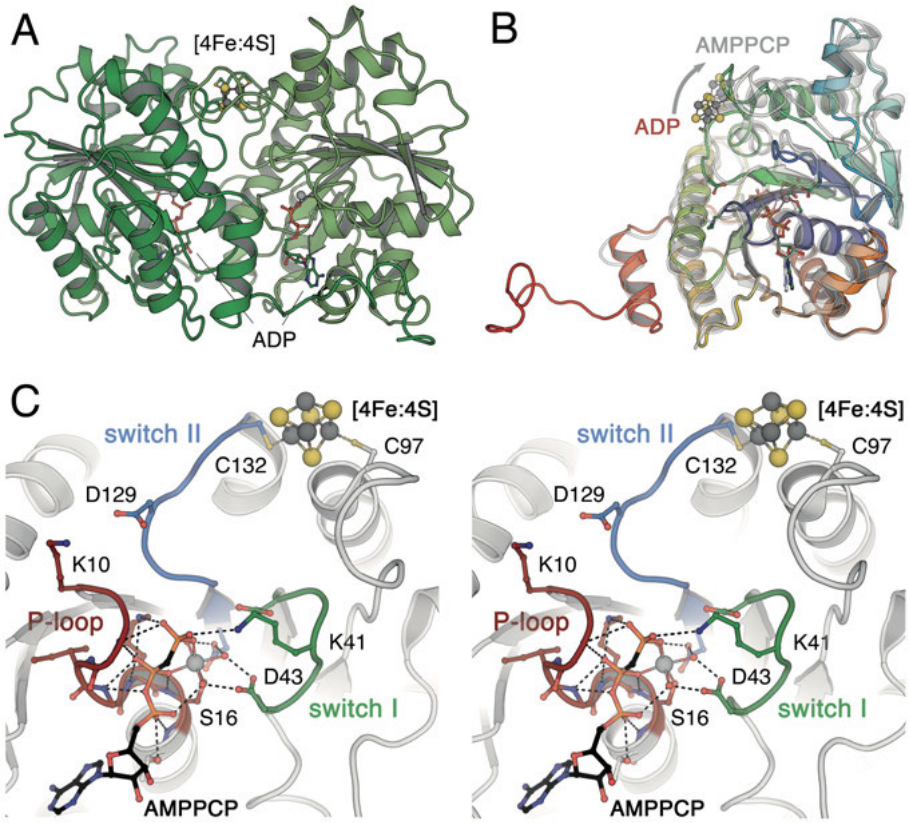


Figure 2. Nitrogenase reductase, the Fe protein. (**A**) Architecture of the homodimeric reductase NifH, the Fe protein of *A. vinelandii* Mo nitrogenase (PDB 1FP6). A cubane-type [4Fe:4S] cluster forms the main connecting element of the monomers and is liganded by two cysteine residues from each. (**B**) Conformational changes in NifH, in a comparison of the ADP-bound state (colored from blue at the N-terminus to red at the C-terminus, PDB 1FP6) and an AMPPCP complex that mimics the ATP-bound form (gray, PDB 1AFK). The rearrangement of the switch I and switch II motifs in the two structures leads to a translocation of the [4Fe-4S] cluster and a hinge motion of the entire dimer. (**C**) Stereo representation of the nucleotide binding site in Fe protein with bound AMPPCP. The nucleotide coordinates a Mg^{2+} cation and is bound in the conserved P-loop and the flexible, adjacent switch I and switch II regions make a connection to the bridging metal cluster.

binding position on dinitrogenase is found, however, the Fe protein is locked in and the two monomers approach each other. The effect of this binding event is twofold: First, the [4Fe-4S] cluster, linked to Cys132 in the switch II region, is pushed towards the interface with the dinitrogenase, shortening the distance for electron transfer to the proximal metal cluster of the catalytic component. Second, Lys10 in this conformation is precisely juxtaposed with the P-loop motif of the other monomer, making the hydrolysis of the nucleotide triphosphate pos-

sible. Recent kinetic studies indicate that electron transfer precedes ATP hydrolysis [17] and the use of non-hydrolysable analogs of ATP (AMPPCP or $ADP{\cdot}AlF_4^-$) made it possible to stabilize and analyze the complex of both component proteins by X-ray crystallography [16, 18].

In nitrogenase, the energy of phosphodiester hydrolysis in the nucleotide must be converted into a form that is productive in promoting the reduction of N_2 at the active site cofactor. It is generally assumed that this is achieved by lowering the effective reduction potential of an electron to allow it to reduce the catalytic site [18], but the mechanisms underlying this process are far from trivial. By changing the conformation of the two switch regions, the effect of ATP hydrolysis is primarily mechanical in nature, inducing a 'power stroke' similar to the twitching motion of a myosin head that leads to muscle contraction. This rearrangement is coupled to (or following) the transfer of an electron from the cubane cluster into nitrogenase, and here the energy of the conformational change of the Fe protein must be passed to the electron, enabling its transfer to the highly reduced cofactor (see Section 3.1). In itself, this successive sequence of energy conversions is highly remarkable, and in order to drive the dissociation of a single N_2 molecule into two ammonium ions (with concomitant H_2 release), this has to happen at least eight times (Figure 1). Not surprisingly, there are ample opportunities for failure along the way.

2.2. Architecture of Dinitrogenases

The Fe protein interacts with dinitrogenase such that in each of the three systems is an $\alpha_2\beta_2$ heterotetramer of approximately 240 kDa. In addition, the vanadium and iron-only proteins contain an additional δ-subunit that does not hold further cofactors and whose role is not entirely clear. Using a distinct set of genes, the three dinitrogenases are $NifD_2K_2$ (Mo-nitrogenase), $VnfD_2K_2G_2$ (V-nitrogenase), and $AnfD_2K_2G_2$ (Fe-nitrogenase). All three systems share a high degree of sequence homology, and moreover, the D and K genes themselves are related to each other, hinting at an early gene duplication event originating from a common precursor protein (Figure 3) [19].

These fundamental subunits then exhibit further internal structure, as each chain consists of three distinct domains that in some instances are connected through additional linker regions. Each domain consists of a basic βαβ-fold of the Rossmann-type, with a four- or five-stranded parallel β-sheet flanked by connecting α-helices [20]. The C-terminal loop regions of these domains interact with the metal clusters in the protein, adhering to the canonical location of cofactor binding sites in Rossmann-fold domains. The peculiar architecture of nitrogenases indicates that the entire complexity of the nitrogen-fixing system can be traced back to a single, globular domain that extended its scope through a triplication within a single gene [19]. While the origin of nitrogenases seemed to be quite obscure for a long time, a clear link has been established to two protein complexes from bacteriochlorophyll biosynthesis, the dark-operative protochlorophyllide reductase (DPOR) and chlorophyllide oxidoreductase (COR) [21].

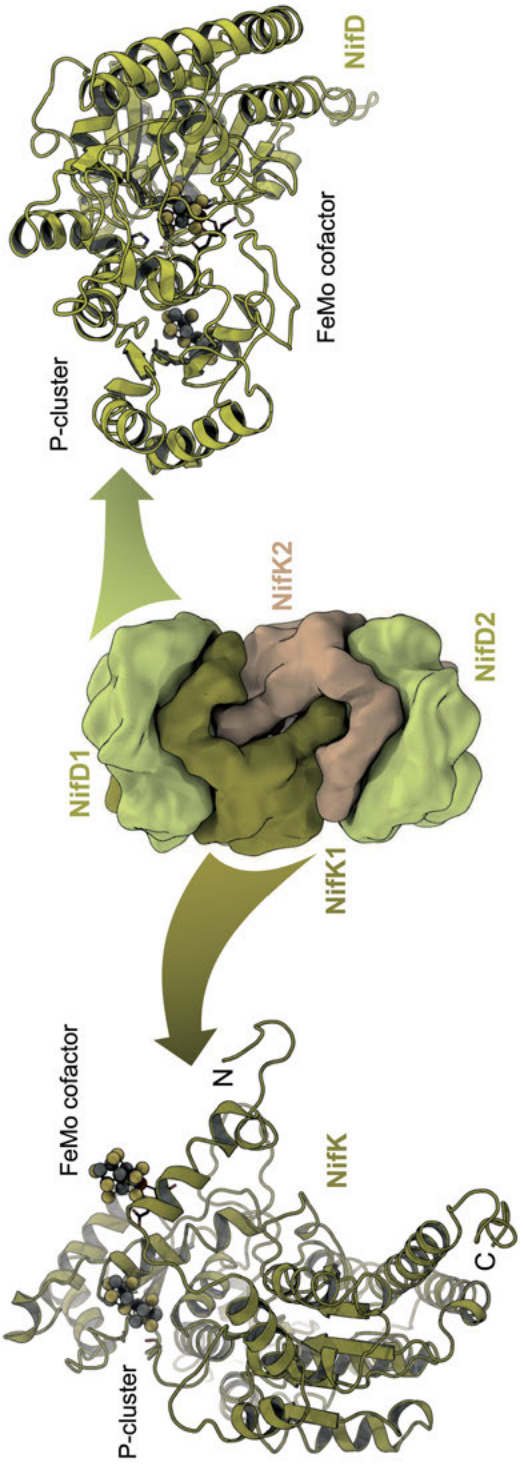


Figure 3. Architecture of the nitrogenase component of the Mo-nitrogenase system. The *A. vinelandii* MoFe protein (center) is a $NifD_2K_2$ heterotetramer with two complements each of the electron-transferring P-cluster and the catalytic FeMo cofactor. In NifD (right), the FeMo cofactor is cradled within the three Rossmann-type domains of the polypeptide chain. NifD and NifK are structurally related, pointing towards an early gene duplication event. P-cluster is located precisely on the pseudo-twofold axis relating NifD and NifK. The NifK chain itself is not in contact with the catalytic cofactor (left).

The recent structural analysis of DPOR has revealed a remarkably similar architecture to the nitrogenases, in the form of a homologous $\alpha_2\beta_2$ heterotetramer that interacts with a dimeric reductase. The cluster complement of DPOR is simpler than in nitrogenases, with a cubane-type [4Fe-4S] cluster both in the reductase and the enzyme, and a large substrate binding site close to the position of the active site cofactor of N_2ases. This structure indicates that nitrogenases may well descend from the bacteriochlorophyll-forming enzymes, which bears the intriguing implication that biological nitrogen fixation may only have evolved *after* the advent of photosynthesis, when life had found a way to conserve such significant amounts of energy that other factors – in this case nitrogen – became limiting [21]. Also, the more recent finding of a set of 'Nif-like' proteins provided further insights, as these consist of a dimeric reductase with clear homologies to the Fe protein, interacting with a second enzymatic component that forms only a homodimer. The role of this system was identified to be in the biosynthesis of cofactor F_{430}, a unique Ni-tetrapyrrole compound found at the active site of methyl-coenzyme M reductase from archaeal methanogens [22, 23]. In the enzyme, now designated CfbDC, the Fe protein $CfbD_2$ has all the canonical features, including a shared [4Fe-4S] cluster at the dimer interface, while the enzymatic component CfbC exhibits the same three-domain architecture as both the D- and K-subunits of the nitrogenases. This arrangement strongly suggests a successive evolution from a system involved in tetrapyrrole modification for (anaerobic) methanogenesis, followed by tetrapyrrole modification for photosynthesis to serving as a scaffold for the insertion of the nitrogenase cofactor, an FeS-based moiety able to catalyze one of the most remarkable chemical reactions in the organismic world.

The concept of a scaffold, distinct from an apoprotein that is fitted with a complete cofactor that was synthesized *ex situ*, is actually the rule rather than the exception for iron-sulfur proteins [24]. The two metal clusters of nitrogenase are among the largest and most intricate metal clusters known to date, and their stepwise assembly and eventual delivery to the protein is a multi-faceted story that has been extensively reviewed elsewhere [25, 26]. At the interface of the D-K dimer, nitrogenases contain a unique [8Fe-7S] moiety designated as the P-cluster that serves as an electron relay for reduction equivalents provided by the Fe protein and the second cluster of the enzymatically active component, the active site cofactor. Nitrogenase cofactors differ for the three known classes of the enzyme and are highly reduced already in their resting state, so that the successive electron transfer required for N_2 cleavage is far from trivial. Interestingly, the P-cluster is typically found in an all-ferrous state – i.e., fully reduced – in nitrogenases as isolated in the presence of a reductant (typically sodium dithionite, $Na_2S_2O_4$). Electron transfer from the Fe protein to the P-cluster therefore is not straightforward, and the current state of knowledge indicates that the docking of reduced and ATP-bound Fe protein first triggers a single-electron transfer from the P-cluster to the cofactor, and only subsequently the electron is transferred from the [4Fe-4S] cluster in the Fe protein. Described as a 'deficit-spending' mechanism, the relevant implication of this concept is that ATP hydrolysis in

the Fe protein acts on the P-cluster rather than lowering the redox potential of its own cubane cluster. Based on kinetic measurements, Seefeldt and coworkers concluded that ATP is only hydrolyzed after electron transfer has occurred, and that the dissociation of the P_i anion is the rate-limiting step in the entire process [17]. N_2 reduction is therefore governed by the Fe protein undergoing a cycle of reduction, nucleotide exchange, complex formation, electron transfer, nucleotide hydrolysis, and complex dissociation for at least eight times per N_2 molecule [27]. This is not a process optimized for high catalytic rates, and indeed, even the most effective Mo-nitrogenase requires more than a second for a single turnover.

2.3. Nitrogenase Catalysis: The Lowe-Thorneley Mechanism

Assuming that the transfer of eight electrons is the minimal mechanistic requirement for any nitrogenase, a kinetic scheme of the enzyme will comprise eight distinct states. Following this exact concept, Roger Thorneley and David Lowe conducted a series of seminal experiments to assess the reactivity of Mo-nitrogenase, resulting in a mechanistic model that is still considered the gold standard in understanding the mechanistic basis of biological nitrogen fixation [28–31]. The eight-electron reaction cycle commonly referred to as the Lowe-Thorneley scheme also revealed unusual properties of this enzymatic catalysis (Figure 4).

Nitrogenases are typically isolated in a reduced state, partly due to the fact that classically sodium dithionite is employed to double as an oxygen scavenger when handling the highly O_2-sensitive components of the system. Nevertheless, the *as isolated* state (E_0) of Mo-nitrogenase required the successive reduction by no less than four electrons to gain the competence to bind and activate dinitrogen in the intermediate state E_4. Note that binding of N_2 is already possible in E_3, but initiating its conversion and effectively breaking the N_2 triple bond requires further reduction to E_4 also here. Each of the four electrons needed to promote the enzyme to state E_4 of course comes through the Fe protein, driven by the hydrolysis of ATP and consequently always at the same potential. At the same time, the progressing accumulation of electrons at the cofactor should make every following reduction event more difficult, up to the point that the reducing power of the Fe protein may no longer be sufficient. Alternatively, significant energy might be lost in the first reduction steps due to an excessive reducing power bestowed upon the Fe protein. A second crucial point addressed in the Lowe-Thornely scheme is that the binding of N_2 to E_4 (or E_3) invariably leads to the release of H_2, and it is this H_2 that is generally considered to be the stoichiometrically required one [8]. Assuming that H_2 formation consumes two electrons in E_4, the enzyme then has two further electrons remaining for the reduction of N_2, progressing right through the most challenging part of the reaction (Figure 4). This is well in line with the finding that the second half of the reaction cycle, involving the final formation of two molecules of NH_4^+, proceeds swiftly and without a further substantial barrier.

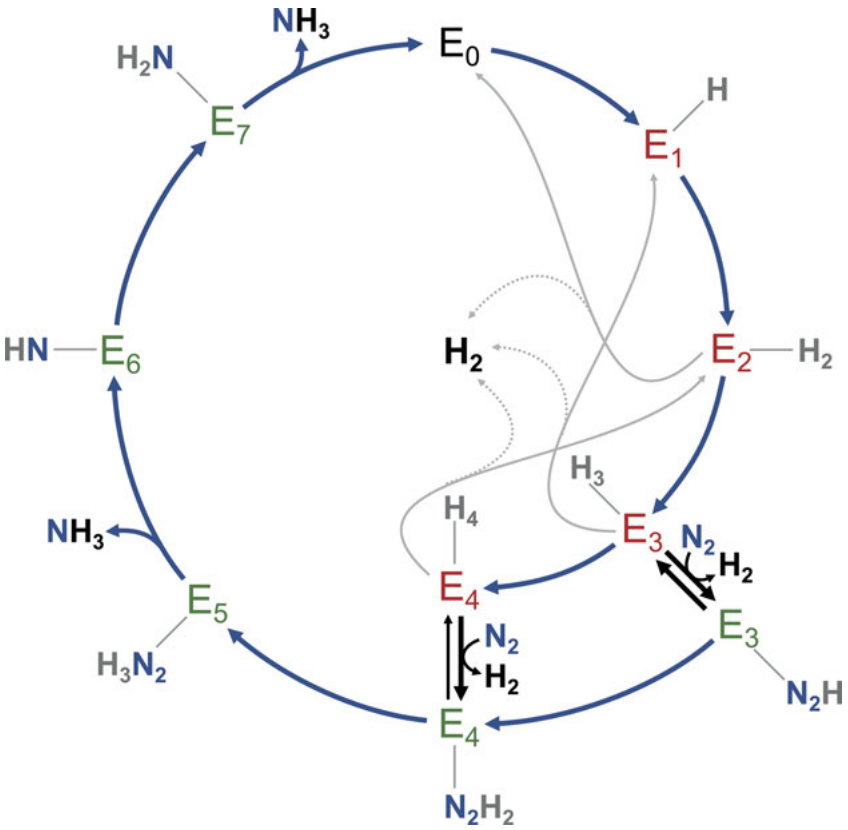


Figure 4. The kinetic scheme for N_2 reduction by Lowe and Thorneley [28–31]. In addition to the 6-electron reduction of N_2 to two molecules of NH_4^+, at least one equivalent of H_2 is released, resulting in an 8-electron process that takes the enzyme from E_0 (resting state) to E_7, where the second product molecule is released. Arrows represent single-electron transfers that are accompanied by protonation events for charge compensation. While in E_1, protonation likely occurs on a bridging sulfide, the transition to state E_2 is already suggested to involve the formation of a first bridging hydride adduct. Even a reversible association of N_2 does not occur before reaching E_3, and only upon N_2 binding to the E_4 state the reaction is committed to proceed to completion. Here, the enzyme has formed two bridging hydride adducts at the cluster surface. These hydrides combine to reductively eliminate H_2, leaving the cluster in a super-reduced state that is able to bind N_2 and transfer two electrons, breaking the triple bond and making the remaining steps of the catalytic cycle energetically straightforward. Suggestion: mention the various intermediates also in the legend, not only in the text, in view of the complexicity of the mechanism. In other words, N_2H, N_2H_2, etc ... in the different states of the enzyme.

2.4. Nitrogenase Is a Hydrogenase: Good and Bad Hydrogen

As detailed above, the stoichiometry laid out earlier and described in detail in the Lowe-Thorneley scheme only represents an idealized outcome of nitrogenase catalysis. Electrons delivered by the Fe protein generate a steady (if slow) flux of electrons towards the active site cofactor, but not all – or even only a minor fraction – of these electrons are transferred to the substrate N_2. A significant portion of the electron flux through all nitrogenases, V- and Fe-enzymes far more than the Mo variant, instead is diverted to the formation of H_2, meaning that the electrons tend to recombine with protons that are abundant in an aqueous milieu and that they are thus lost for the reaction. This unwanted release of H_2 can occur at any point once the enzyme has accumulated at least two electrons and importantly, before the substrate N_2 is bound in state E_4. It puts the enzyme back in the Lowe-Thorneley cycle by two steps and can thus be interpreted as a failure in the attempt to reach the state competent in N_2 binding. Unfortunately, there is presently no data available to address the question from which of the states E_1 through E_4 most of this adventitious hydrogen gas is produced, but the accumulating charge on the active site cofactor would intuitively indicate that the power to reduce protons increases with the progression of the enzyme along the Lowe-Thorneley cycle. In an enzyme system that hydrolyzes two molecules of ATP to allow a single electron to even reach the active site cofactor, this appears to be wasteful in a way that is entirely uncharacteristic for biochemical reactions that have been optimized through billions of years and millions of generations of organismic evolution. This side reaction, however, does not make nitrogenase necessarily another hydrogenase enzyme. Even under high overpressure and with chemical electron acceptors, nitrogenase is not reduced by H_2 gas, so that the crucial criterion of reversibility of the H_2 formation reaction is not fulfilled. There is, however, a mechanistic twist to the interaction of nitrogenases and H_2 that is presumably of the utmost mechanistic importance. Under turnover conditions, i.e., in the presence of reductant, ATP and a suitable substrate such as N_2 or acetylene (C_2H_2), Mo-nitrogenases were found to form HD in the presence of D_2, implying that D_2 gas is activated and converted [32, 33]. While the basis for this surprising finding remained enigmatic for a long time, the same was found for both V- and Fe-nitrogenases [34], further supporting the assumption that the underlying mechanism of reduction is largely identical in all three variants.

2.5. A Mechanism for Nitrogenase

While the kinetic scheme by Lowe and Thorneley provided an outline for N_2 conversion by nitrogenases and many of the implied rate constants for the interconversion of reaction intermediates were experimentally determined, the formulation of a *bona fide* mechanism of nitrogenase action required at least one more crucial component, the molecular structure of these intermediates. Defin-

ing the site and mode of binding to the active site represents an indispensable confirmation for the postulated sequence of events during catalysis, but would also substantially refine the picture obtained from kinetic and spectroscopic measurements. The most obvious path towards defining the intermediates of the nitrogenase reaction would be to actually determine their structure directly and unambiguously through X-ray crystallography. Since the first structure of the Mo-nitrogenase from *Azotobacter vinelandii* presented by Rees and coworkers in 1992 [14, 35, 36], three-dimensional structural data has been crucial for defining and understanding the entirely unprecedented properties of this enzyme. However, in spite of a multitude of structural analyses carried out on a series of different orthologs of the enzyme from various organisms [37–39], it took more than two decades to crystallographically describe a state of nitrogenase that was not the chemically inert resting state, E_0, in the Lowe-Thorneley scheme.

3. METAL CLUSTERS OF NITROGENASES

In spite of all the technical difficulties, biochemistry, X-ray crystallography and spectroscopy ideally combined to flesh out our current picture of nitrogenase function. The most relevant and at the same time most unusual aspects of this class of enzymes are its metal sites. The three clusters in Fe protein and MFe protein (M = Mo, V, Fe) are variations on the theme of iron-sulfur sites in proteins, but all have unique features that set them well apart from the canonical clusters found as redox centers or catalytic sites in a vast number of biological machines.

3.1. The [4Fe-4S] Cluster of the Fe Protein

Among the three types of iron-sulfur centers found in the nitrogenase system, the cluster of the Fe protein comes closest to a canonical cubane-type metal site known from many different members of the ferredoxin family and scores of iron-sulfur-based enzymes in nature. It features the typical, distorted cubane structure and a symmetric coordination by four cysteine residues (Figure 5), and its function is in single-electron transfer, receiving its redox equivalent from a soluble flavodoxin or ferredoxin in the cytoplasm and passing it on to the P-cluster (see Section 3.2). The redox transition relevant to this transfer is the one commonly found in low-potential [4Fe-4S] ferredoxins, from the reduced $S = 1/2$ state (total charge $+1$) to an oxidized $S = 0$ state with a $+2$ total charge [40].

The first of the unusual features of this cluster, however, is its location within the Fe protein. With only two cysteine ligands (C97 and C132 in *A. vinelandii* NifH) supplied by each monomer, the cluster is bound precisely on the twofold symmetry axis of the Fe protein dimer, where it forms the hinge in the flexible

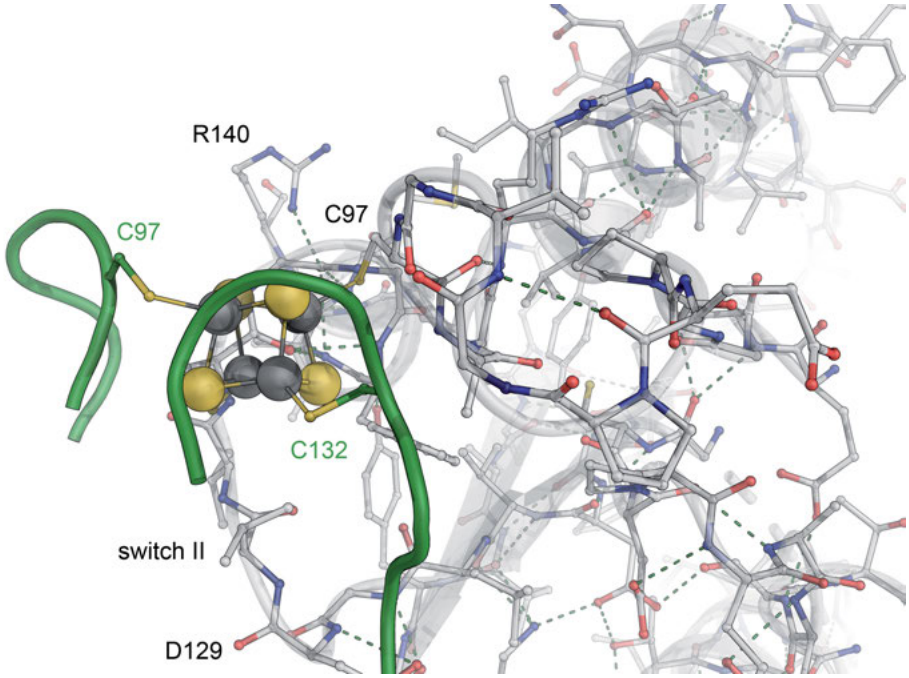


Figure 5. Cluster environment in Fe protein. While the metal center of NifH is a canonical, cubane-type [4Fe-4S] cluster, is location at the interface of the Fe protein homodimer is peculiar. Each protein chain only provides two ligands to the cluster, which in addition functions as the main connecting element in the dimer and also as a hinge for the flexible relative movements of the two protomers that are controlled by the binding and hydrolysis of ATP.

quaternary structure of the reductase (see Section 2.1) [14]. As the Fe protein dimer changes its conformation dependent on the redox state of the cluster, the bound nucleotide and the formation of a complex with the dinitrogenase component, the surface exposure of the cluster, and also the potential to closely approach the P-cluster in the complex changes significantly. In conjunction with the highly reduced state of the P-cluster (see Section 3.2.2) this is presumably of the utmost mechanistic relevance, and it quite definitely is part of the finely tuned interplay of the two protein components that go through a series of precisely timed steps in order to convert the energy of ATP hydrolysis into the reducing power of an electron and enable its transfer to the catalytic cofactor.

The Fe protein has a second peculiarity linked to its metal cluster. Contrary to almost all other known [4Fe-4S] clusters, Fe proteins can be further reduced to an all Fe(II) ([4Fe-4S]^{0} state using chemical reductants such as Ti(III) citrate [41]. While the reductant may be non-physiological, super-reduced NifH was indeed shown to bind and hydrolyze ATP and transfer two electrons to the MoFe protein, leaving the Fe protein cluster in the oxidized [4Fe-4S]$^{2+}$ state [42]. A

three-dimensional structure of this state of NifH did not show any significant structural rearrangements [43], but the technical possibility of transferring two electrons instead of one in a single cycle of ATP hydrolysis potentially has profound implications for the energetics and stoichiometry of the nitrogenase reaction.

3.2. The P-Cluster as an Electron Relay

When reduced the Fe protein with bound ATP forms a complex with the nitrogenase, and the binding position of the reductase component was found to be precisely on the pseudo-twofold symmetry axis relating the D- and K-subunits of the enzyme [18]. Here, the Fe protein cubane cluster is able to approach its immediate electron acceptor, the first of the two large iron-sulfur clusters of the dinitrogenase. This is P-cluster, a [8Fe-7S] moiety that is fully conserved through all three nitrogenase variants and that in its reduced state shows the symmetric architecture of a double-cubane cluster fused through one of its acid-labile sulfide ions, S1. Three cysteines, each from the D- and K-subunits, coordinate the metals, and the cluster is centered with its S1 sulfide on the pseudo-twofold DK-axis [44]. In nitrogenase as isolated, the P-cluster is observed in the diamagnetic P^N state, with all eight Fe ions in the Fe(II) state, and pairwise antiferromagnetic coupling leads to a total spin $S = 0$ (Figure 6A). While this state of the cluster can be oxidized by one (P^{1+}), two (P^{2+}), or even three electrons (P^{3+}), all attempts at further reducing the P^N state have not been successful. P^N likely represents the resting state of P-cluster, and consequently the transfer of an electron from the Fe protein to P-cluster is not unproblematic. The original idea how this problem is overcome in the enzyme was that the hydrolysis of ATP lowered the midpoint potential of its [4Fe-4S] such that electron transfer to the P^N cluster becomes feasible. Structural data on the Fe protein in different nucleotide-bound states was conceptually in line with this, showing that two extended α-helices changed their relative orientation towards the cluster, realigning their helix dipoles with respect to the cluster to potentially change its midpoint potential. It remained unclear, however, whether such a shift in potential was sufficient to inject an additional electron into an all-Fe(II) Fe-S cluster.

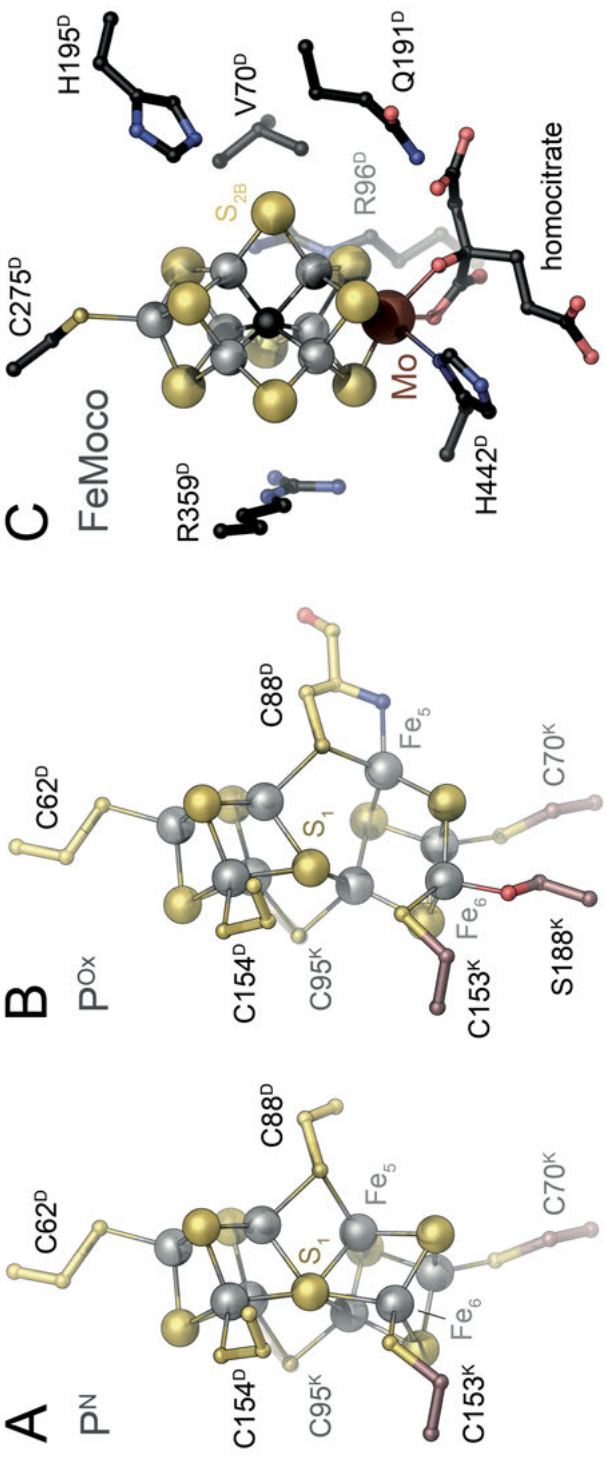


Figure 6. Metal cofactors of the MoFe protein. (**A**) As isolated in the resting state E_0 and in the presence of dithionite, the [8Fe-7S] P-cluster is obtained in an all-Fe(II) state designated P^N. The two fused cubane substructures are highly symmetric and the symmetry operation relating the two halves coincides almost with the pseudosymmetry axis relating NifD to NifK. (**B**) Two-electron oxidation by chemical oxidants or limited exposure to O_2 leads to the P^{Ox} state, in which the cluster loses its symmetry around the central sulfide S1. Instead, Fe5 and Fe6 undergo a conformational change releasing their coordination to the sulfide and moving outwards towards harder ligands. For Fe6 this is the conserved serine 188 in NifK, while Fe5 in P^{Ox} is coordinated by the backbone amide nitrogen of C88 of NifD. (**C**) The catalytic FeMo cofactor is a unique [Mo-7Fe-9S-C]-homocitrate moiety liganded by the protein exclusively through the apical cysteine 275 and histidine 442. The residues of the second coordination sphere of the cluster are highly conserved among different species and include two positively charged arginines 96 and 359, the catalytically relevant glutamine 191 and histidine 195 and the bulky valine 70, whose replacement affects the substrate spectrum of the enzyme as well as the stability of reaction intermediates.

3.2.1. Redox-Dependent Conformational Changes

With the initial structural analyses of the MoFe protein, some ambiguity remained concerning the structure of the P-cluster [36]. The reason for the unclear electron density maps around the metal site were only understood when Rees and coworkers presented an improved structural model that revealed unusual conformational changes in the cluster, leading to different structures in the reduced (P^N) and oxidized (P^{2+}) states [44]. The symmetric dicubane of the P^N state was confirmed in this analysis (Figure 6A), but the structure of an oxidized form, identified spectroscopically as P^{2+}, showed a movement of two of the eight iron ions, concomitant with an exchange of protein ligands. Fe5 and Fe6 were both found to break their coordination to the central sulfide S1, with Fe5 moving towards the backbone amide of residue cysteine 88 of NifD, notably requiring the unfavorable deprotonation of the amide nitrogen as a prerequisite for metal coordination (Figure 6B). Fe6 also moved away from S1, selecting the largely conserved serine 188 of NifK as a preferred ligand in the oxidized state. Chemically, both transitions involve the exchange of a soft ligand (sulfide) for a hard ligand (N or O), in parallel with a two-electron oxidation. As Fe^{3+} also is harder in character than Fe^{2+}, this is generally taken to indicate that in the P^{2+} state of the P-cluster Fe5 and Fe6 are localized Fe(III) sites. The observed conformational change is fully reversible upon reduction of the cluster, further supporting the hypothesis that the phenomenon is indeed of functional relevance for the catalytic cycle of nitrogenases. A caveat remains, however, going back to the ATP/e^- stoichiometry of the nitrogenase reaction. Although the Fe protein, as well as its physiological electron donor, a flavodoxin, are technically capable of two-electron transfer (see Section 3.4.1), it is generally assumed that only one electron is transferred in each cycle, and consequently the same is true for the P-cluster. Here, the relevant redox states are P^N and P^{1+}. From the available structural data on the P-cluster, one might infer that the one-electron oxidized P^{1+} state will only have one iron ion change its place (and ligand), and considering both the complication of deprotonating the amide nitrogen of cysteine 88 (NifD) and the complete conservation of serine 188 (NifK) in all known nitrogenase sequences, this should likely be Fe6 rather than Fe5. Experimental evidence for this hypothesis was presented only most recently, when Peters et al. carried out a redox titration of intact protein crystals to define a series of structures at defined posed potentials [45]. The structure of a one-electron oxidized P-cluster obtained through this strategy indeed confirmed the hypothesis, showing Fe6 to be liganded by serine 188 (NifK).

More recently, Tezcan et al. reported on the structure of the nitrogenase MoFe protein from *Gluconacetobacter diazotrophicus* [38]. This ortholog belongs to a family of nitrogenase proteins that lack the conserved serine 188. The determination of the structures of both the P^N and the P^{2+} state of this protein showed that an analogous conformational change takes place in the cluster, but that a different residue, tyrosine 98 (NifK), now serves as an external ligand to Fe6. As might be expected, this tyrosine was found to be conserved in all those sequences lacking the serine that otherwise plays this role.

3.2.2. Electron Transfer through the P-Cluster

The P-cluster being in an all-Fe(II) state in the resting state enzyme raises the question how electron transfer from the Fe protein can lead to further reduction of this site, even considering the hydrolysis of two molecules of ATP as a driving force. According to a recent kinetic study, the answer to this question seems to lie in the sequence of electron transfer events in the enzyme system and their coordination with ATP hydrolysis. Seefeldt and coworkers postulated a mechanism, in which the complex formation of MoFe protein with reduced Fe protein first triggers electron transfer from the P-cluster to the active site cofactor, and only subsequently the resulting hole in the P-cluster is replenished from the [4Fe-4S] site in the Fe protein [46]. The significance of this is that the energetically driven electron transfer step in the entire process originates at the P-cluster, with the implication that ATP hydrolysis in the Fe protein is not used to lower the midpoint redox potential of the [4Fe-4S] cluster in the reductase component. Later studies by the same group indicated that ATP hydrolysis only occurs after both electron transfer events [47], and finally that the rate-limiting step in the entire electron transfer cycle between the Fe protein and the MoFe protein is the release of P_i after ATP hydrolysis [17]. Taken together, the emerging model of electron transfer is such that the Fe protein must obtain an electron (from flavodoxin or ferredoxin) and exchange bound ADP for ATP before it gains the capacity for complex formation with the dinitrogenase. Apparently, the complex-forming event itself is what triggers the ejection of an electron from the P-cluster into the catalytic cofactor, followed rapidly by reduction of the P-cluster from the [4Fe-4S] cluster of the Fe protein. Either of these events then provides the signal for ATP hydrolysis in the Fe protein, likely causing a conformational change that is prerequisite for the dissociation of the electron transfer complex and serves to make the process irreversible.

To date, several structures of complexes of the two components of nitrogenase have been described. The first structure was obtained by stabilization of the interacting state of the two proteins with $ADP{\cdot}AlF_4^-$ by Rees and coworkers [18]. This compound is generally seen as a stable analog for the transition state of ATP hydrolysis, and following the above logic regarding the sequence of events during the interaction of the two nitrogenase components, this structure supposedly represents a state where electron transfer has already taken place. Subsequently, the same group reported several further complex structures obtained either in the absence of a nucleotide, with bound ADP or with adenosine 5′-[β,γ-methylene]triphosphate (AMPPCP), a non-hydrolysable analog of ATP [16]. Interestingly, the AMPPCP complex showed an identical arrangement to the one with $ADP{\cdot}AlF_4^-$, while the interaction modes of the two proteins in the other complexes were different. The authors interpreted this as a probing of the surface of the nitrogenase by the Fe protein, and later also reported a structure obtained with a mixture of ADP and AMPPCP, where they find a single Fe protein dimer to bind one of these nucleotides in one subunit, and a second in the other [48]. Thus, at present, it is difficult to fully align the existing structural and biochemical mechanism into a consistent model of electron transfer in the nitrogenase system.

3.3. The Catalytic Cofactor

The final objective of the complex electron transfer pathways within the nitrogenase system is to sequentially transfer high-energy electrons to the active site cofactor. This unique moiety is the site of substrate reduction in the enzyme system. It is a complex organometallic cluster that is synthesized entirely *ex situ* and inserted into apo-nitrogenase only in the final step of the biogenesis of the enzyme. Its intricacy and the range of unusual features it presents have long prevented an in-depth understanding of its properties and functionality. Only recently, clear evidence begins to accumulate and although the site remains to reveal its final secrets, a consistent picture is beginning to emerge.

3.3.1. The Unique Cofactor of Nitrogenases

As is the case for all other metal sites in the nitrogenase systems, the active site cofactor is an iron-sulfur based moiety, but differs fundamentally from all other known clusters of this type. Beside seven iron ions, the cofactor of Mo-nitrogenase contains an additional heterometal (obviously Mo), and is thus commonly referred to as 'FeMo cofactor' (Figure 6C). It also incorporates an organic ligand, homocitrate, to molybdenum that is formed from 2-oxoglutarate and acetyl-CoA by NifV, a designated homocitrate synthase [49, 50]. In the absence of any precedent, the nature of this cofactor was a matter of speculation until and even after its three-dimensional structure became known, and the complex history of understanding its properties deserves a detailed review.

3.3.2. Three Classes of Cofactors

The FeMo cofactor remains the best-studied of all nitrogenase cofactors, and being the main point of differentiation between the three classes of nitrogenases, the related cofactors of the V- and Fe-nitrogenases are structurally and functionally distinct. Naming conventions designate the two clusters as 'FeV cofactor' and 'FeFe cofactor', which already implies that a major difference between the clusters lies in the employed heterometal. Indeed, in organisms such as the free-living soil diazotroph *A. vinelandii* that contains all three alternative systems, Mo-nitrogenase is preferred, leading to the production of the V-dependent system only under conditions of Mo depletion. Similarly, the least active Fe-nitrogenase is only produced if neither of the two required heterometals is available [12]. For most of the history of nitrogenase research, the bulk of all investigations has focused on the Mo system, and only most recently structural information for a V-nitrogenase was reported. The following discussion of the structure and properties of the nitrogenase active site cofactor thus largely refers to the history of FeMo cofactor research, and will be followed by newer findings on the FeV cofactor and its differences to the molybdenum system.

4. ATOMIC AND ELECTRONIC STRUCTURE OF THE FeMo COFACTOR

4.1. Understanding the Cofactor Structure

With the availability of isolated protein [51] and an increasing amount of data from a variety of spectroscopic techniques, the unique nature of the metal clusters of nitrogenase became clear early on. Numerous attempts were made to generate synthetic models that would reproduce the observed properties of the cofactors of the enzyme, and the inclusion of a molybdenum site within (or directly coupled to) an iron-sulfur-based scaffold was generally assumed [52–54]. The analysis of P-cluster indeed came close, with an assignment as a pair of [4Fe-4S] clusters following p-xylylenedithiol extraction [55]. However, this procedure failed for the cofactor, already providing a hint at its unusual features.

4.1.1. The First Structural Model for the FeMo Cofactor

All speculations on the structure and nature of the FeMo cofactor were ended – for the time being – with the first successful crystallographic analysis of the MoFe protein from *A. vinelandii* by Kim and Rees in 1992 [35, 36]. The first structural model at a resolution of 2.7 Å was described as consisting of two incomplete cubane-type subclusters, a [4Fe-3S] and a [Mo-3Fe-3S] unit, μ^2-bridged by three non-protein ligands, two of which were already assigned as sulfides. The structure also revealed that the entire site was only fixed in the protein through coordination to two amino acid side chains, Cys 275 to Fe1 and His 446 to Mo. It also showed the position of the previously identified homocitrate molecule and its coordination to molybdenum [36]. The most striking feature of this novel metal site, however, was a large internal cavity with a diameter of approximately 4 Å, surrounded by the six central, coordinatively undersaturated Fe ions. This unusual site for once appeared to be a prime candidate for a possible, if somewhat tight, ligand binding site [56], but also posed problems for explaining the structural rigidity of the cluster in the first place, in particular due to the direct juxtaposition of six metal cations. Computational approaches resorted to the inclusion of metal-metal bonding in order to reproduce the structural features of the site, but at the same time the assembly did not show any reactivity towards substrates without prior activation. Nitrogenase reduced an entire series of small molecules, and as Thorneley and Lowe had shown, the binding of N_2 required the reduction of the enzyme by four electrons. Even the substrate acetylene (HC≡CH) that only undergoes a two-electron reduction to ethylene ($H_2C{=}CH_2$), required activation, and reduction of the MoFe protein by at least two electrons [57], and so for the time being, the mode of substrate binding to cofactor and mechanism of reduction remained largely unknown.

4.1.2. A Central Ligand in the FeMo Cofactor

Structural analyses of nitrogenases by the Rees group meanwhile had provided models of increasing resolution, starting from the initial analysis at 2.7 Å [35, 36], via further refinement at 2.2 Å [56] to a model at 2.0 Å [44]. These improvements provided a more precise structure of P-cluster, revising its original assignment of a [8Fe-8S] moiety to [8Fe-7S] and highlighting the structural changes involved in its redox transitions (see Section 3.2.1), but did not provide new insights regarding the FeMo cofactor. A higher-resolution structure was subsequently obtained for the orthologue from *Klebsiella pneumoniae*, consistently showing a largely identical structure of the FeMo cofactor at a resolution of 1.6 Å [37]. In the $2F_o–F_c$ difference electron density maps of these analyses, the empty central cavity of the cluster was very well defined, leaving little room for any alternative interpretation.

Therefore, it came as a surprise when in 2002 a change of crystallization protocols led to crystals of *A. vinelandii* MoFe protein that diffracted to a far higher resolution of 1.16 Å. This analysis crossed the borderline to true atomic resolution, where every individual atom is indeed represented by a distinct maximum in the electron density map, and provided an unprecedented level of detail for the entire enzyme [58]. It allowed to resolve and refine a mixture of the P^N and P^{2+} states in the P-cluster, but most prominently it revealed an unambiguous and well-defined electron density maximum in the central cavity of the cluster. At the given resolution, this new ligand could be identified as a μ^6-coordinated light atom, representing a reduced species of either C, N or O. This discovery of course gave rise to the immediate question whether such a central light atom was only present in this new structure, or whether the feature had simply been overlooked in earlier analyses. Considering the multitude and consistence of earlier structures, as well as the high resolutions of 2.0 Å (*A. vinelandii*) and 1.6 Å (*K. pneumoniae*) that had been obtained, the latter seemed highly unlikely, but the protein that eventually yielded the atomic-resolution structure had not been treated in a different way than in earlier studies. The solution was eventually found in the unique structural properties of the metal cluster itself. The very center of the cluster is not only surrounded by the six prismatically arranged iron ions Fe2-Fe7, but it is also equidistant to all nine sulfide ions in the arrangement (Figure 7A). The reconstruction of the spatial electron density distributions from diffraction data that underlies crystallographic structure determination is based on the interconversion of both these representations of information *via* a Fourier transform. Hereby, the diffraction data provides the coefficients (structure factors) for the Fourier synthesis of an electron density map. The perfect representation of an electron density distribution in this way technically assumes an infinitely high resolution, which will not be the case in practice. As in any Fourier analysis, an incomplete set of Fourier coefficients then leads series termination artifacts that manifest as periodic noise (‘ripples’) in an electron density map. Such ripples typically are small and tend to compensate each other in a complex arrangement of multiple atoms, so that they rarely become an issue that needs

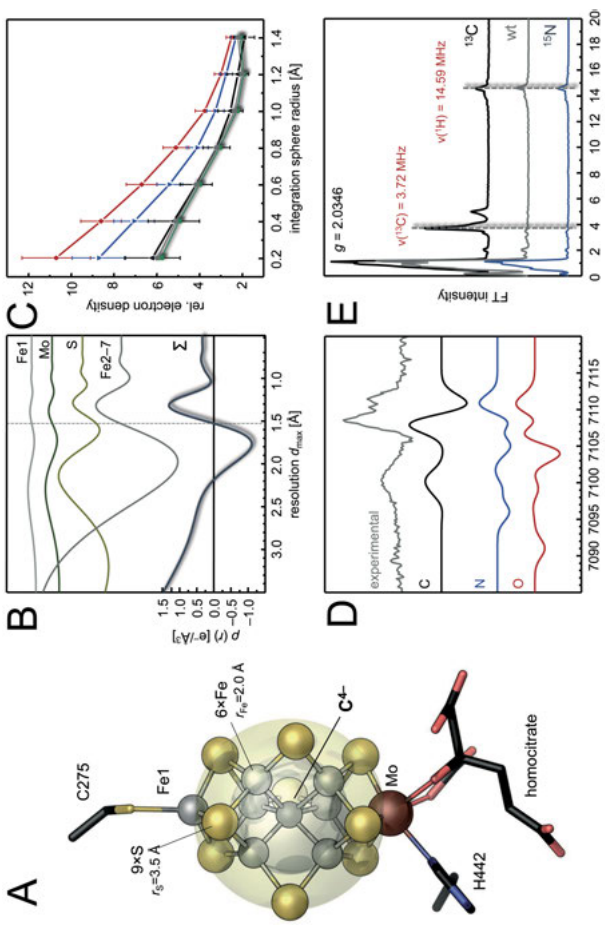


Figure 7. Determination of the structure of the FeMo cofactor. (**A**) FeMo cofactor in the *A. vinelandii* MoFe protein (PDB 3U7Q); in the [Mo-7Fe-9S-C]-homocitrate moiety, the central carbide (C^{4-}) is equidistant from the six central irons (Fe2–Fe7, gray spheres), and also from all nine sulfides (yellow spheres). This unique geometry forms the basis for the amplification of Fourier series termination artifacts. (**B**) Resolution-dependent profile of the artifactual influence of the surrounding atoms at the position of the central carbide. While the influence of the apical F1 and Mo are negligible, the added effects of the multiple sulfides and irons on the two shells add up to a significant artifact, sufficient to obfuscate the central atom at resolutions lower than 1.55 Å. (**C**) The high resolution of 1.0 Å of the obtained crystal structure allowed for a statistical analysis of the diffraction behavior of all carbon (black), nitrogen (blue) and oxygen atoms (red), revealing that the properties of the central atom are best explained by carbon. (**D**) The HERFD-XAS analysis of Mo-nitrogenase revealed features that were well reproduced in a simulation containing a central C (black), but not with either N (blue) or O (red). (**E**) X-band three-pulse ESEEM of wild type (wt) and ^{13}C and ^{15}N-labelled MoFe protein showed a ^{13}C resonance indicative of the nature of the interstitial atom.

to be taken into account. In the FeMo cofactor, however, the unusual degree of symmetry of the cluster structure with respect to its central position makes this effect highly significant. While at limited resolutions the effect of a single iron ion or a single sulfide might be negligible, the central cavity of the cofactor is where this effect adds up for six iron and nine sulfur ions (plus the apical metals) to generate a defined profile of electron density that entirely represents an artifact attributable to the geometry of the cofactor itself (Figure 7B). Most peculiarly, the result of this additive ripple effect in cofactor is a *negative* electron density feature right at the cluster's center. This feature occurs in a resolution range between 2.2 Å and 1.55 Å, and is sufficiently pronounced to mask the presence of a light atom at the central position (Figure 7B). The limit of 1.55 Å implies that unfortunately this atom was not yet visible in the 1.6 Å *K. pneumoniae* structure, but appeared very prominently and unambiguously at 1.16 Å in *A. vinelandii* MoFe protein [58]. It quickly became evident and has been confirmed by multiple structural studies since, that the central atom is an integral structural feature of the nitrogenase cofactors. Its nature remained a matter of debate at the point of its discovery, as did (and does) its role during catalysis [59].

4.1.3. The Nature of X: Identification of a Central Carbide

When the central ligand of the cofactor was discovered, it was designated as a 'light atom X' (with X being either C, N, or O) [58]. At the resolution of 1.16 Å, the observed electron density features were best explained by a central N atom, coordinated in a μ^6-ligation within the trigonal prism formed by Fe2-Fe7. In spite of the high resolution, however, an ambiguity remained regarding this assignment. The presence of a central N atom in the cluster core could possibly be rationalized as bound product following N_2 cleavage, but this initial hope for a direct hint at the enzymatic mechanism of nitrogenase was soon shattered when Hoffman and coworkers first reported that ENDOR and ESEEM spectroscopies did not provide any evidence for an exchangeable N atom [60], and subsequently refined their study to point out that any nitrogen should have been detected through their approach, so that the central atom should be of a different nature [61]. Density functional theory (DFT) calculations of the cofactor were (and are) based on a broken-symmetry analysis by Noodleman and coworkers [62, 63]. These authors subsequently studied the influence of a central ligand and found either O^{2-} or C^{4-} to be unlikely as a central atom [64], and Hoffman and coworkers also concluded that they could not assign the central ligand unambiguously [65]. The nature of X thus remained elusive, as did its significance for catalysis. Several strategies to resolve this issue were pursued in the following years, but final clarification only came almost a decade after the initial discovery of the central ligand. Using Fe K_β X-ray emission spectroscopy, DeBeer and co-workers showed that an experimental valence-to-core X-ray emission difference spectrum of intact MoFe protein and a cofactor-free variant

(Δ*nifB* MoFe) was in far better agreement with the calculated signatures of an interstitial carbide (C^{4-}) that with that of a nitride (N^{3-}) or an oxide (O^{2-}) (Figure 7D) [66]. In parallel, a new crystallographic analysis of the protein at a resolution of 1.0 Å (Figure 7C), in conjunction with ESEEM studies of ^{15}N- and ^{13}C-labelled MoFe protein (Figure 7E) provided consistent evidence for the interstitial C^{4-} [67]. With three lines of experimental evidence now in firm agreement, it was the establishment of *S*-adenosylmethionine (SAM) as the carbon donor during cofactor biogenesis on the radical SAM enzyme NifB by Ribbe and coworkers that provided final closure to the debate [68]. This did not, however, reveal what role, if any at all, this central carbide plays for nitrogenase catalysis. It also did not advance the question regarding the site of substrate binding to the cofactor, as a carbon species quite obviously is not a reaction intermediate of dinitrogen cleavage. A series of theoretical studies that subsequently addressed the revised cofactor structure also did not reach an agreement on the effects that this unusual ligation had on the reactivity of the cluster [69–72].

4.2. The Electronic State of the FeMo Cofactor

With the atomic structure of the FeMo cofactor finally clarified, most of the relevant functional questions still remained very much unanswered. In particular, the underlying electronic structure of the site was under debate, leading to substantial variations in the outcome of theoretical studies depending on the underlying assumptions. In its resting state, E_0, the FeMo cofactor is a $S = 3/2$ system and being in a tetrahedral environment with weak-field sulfide ligands, all seven Fe ions will be in a high-spin configuration, leaving each with a total spin of $S = 5/2$. The possible coupling schemes for these sites were investigated by Noodleman and coworkers using spin-polarized broken-symmetry density functional theory (BS-DFT). Their energetically favored coupling scheme, BS7 [63], has since been widely used and forms the basis of the subsequent discussion. It features four Fe sites in a *spin-up* and three Fe sites in a *spin-down* configuration, while maximizing the amount of antiferromagnetic coupling of each site with its neighbors. Assuming all sites to be Fe(III), this leaves the total spin of the coupled system at $S = 5/2$, not counting the Mo ion, which was assumed to be Mo(IV) based on ^{57}Fe and ^{95}Mo ENDOR [73], a typical charge state for the metal in biological systems. Applying ^{57}Fe Mössbauer spectroscopy, Burgess and Münck assigned all seven Fe ions of the FeMo cofactor with a configuration of 4 Fe(II) and 3 Fe(III). Notably, however, this study predates the discovery of the central carbide in the cluster [74]. Alternative assignments were 2Fe(II)5Fe(III)Mo(IV) or 6Fe(II)1Fe(III)Mo(IV), both of which can also give rise to an overall $S = 3/2$ state. The remaining ambiguity regarding the electronic structure of the cofactor further hindered the emergence of a commonly accepted model to guide theoretical approaches, and the ensuing questions regarding the function of the site in dinitrogen reduction remained open.

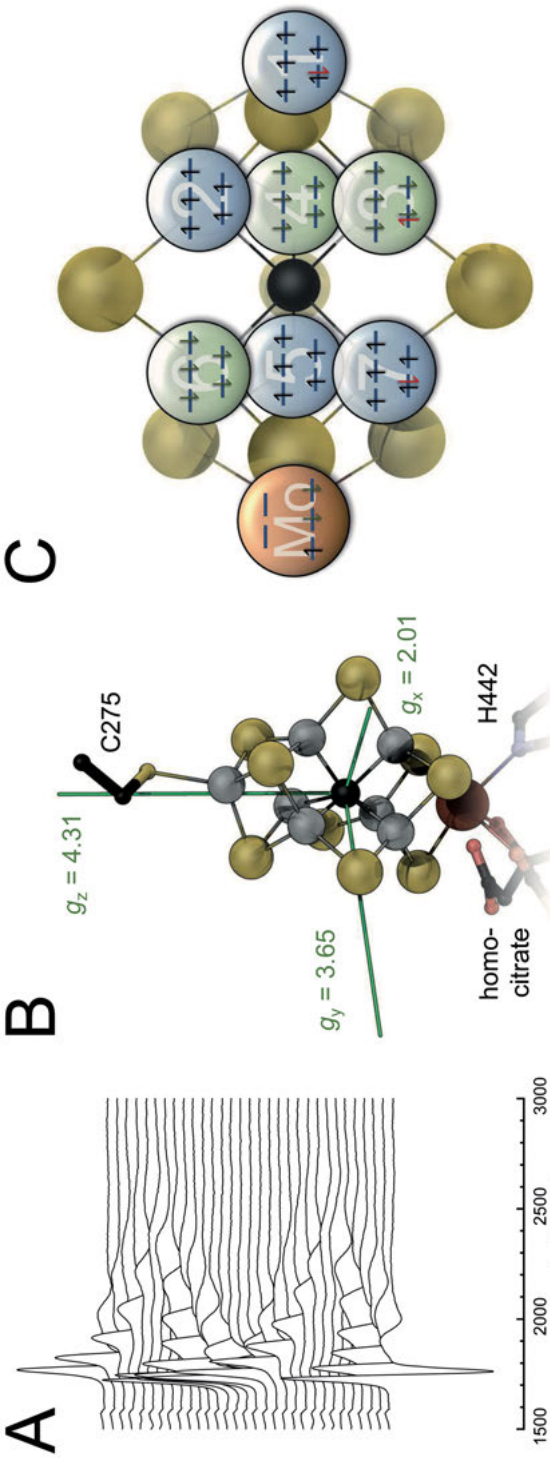


Figure 8. Elucidation of the electronic structure of the FeMo cofactor. (**A**) Single crystal EPR spectra of the MoFe protein from *A. vinelandii*; individual spectra were generated by a 5° rotation of the EPR tube. Tracing the angular dependence of the resonances allows for deriving the spatial orientation of the apparent *g*-tensor of the $S = 3/2$ system. (**B**) Overlay of the derived, apparent g-tensor with the structure of the FeMo cofactor, placed arbitrarily in the cluster center. The largest component of the tensor ($g_z = 4.31$) aligns with the three-fold symmetry axis of the cluster, while the shortest component ($g_x = 2.01$) points towards the mechanistically relevant Fe2-Fe6 edge. (**C**) Spin-localized model of the electron distribution in the resting state E_0 of FeMo cofactor, as derived from broken-symmetry DFT calculations, SpReAD analysis and Mo XAS.

4.2.1. Single-Crystal EPR of the FeMo Cofactor

The spectroscopic signature of the *as isolated* nitrogenase FeMo cofactor in electron paramagnetic resonance spectroscopy (EPR) is deceptively simple, but the correlation between the observed apparent *g* tensor with principal components $g_x = 2.01$, $g_y = 3.65$ and $g_z = 4.31$ and the structure of the cluster with its distinct three-fold symmetry along the longest axis is not immediately apparent. This question was clarified by an EPR study at X-band on single crystals of the enzyme, where the orientation of the molecular structure of the enzyme was determined by X-ray diffraction on the very same crystal that was subsequently rotated in the magnetic field of the EPR spectrometer (Figure 8A) [75]. Here, the longest principal component of the apparent *g* tensor, g_z, oriented fully (within error) along the threefold cluster axis. Of the other principal axes, $g_x = 2.01$ was in the plane formed by Fe1, Fe2, Fe6, and Mo in FeMo cofactor (Figure 8B). This observation was noted, but did not relate to any known functional property of the cluster at the time. It was, in retrospect, the first indication that the threefold symmetry of the cluster is not reflected in its electronic properties and that Fe2 and Fe6, together with the apical metals of this moiety, play a very different role. However, at the time this information could not be integrated, mainly due to the lack of a reliable model for the electron distribution within the site.

4.2.2. A Biological Mo(III) Site

The primary role of molybdenum in biochemistry is in catalyzing two-electron transfer reactions such as the one from nitrate to nitrite (nitrate reductase) or from dimethylsulfoxide to dimethylsulfide (DMSO reductase). To this end, the element typically alternates between the two diamagnetic redox states Mo(IV) ($4d^2$) and Mo(VI) ($4d^0$). The assignment of the ion in the FeMo cofactor as Mo(IV) was in line with this and supported by experimental data, as mentioned above (see Section 4.2). Also, in order to yield the observed $S = 3/2$ system, Mo(IV) would imply either 2, 4 or 6 ferrous (d^6) irons in the resting state of the cofactor. Data from *spatially refined anomalous dispersion* (SpReAD) studies on the enzyme (see Section 4.2.3), however, pointed to three Fe(II) sites in the cluster, which necessarily required an overall uneven number of electrons on the heterometal for a total $S = 3/2$. Assuming that the less stable Mo(V) state was most likely, DeBeer and coworkers conducted X-ray absorption spectroscopy (XAS) measurements of MoFe protein and in parallel also on model compounds, in particular on a $(Et_4N)[(Tp)MoFe_3S_4Cl_3]$ complex originally synthesized by Holm [76]. The authors used high energy resolution fluorescence-detected X-ray absorption spectroscopy (HERFD-XAS) at the Mo K-edge on this and other well-characterized Mo complexes as well as on the FeMo cofactor, finding that both the pre-edge and rising edge features were inconsistent with Mo(IV) for the enzyme and the model complex $(Et_4N)[(Tp)MoFe_3S_4Cl_3]$. Instead, the lower energy of the edge position indicated a more reduced species. DFT calculations not only supported this assignment as Mo(III), but also revealed that the resulting d^3 configuration of the heterometal represented a non-Hund $\alpha\alpha\beta$ ground

state due to electronic coupling with the surrounding, partially ferromagnetically coupled Fe sites [77]. This was significant for the rationalization of the overall spin system of FeMo cofactor, as it implied that the contribution of Mo to the total spin of the cluster was one-electron spin, redefining the combinatorial requirements for the redox states of iron in the cluster. The calculations showed that the crucial features of the $(Et_4N)[(Tp)MoFe_3S_4Cl_3]$ complex – Mo(III) and the spin-coupled non-Hund ground state – are most likely retained in the larger cofactor of the enzyme.

4.2.3. Assignment of Redox States by SpReAD Analysis

Nitrogenases contain only a one heterometal atom, making XAS an ideal and highly precise method for its characterization. For iron, the situation is quite different. Taking into account the P-cluster and the FeMo cofactor, the enzyme contains 15 distinct Fe sites, all of which will yield overlaying signals, making the deconvolution into their individual contributions and thus an assignment of absorption properties to the Fe sites virtually impossible. At the same time, the absorption of X-rays by the core electrons of an element at energies around an absorption edge have a profound effect on the diffraction properties of this particular element, in that they break the internal centrosymmetry of the diffraction process described by Friedel's law [78]. While in general the intensity of a reciprocal lattice point at a position (h, k, l) is identical to one at position $(-h, -k, -l)$, albeit with opposite phase angles, this relationship breaks down in the proximity of an absorption edge, leading to an *anomalous difference* Δ_{ano} in intensity, which is directly proportional to the X-ray absorption coefficient and thus provides an information equivalent to an X-ray absorption spectrum. As noted earlier, however, this information is encoded within a three-dimensional reciprocal lattice that can be sampled in a diffraction experiment, adding spatial resolution to the absorbance measurement [79]. The magnitude of Δ_{ano} can thus be sampled at any given point in space, and as the effect exclusively originates from the anomalous scatterers, it can then be refined individually for each atom of the given type. Through collecting a series of diffraction data sets along an absorption edge of the desired element, the refined values of Δ_{ano} thus provide a direct approximation of the individual absorption edges of single atoms in space. The SpReAD method was successfully applied to the well-understood test case of the [2Fe-2S] ferredoxin Fd4 from *Aquifex aeolicus* that contains localized Fe(II) and Fe(III) sites in its reduced state, due to the antiferromagnetic coupling of the two metals in the cluster [79]. In the refined data, the individual edge positions for Fe(III) where shifted higher in energy by approximately 2 eV with respect to the ones for Fe(II), in line with expectations from XAS and theory.

In this test case, the structure in question contained a dimer of the single-cluster protein Fd4 in the asymmetric unit, with a total of four individual Fe sites to be refined. Applying the SpReAD method to nitrogenase involved at least a complete $NifD_2K_2$ heterotetramer with two copies each of the P-cluster and the FeMo cofactor, amounting to 30 Fe sites. The collection of multiple datasets

from a single crystal initially raised the question whether photoreduction by X-rays might turn out to be a major obstacle for the applicability of the method. Here, experiments with different types of metal sites showed that this was indeed severe for sites with a positive midpoint redox potential, while most iron-sulfur clusters proved to be largely or entirely inert towards photoreduction within experimental error.

The FeMo cofactor is a $S = 3/2$ state as isolated, and the BS7 coupling scheme established by Noodleman and coworkers [62] in conjunction with a spin-coupled Mo(III) site in a non-Hund ground state [77] implies a formal distribution of 4 Fe(III) and 3 Fe(II). In an iron-sulfur cluster, spin delocalization across ferromagnetically coupled metal centers should be the rule, as was predicted for model complexes [77], but at the same time the maximized antiferromagnetic coupling in the BS7 scheme should support more localized spins. In a SpReAD analysis of the resting state of the FeMo cofactor, the refinement of individual anomalous contributions indeed indicated that Fe1, Fe3 and Fe7 seemed to be more reduced (lower energy edges), while Fe2, Fe4, Fe5, and Fe6 were more oxidized. An averaged absorption profile of the eight Fe sites of the all-Fe(II) P-cluster was used as an internal reference for Fe(II) and aligned almost perfectly with the profiles of Fe1, Fe3, and Fe7. Following this reasoning, the remaining four Fe ions would then be Fe(III), adding up to a well-defined system with the observed total spin of $S = 3/2$ [80]. The data was in agreement with the valence electrons being almost fully localized on individual Fe sites (Figure 8C). Through a reinterpretation of earlier data from Mössbauer spectroscopy by Münck and coworkers [74] and subsequent DFT calculations, the SpReAD analysis was largely confirmed, although the respective ferromagnetic coupling of Fe3/Fe4 and Fe5/Fe7 resulted in delocalization that was not reproduced in the SpReAD data [81]. Most importantly, however, both approaches agreed on largely localized electrons in the system and on Fe2 and Fe6 as the most oxidized sites in the E_0 state of FeMo cofactor [82].

4.3. Ligand Exchange in Nitrogenase Cofactors

While moving forward on the way to a comprehensive electronic description, this deepened understanding of the FeMo cofactor still failed to reveal clues regarding the precise site and mode of substrate binding. Even since before the first structure determination of the enzyme, synthetic inorganic chemistry has provided numerous suggestions for the nature of the nitrogenase clusters in the form of model compounds. The pioneering work of the Holm group had produced complex clusters (Figure 9A) [83, 84] with similarity to the multi-metal moieties that by then were known from the structures reported by Kim and Rees [35, 36], and Tatsumi and coworkers found that a topological equivalent of P-cluster can be assembled in a single step from simple [2Fe-2S] precursors at high yields [85], and Holm achieved a similar model that included apical molybdenum ions along an entirely different route [86] (Figure 9B). With a central carbide ion none of the metal sites of the cluster remained coordinatively

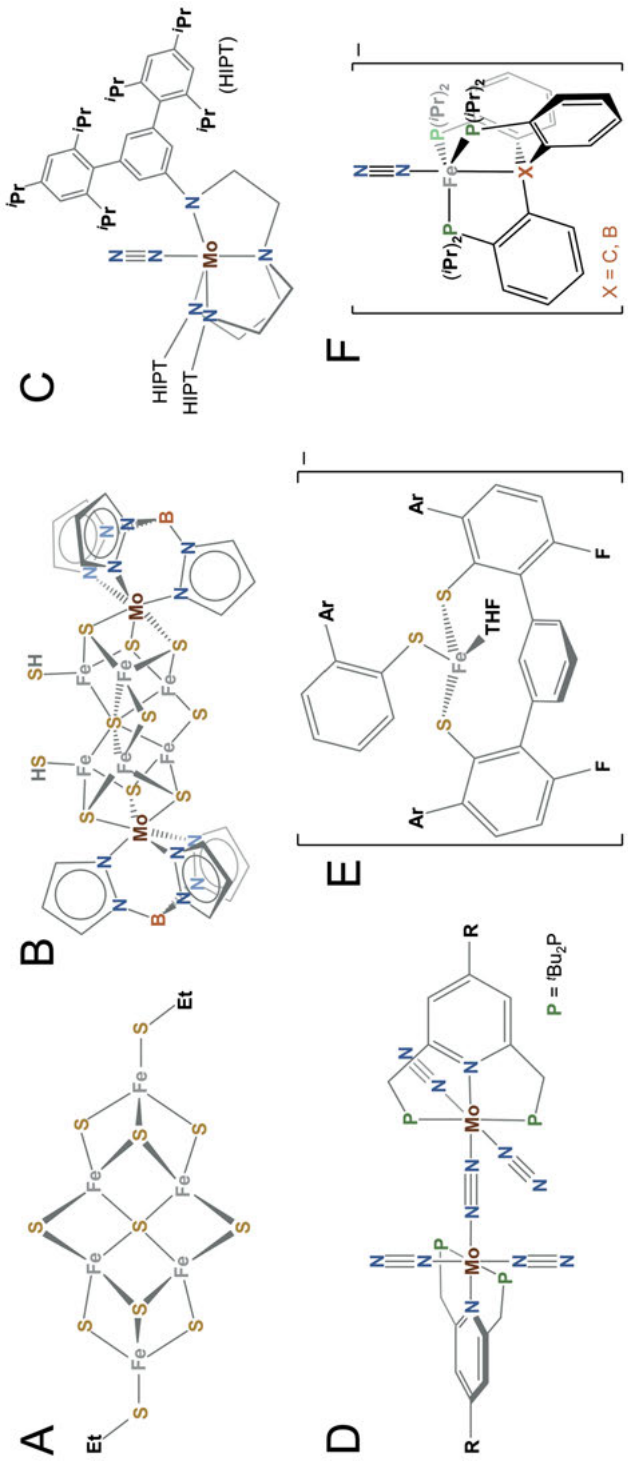


Figure 9. Synthetic models addressing the structural and functional properties of nitrogenase clusters. (**A**) The early $[Fe_6S_9(SEt)_2]^{4-}$ complex by Holm and coworkers [84], providing strong topological analogies to P-cluster. (**B**) The $[(Tp)_2(Mo)_2Fe_6(S_9(SH)_2]_2$ cluster by Zhang and Holm [86] underlines the similarities between the FeMo cofactor and the P^N state of P-cluster. (**C**) The Mo-HIPT-triamidoamine complex by Schrock [91] was the first synthetic catalyst for N_2 reduction, yielding approximately 8 equivalents of NH_3. (**D**) Nishibayashi [92] introduced PNP pincer ligands to Mo that effect the reduction and cleavage of the bridging N_2 molecule. (**E**) In the Fe *tris*-thiolate compound by Holland [101], the binding of N_2 leads to the elimination of a bound sulfide. (**F**) The tetradentate triphosphine ligand of Peters [88] was inspired by the concept of an *exo* binding of N_2 to a core Fe of the FeMo cofactor, and achieved good catalytic efficiencies.

unsaturated. Studies hypothesized on the six central iron atoms as the most likely binding sites for any ligand [87]. While the iron atoms in these positions are tetrahedrally coordinated, the Fe ion is closer to the plane formed by its three sulfide ligands than to the center of the tetrahedron, implying that an external ligand could distort the arrangement towards a pentacoordinate trigonal bipyramidal geometry. A series of model complexes by Peters and coworkers have built and expanded on this motif, leading to catalytically active compounds for N_2 reduction based on a single, highly-reduced iron site with a *trans*-axial Si, B or C in an arrangement with three coplanar phosphine ligands (Figure 9F) [88] that in later studies could be replaced by sulfides [89]. Ranking among the most efficient small-molecule catalysts for mild activation of dinitrogen, these compounds still require a strongly reduced metal site as well as electron donors with a low midpoint potential that are not obviously compatible with the aqueous milieu the enzymes reside in. The role of the central iron sites for substrate binding thus remained under debate, but in particular due to the absence of an obvious alternative binding site, the position remained among the favorites in literature.

A second perspective on substrate binding gained new interest with the assignment of Mo(III) in the FeMo cofactor. Coucouvanis and coworkers early on proposed Mo as a candidate for substrate ligation in nitrogenase [90], and most remarkably, the first catalytically active model compounds for N_2 reduction by Yandulov and Schrock (Figure 9C) [91] and Nishibayashi et al. (Figure 9D) [92] both were Mo-based entities. In these systems, the stabilization of a Mo(VI) nitride typically represents a crucial step in the reaction mechanism, while a reduced Mo(III) species is required for N_2 binding [93, 94]. In the FeMo cofactor, the octahedral coordination of Mo(III) is complete, so that the binding of any further entity would necessarily require the dissociation of an existing ligand. With the unusual coordination of the Mo ion by homocitrate, this possibility cannot easily be ruled out. Coordinating the metal through a carboxylate oxygen and its hydroxyl group, homocitrate is located in the heart of the nitrogenase enzyme, adjacent to an internal cavity filled with water molecules. While likely related primarily to the maturation process of the enzyme, in which the insertion of the catalytic cofactor into the apoprotein matrix represents the final step, this cavity also provides substantial conformational freedom for the organic ligand, and its dissociation from the heterometal during the catalytic cycle is – at the very least – not sterically impossible.

Lessons from synthetic chemistry thus seem to imply that both metals found in FeMoco can be suitable coordination sites for the substrate of the enzyme, but a distortion of the cluster from its resting state conformation, or even the exchange of a ligand will be required to accommodate for an N_2 molecule. According to the Lowe-Thorneley scheme for nitrogenase catalysis, this only occurs after a four-electron reduction of the cluster, and a rearrangement of the cofactor in preparation for ligand binding may well occur in this process. However, none of the available structural data provided any indication to discriminate the hypothetical modes of binding of N_2 or any of the other known substrates that can be reduced by nitrogenase.

4.3.1. Binding of Carbon Monoxide and Selenide

A known data point of potential relevance for elucidating the modes of ligand binding to dinitrogenase was the inhibition of the reduction of any substrate other than protons by carbon monoxide [95]. The diatomic molecule is a strong-field ligand that readily binds to various metal cations in a terminal or μ-bridging fashion. In nitrogenase, its binding occurs at a less reduced state than the E_4 state required for N_2 binding, most likely at the two-electron-reduced E_2 level. In consequence, there is no direct competition of N_2 and CO for binding to the metal cofactor even if both were binding to the same site, and indeed the observed mode of inhibition was non-competitive [96]. Upon further analysis, two distinct CO-bound states could be differentiated by EPR and vibrational spectroscopy, termed the low-CO and high-CO states with one and two molecules of the ligand bound to the cofactor, respectively [96]. Cramer and coworkers identified a series of photolysis products from the high-CO state indicative of a variety of binding modes [97, 98], but structural evidence for this particular ligand-metal interaction was lacking. Hereby, a major complication for the structural characterization of a CO adduct of the FeMo cofactor was that its formation required the enzyme to reach the E_2 state, which can only be achieved under turnover conditions, i.e., in the presence of Fe protein, a suitable electron donor, and ATP. The dynamic nature of the Fe-MoFe complex interaction precludes this to be carried out *in crystallo*, so that the experimental challenge that had to be overcome was to inhibit the enzyme under electron flux to the active site, then isolate inhibited the MoFe protein, grow crystals to a suitable size for diffraction experiments and flash-cool these crystals in liquid nitrogen. Rees and coworkers succeeded in this endeavor, leading to a high-resolution structure of CO-inhibited MoFe protein in the low-CO state [99]. Defying all expectations, the crystallographic analysis at 1.5 Å resolution unambiguously revealed the position and mode of CO binding (Figure 10A). Most notably, the second outstanding observation in this structure was that the CO ligand did not simply add to the cofactor, but would replace and occupy the position of a constituent atom of the heterometal cluster, the μ^2-bridging sulfide S2B. The resulting bridging carbonyl connecting iron sites Fe2 and Fe6 is a common chemical motif, but without a three-dimensional structure no available methodology had been able to anticipate the unprecedented replacement of a sulfide. Furthermore, the binding of CO was reversible, and when CO was removed under continued turnover, a subsequent structural analysis revealed that sulfide S2B had returned to its original bridging position, taking the enzyme back to the well-known resting state [99]. The fate of this sulfide in the CO complex remained unclear, as no binding site for the HS^- anion that is expected at physiological pH was found. The CO complex of course did not show the binding for a substrate molecule, but rather an inhibitor. Nevertheless, the binding of CO and sulfide replacement was revealed to be reversible, making a functional role for this type of conformational flexibility a serious point for further consideration. In an important study on a synthetic Fe-S model compound, Holland et al. showed that addition of N_2 in their system

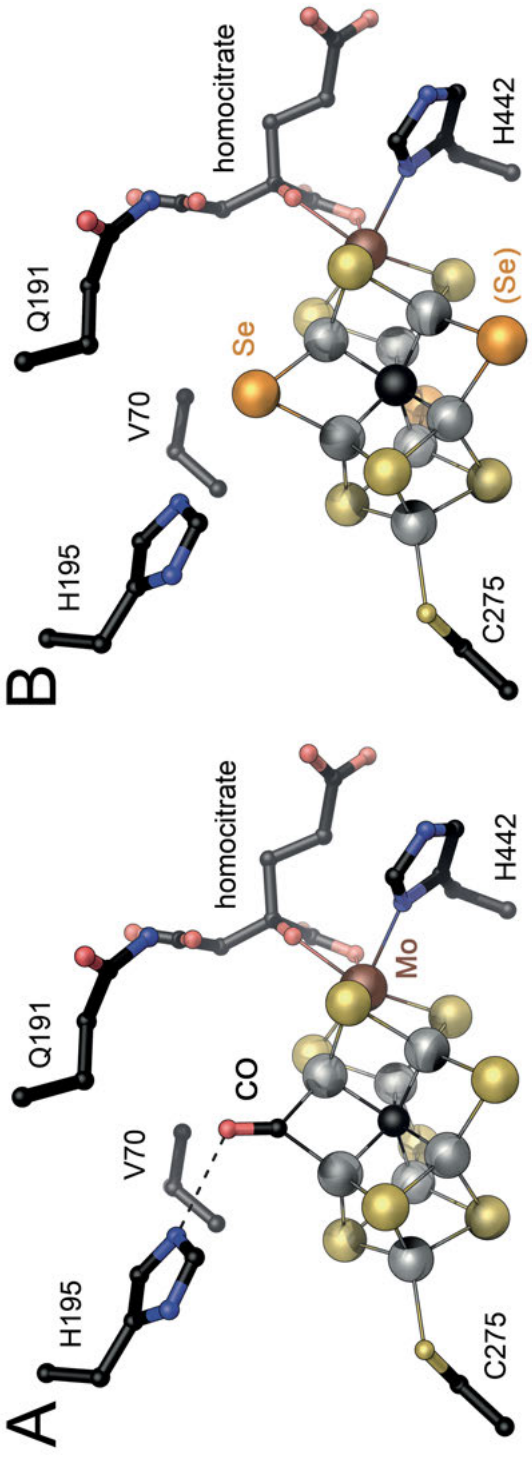


Figure 10. Binding of ligands to the FeMo cofactor. (**A**) The CO adduct was generated under turnover conditions by Rees and co-workers, leading to the reversible replacement of the bridging sulfide S2B [99]. (**B**) By turning over the enzyme with $SeCN^-$, S2B and, to a lesser degree, also S3A and S5A, could be replaced by Se, as shown by anomalous dispersion measurements near the the K-edge of selenium.

led to Fe–S bond dissociation and the formation of an Fe-N_2 adduct, highlighting that the assumption of the binding of a small ligand leading to Fe–S bond dissociation is indeed chemically feasible (Figure 9E) [100, 101].

The reversible nature of CO binding and the temporary replacement of sulfide S2B, which could not be reliably located in the CO-bound structure, prompted Rees and coworkers to also attempt the replacement of S2B with a spectroscopically and structurally more easily detectable selenide, Se^{2-}. While this was not successful through the simple addition of the selenide anion, the desired outcome was achieved by turning over the enzyme with selenocyanate ($SeCN^-$). The reduction of the known substrate cyanide (CN^-) to ammonium (NH_4^+) and methane (CH_4) left the enzyme in a situation corresponding to the resting state E_0, but an analysis of the electron density maps of a high-resolution structure – and in particular the anomalous scattering signal of selenium – provided clear evidence that Se^{2-} had replaced S^{2-} at the bridging position between Fe2 and Fe6 [102]. Moreover, under continuous turnover the authors not only observed the quantitative replacement of S2B, but successively also of the two remaining bridging sulfides, S3A and S5A (Fig. 10B). The remarkable flexibility of these sites is a feature of the unique cofactor structure, providing further indication that the central carbide acts to add stability to the arrangement, providing a central pivot that allows the peripheral bridges to flexibly exchange in a manner that likely depends on the oxidation state and the concomitant protonation of the bridging atoms in order to facilitate their dissociation from the metal core. Whether all three replacements play a functional role during nitrogenase catalysis remains to be clarified, but this experiment convincingly revealed the unprecedented amount of structural flexibility offered by the complex iron-sulfur-based metal clusters encountered in ancient enzymatic pathways [103].

4.3.2. Nitrogenase Catalyzes Fischer-Tropsch Chemistry

By providing the first ligand-bound structure, CO had been highly instrumental for understanding the unprecedented flexibility and complexity of the nitrogenase active site cofactor. As a strong-field ligand, CO is well known to form stable complexes with various metal cations, but in nitrogenase, a redox enzyme able to cleave the N_2 triple bond with a bond enthalpy of –946 kJ·mol^{-1}, the fact that CO inhibits its activity rather than being itself subject to reductive conversion was quite remarkable. It was nevertheless unexpected when in 2010 Ribbe and coworkers found that the alternative V-dependent nitrogenase enzyme from *A. vinelandii* did in fact reduce CO, although at very low rates [104]. Moreover, the main product of this reaction was not CH_4, but 93 % of the carbon were found as ethylene [105], implying that C–C bond formation regularly occurs between two substrate CO molecules, lending potential mechanistic relevance to the hi-CO state described above [106]. While the yield of methane was below 1 % in these reactions, further reduction products were ethane (3 %) and propane (2.5 %). V-nitrogenase thus has the ability to directly form hydrocarbons

from CO, a reaction analogous to the well-known Fischer-Tropsch process. The industrial process largely yields the comparatively low-value product methane that requires further energy-intensive conversions to obtain useful products, while the formation of a carbon chain is the primary outcome of the enzyme-catalyzed reaction. This finding was of potential relevance for applications such as the production of biofuels from CO, and subsequent work by the same authors showed that, other than N_2 reduction, the conversion of CO can also be achieved in an ATP-independent fashion using isolated the FeV cofactor in conjunction with chemical reductants. Interestingly, in these experiments the formation of hydrocarbon chains up to a length of 7 carbon atoms was observed, and consistently a significant proportion of the reaction products retained a terminal unsaturation [107]. The different product spectrum obtained with isolated cofactor likely means that in the confined environment of the active site cavity of the enzyme longer chains cannot be formed for steric reasons. Subsequent studies then found that the reduction of CO also occurs in Mo-nitrogenase, but that the activity towards this substrate is lower than that of V-nitrogenase by a factor of 800 [108], which is well in line with the observation that the CO complex of the MoFe protein was stable on the timescale of a crystallization experiment. CO thus largely acts as an inhibitor of substrate reduction by nitrogenase (with the important exception of proton reduction to form H_2), but the fundamentally different reduction rates observed for MoFe and VFe proteins have been one of the most outstanding reminders that the two variants of nitrogenase enzymes exhibit differential functionalities.

5. VANADIUM NITROGENASE AND THE FeV COFACTOR

5.1. Alternative Nitrogenase Systems

Today, the free-living soil bacterium *A. vinelandii* is the prime model diazotroph for the biochemical and biophysical analysis of nitrogenases. Its genome encodes all three known nitrogenase variants, the Mo-, V-, and Fe-dependent enzymes, in four distinct loci on its chromosome [109]. *A. vinelandii* preferably produces molybdenum nitrogenase, the Nif system, which is also the most active of the three, and switches to the vanadium nitrogenase, the Vnf system, upon molybdenum depletion [110]. Only if the element vanadium is additionally absent, the third system, Anf, is induced, producing a nitrogenase variant that solely relies on iron as a metal for electron transfer and catalysis. The catalytic activity of V-nitrogenase towards N_2 amounts to approximately one third of that of Mo-nitrogenase, and the Fe-dependent enzyme then shows only about half of the activity of V-nitrogenase. All this is in good agreement with a differentiated pattern of expression regulated by the availability or lack of the respective heterometal for the formation of the active site cofactor. Interestingly, this simple picture has more recently become somewhat complicated not only by the finding

that only V-nitrogenase reduces CO (see Section 4.3.2), but also by a more recent report indicating that Fe-nitrogenase, but not its two heterometal-dependent relatives, is able to convert CO_2 directly. Harwood et al. observed that this reaction can sustain the growth of a methylotrophic co-culture, indicating that the product of the reaction is (at least in part) methane [111]. Furthermore, while a multitude of structural and accessory genes for all three nitrogenases are encoded in a distinct *nif*-, *vnf*- and *anf*-cluster, respectively, the aforementioned NifB protein is part of a fourth locus. It provides a topologically complete cluster precursor of presumed stoichiometry [8Fe-9S-C], and this moiety – termed NifB-co or L-cluster – serves as precursor for the active site cofactors of all three nitrogenase systems [112].

5.1.1. Architecture of VnfH

Although in *Azotobacter vinelandii* the three variants of nitrogenase share this NifB-co precursor, each system features a unique, if closely related, Fe protein component. These orthologs of NifH from the Mo nitrogenase are VnfH for vanadium nitrogenase, and AnfH for the iron-only enzyme. The H-genes are coexpressed with the according structural genes for the three nitrogenases, and all three share an architecture consisting of a homodimeric protein with the structural features of a P-loop NTPase, with the connection between the monomers formed almost exclusively by the shared [4Fe-4S] cluster, for which each monomer provides two cysteine ligands. *A. vinelandii* NifH and VnfH have a sequence identity of 91 %, and the recently reported three-dimensional structure of VnfH confirms the expected degree of similarity (Figure 11A) [113]. Notably, both Fe proteins can functionally complement each other, with no discernible difference in functionality with either system. The structure of VnfH has an ADP molecule bound to each monomer, leading to a conformation of the dimer that leaves a central cleft between monomers. The fact that the ADP-bound conformations of NifH and VnfH are virtually identical provides further support that this is not a random and flexible conformation, but indeed a defined state that is retained as a feature of the protein. This ADP-bound state of Fe protein supposedly represents the point in the reaction cycle where an electron has been transferred and the bound ligands are the remnants of the hydrolysis of ATP that was triggered by complex formation with the actual nitrogenase. From here, Fe protein obtains another electron from its redox partner to then exchange its ADP ligands for ATP, readying the system for a new round of electron transfer to the enzymatically active component. The structures were obtained in the reduced state, and it remains to be clarified whether a change of redox state of the [4Fe-4S] cluster bears structural consequences for the conformation of the Fe protein.

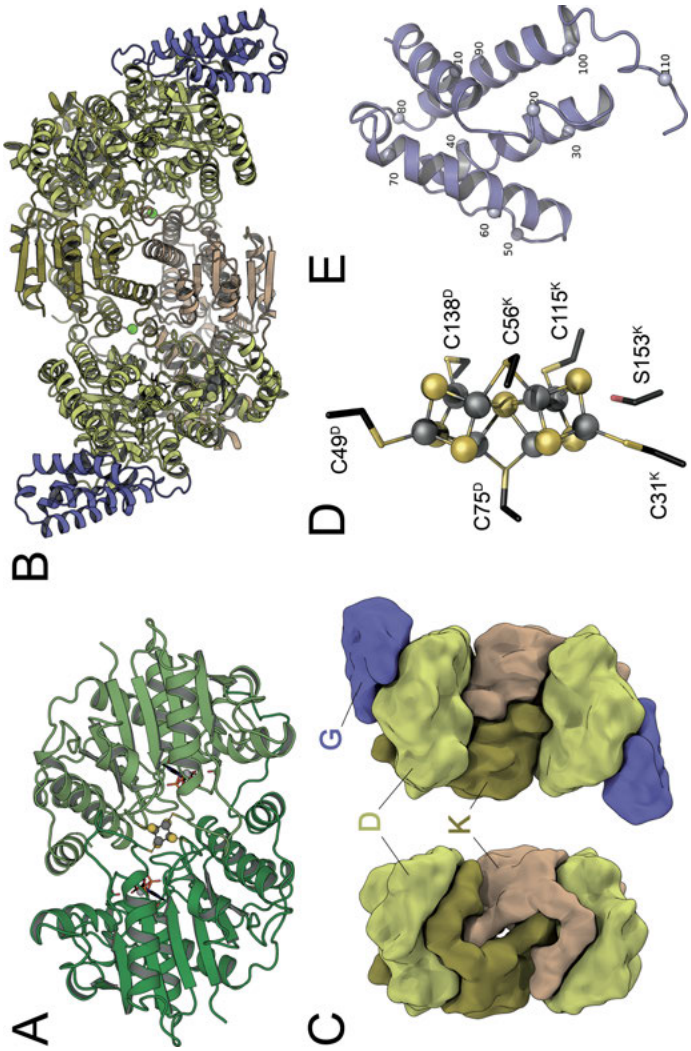


Figure 11. The V-dependent nitrogenase from *A. vinelandii*. (**A**) VnfH, the Fe-protein component is almost identical to NifH and able to functionally complement its homolog. (**B**) VFe protein is a $VnfD_2K_2G_2$ heterohexamer of 230 kDa; while the D- and K-subunits show high homology to NifD and NifK, respectively, the VnfG subunit (purple) is absent in the Mo-dependent enzyme. (**C**) A low-resolution surface depiction of the MoFe and the VFe protein highlights the similarities; the crescent-shaped extension of the N-terminus of NifK gives the impression of a structural difference that is not reflected in the environment of the two metal clusters. (**D**) In the VFe-protein, P-cluster is almost identically arranged to the electron-transfer site in MoFe protein, but typically only undergoes a one-electron transfer reaction accompanied by a change in the position of Fe6. (**E**) The additional VnfG subunit forms a four-helix bundle without additional cofactors and binds exclusively to VnfD; its C-terminal extension is in close proximity to the presumed entry site for protons to reach His180.

5.1.2. Architecture of VnfDKG

When a three-dimensional structure for the VFe protein – once more from *A. vinelandii* – was solved in 2017, it revealed the expected architecture of a heterotetrameric D_2K_2 assembly that shared its key structural features with the known MoFe protein (see Section 2.2, Figure 11B) [114]. The pseudosymmetric D- and K-subunits held a [8Fe-7S] P-cluster at their interface, and in spite of earlier reports suggesting that this cluster differed from its analogue in MoFe protein, the dithionite-reduced state had the exact same architecture of two cubane-type subclusters fused *via* a central μ^6-sulfide (Figure 11D). Similarly, the active site cofactor – here a FeV cofactor – was cradled between the three Rossmann-type subdomains of the VnfD subunit, in full analogy to the FeMo cofactor in Mo nitrogenase (see Section 5.2). With an overall similar fold and subunit arrangement, a structural difference between the two nitrogenases was found in a N-terminal extension of NifK that is absent in VnfK. While in MoFe protein the N-terminus of NifK wraps around the surface of the adjacent NifD in three extended α-helices, the shorter VnfK chain of the VFe protein forms a β-strand that adds itself to the central β-sheet of the third Rossmann-domain of VnfD, conveying a visible difference to the outward appearance of both proteins that obfuscates the high degree of conservation of the features known to be relevant for the correct operation of both enzymes (Figure 11C). As in the MoFe protein, the K-subunits of the VFe protein are also bridged by two metal cations in a symmetrical arrangement. However, while in the MoFe protein these sites were initially modelled as Ca^{2+} [35, 58] and later shown to be at least partly replaced by a 16th iron ion [115], the binding geometry and electron density observed in the 1.35 Å resolution structure of the VFe protein strongly indicated that the ion employed here is Mg^{2+} [114]. In either case, no indications exist that the bridging cation is of relevance beyond a structural role in stabilizing the heterotetrameric arrangement of the nitrogenases, but its presence and positioning are interesting also in the context of an observed half-side reactivity of the enzyme [116, 117], which implies some kind of conformational communication between the two DK-heterodimers that still remains to be understood in molecular detail.

As a major differentiation from Mo-nitrogenase, the VFe protein contains an additional subunit encoded by the *vnfG* gene located immediately downstream of the structural genes *vnfDK* in the *A. vinelandii* genome [109]. Little was known about the structure, role, and even the stoichiometry of this additional component prior to the availability of a crystal structure [114]. VnfG is a 113 aa protein that folds into a globular domain with four α-helices (Figure 11E). Two copies of VnfG were found in the VFe protein, conveying an overall composition of a $VnfD_2K_2G_2$ heterohexamer of approximately 240 kDa. VnfG is in exclusive contact with VnfD, it holds no cofactor and its location will presumably not make it interfere with the Fe protein VnfH if the resulting complex of the two components is similar to the one observed for NifH and the MoFe protein [114].

5.2. The FeV Cofactor Contains an Unexpected Ligand

The catalytic heart of the VFe protein is the FeV cofactor, and the general expectation prior to the availability of a crystal structure was that the overall structure of the FeMo cofactor would be retained, but the apical heterometal molybdenum was replaced by vanadium. This is in line with the current understanding of cofactor biogenesis, where in *A. vinelandii* only a single copy of *nifB* is present that should therefore be required for the synthesis of the precursor L-cluster for all three nitrogenases. In contrast, heterometal insertion occurs on the scaffold complex EN, a structural homolog of the enzymatic component itself [118], and here the *nif* cluster comprises a NifEN machinery, while the *vnf* cluster features a separate VnfEN system [109]. The structure of the VFe protein confirmed the expectations. The FeV cofactor is a heterometal cluster with a central, interstitial carbide as predicted by XAS [119], and it contains a single vanadium ion taking the position of Mo in the FeMo cofactor. As in the case of its counterpart in the MoFe protein (Figure 12A), the vanadium ion is coordinated by a histidine residue from the protein (VnfD) and an organic *R*-homocitrate ligand that is synthesized from 2-oxoglutarate and acetyl-coenzyme A by the NifV protein, another nitrogenase maturation factor for which only a single copy is present in the *A. vinelandii* genome. Interestingly, the bond distances to vanadium in the VFe protein were almost identical to those found at Mo in the MoFe protein, leading also to an overall structure of the metal cluster that deviated only minimally from FeMoco [114]. However, the structure also revealed one modification at the cluster that was entirely unexpected. While both cofactors are built around a central trigonal prism of six iron ions with an interstitial carbide, the FeMo cofactor features three μ^2-sulfides, termed S3A, S5A and S2B that form the characteristic edges of the cluster [58]. In the FeV cofactor, only S2B and S5A were retained, while the S3A sulfide that bridges Fe4 and Fe5 was replaced by a 1,3-bridging tetraatomic ligand [114]. At 1.35 Å resolution, the electron density map allowed to assign this ligand as carbonate, CO_3^{2-}, showing symmetric bond distances as well as a hydrogen-binding environment that excluded protonation of any of its oxygen atoms (Figure 12B). However, the presence of a nitrate anion, NO_3^- (instead of CO_3^{2-}), could not be excluded based on the electron density analysis alone, as both ligands have the exact same number of electrons and thus yield virtually indistinguishable features. Strong arguments for carbonate were its straightforward availability as the (hydrated) product of the oxidation of glucose that supplies the energy for diazotrophy in *A. vinelandii*, as well as the fact that the mere presence of nitrate in the cell is one of the triggers leading to an immediate 'switch-off' of nitrogenase activity, being a signal that bioavailable nitrogen is present and the energy required for nitrogen fixation may be saved.

The 1,3-bridging binding mode of carbonate leads to a Fe4-Fe5 distance of 2.76 Å [114], compared to 2.61 Å in the FeMo cofactor [67], representing the most significant change in metal-metal distances in the two clusters. The functional role of this ligand change remains unclear. Within VnfD, carbonate is tightly embedded in a loop of the protein (residues 335–340 in *A. vinelandii*)

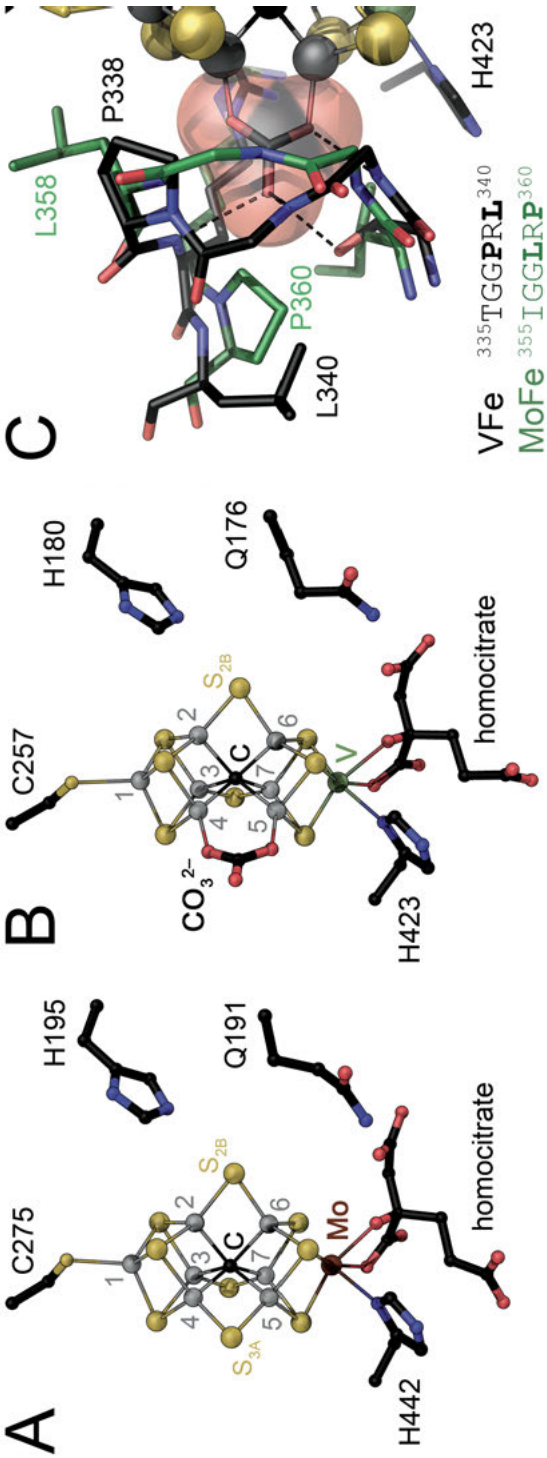


Figure 12. The FeV cofactor of V-nitrogenase. (**A**) Structure of the FeMo cofactor with numbering of the Fe ions and bridging sulfides; structurally and catalytically relevant amino acid residues in the surrounding of the cluster are shown as sticks. (**B**) The FeV cofactor shares the overall topology of the FeMo cofactor; Mo is replaced by V, while the protein surroundings are highly conserved, the bridging sulfide S3A in the FeMo cofactor is replaced by CO_3^{2-}. (**C**) The binding of CO_3^{2-} requires a structural change in the loop that in the MoFe protein cradles atom S3A. In the VFe protein, a proline and a leucine are swapped with respect to the MoFe protein, leading to the enlargement of a binding pocket that is required to accommodate CO_3^{2-}.

that only differs from the corresponding loop in NifD by the swap of a proline and a leucine residue (Figure 12C). This is sufficient to generate a cavity in VnfD that can hold carbonate, while the shifted proline in NifD would block this space. Considering the *ex situ* biosynthesis of the nitrogenase cofactors and their insertion into apo-nitrogenases only in the very last step, it is assumed that during FeVco biogenesis the S3A ligand that is still present in the L-cluster precursor of all cofactors, is removed and exchanged for carbonate. Once the cluster is inserted into its target protein, the tight interaction with the aforementioned loop region then precludes any further dissociation of the ligand. There are no indications to date that the carbonate ligand of the FeV cofactor is directly involved in VFe protein catalysis, in particular also not in the reduction of CO that occurs here. Although the respective factors remain to be identified, the removal of sulfide and insertion of carbonate is a multi-step process that will not occur spontaneously. It might take place on VnfEN, but likely requires further enzymes, and the benefit of this process for the enzyme is not yet understood. V-nitrogenase only shows approximately one third of the catalytic activity of Mo-nitrogenase, and one might assume that the insertion of carbonate represents an optimization without which the performance of the alternative system might be even poorer. Alternatively, the carbonate ligand is required to modulate the electronic structure of the FeV cofactor in order to allow it to be functional in N_2 reduction in the first place. Interestingly, XAS revealed that vanadium in the FeV cofactor is also in the (III) oxidation, like Mo in FeMo cofactor [120]. As a consequence, V has a d^2 configuration, while Mo in FeMo cofactor is in a non-Hund d^3 configuration [77]. Thus, if both clusters reside in a $S = 3/2$ multiplet ground state as commonly assumed, the same considerations as made for MoFe protein above (see Section 4.3.2) imply that one Fe site in the FeV cluster must be more reduced than in its Mo-containing counterpart. The role of the additional ligand may be related to this, and here the singly charged nitrate would have been appealing, as it would allow for the overall charge of the cluster to stay the same as for FeMo cofactor. A carbonate ligand does not provide this charge compensatory effect, but may still fundamentally influence the spin distribution within the system and thus its reactivity towards substrates.

5.3. Ligand Binding in a Turnover State

In spite of the carbonate ligand, the three-dimensional structure of vanadium nitrogenase highlighted the high degree of conservation between the different types of nitrogenase enzymes. Beyond cofactor topology, the homocitrate ligand and the interaction with the protein exclusively through a cysteine residue to Fe1 and a histidine to the heterometal, it is also the surrounding protein environment in the second coordination sphere of the cluster that remained almost unchanged. In particular, the bridging sulfide S2B of the FeV cofactor is in close proximity to two conserved residues, His180 and Gln176 in VFe protein. When their respective counterparts in the MoFe protein, His195 and Gln191, were exchanged by site-directed mutagenesis, the effect on nitrogenase activity had

been substantial [121, 122], although in all structures the glutamine residue was observed to point away from the cofactor itself. Its role therefore remained somewhat undefined, while the adjacent histidine formed part of a tight and conserved hydrogen-bonding network that extends up to the protein surface and suggests a possible role in proton provision during the reduction of substrates [123], although alternative pathways were also proposed based on theoretical considerations [124]. Sulfide S2B stands out as the ligand that is exchanged for CO under turnover conditions in the MoFe protein (see Section 4.3.1) and also the first of the sulfide bridges to be replaced by selenide [99, 102].

The initial structural analysis of the VFe protein revealed a resting state structure with S2B firmly in place, and the corresponding samples showed an EPR signature that was in line with data published previously from various laboratories [114]. However, the further optimization of isolation and crystallization protocols for the VFe protein led to consistent changes in the observed EPR signals, in that not the position of the signals in the complex spectrum, but rather their relative intensities shifted in a reproducible manner [110]. Although the FeMo and FeV cofactors are remarkably similar in most aspects, it bears mentioning that the EPR signatures of the latter are quite distinct and indeed significantly more complex than the former. The EPR spectrum of the MoFe protein in the resting state (E_0) is a broad, rhombic $S = 3/2$ signal with apparent g-values of 2.01, 3.65, and 4.31 (Figure 13A, orange) that orients with its g_z component along the pseudo-threefold axis of the cofactor (Figure 8B) [75]. The currently proposed electronic model for the cluster is well in line with the observed spectrum and SpReAD analysis indicates the three unpaired electrons to be largely localized to Fe1, Fe3, and Fe7 of the cluster [80]. In the VFe protein, the width of the signal is quite similar at nearly 3000 G, and bands at apparent g-values similar to the ones for the MoFe protein are present. However, the most prominent feature of the EPR spectrum is a far narrower, nearly axial signal at 3470 G, and the spectrum also contains a further band at 1300 G (Figure 13A, black). Based on EPR and Mössbauer studies, Münck and coworkers have described the spectra of the MoFe protein with a system of effective g-values of $g_x = g_y = 2.003$ and $g_z = 2.03$ [125]. The features of both the MoFe and the VFe protein can then be explained by a difference in the rhombicity parameter E/D representing the lower Kramer's doublet of the $S = 3/2$ system that leads to substantial variations in the apparent g tensor for both systems. The spectrum of the VFe protein is frequently rationalized as a mixture of an $S = 1/2$ signal at 3470 G, the broad $S = 3/2$ signal and an additional minor $S = 5/2$ contribution at 1300 G, meaning that contrary to the MoFe protein, this nitrogenase orthologue is poised between at least three electronically distinct states. The changes observed during optimization of the isolation procedures can be rationalized as a change in the population of these states, but not their chemical nature (Figure 13A, green).

In spite of differences in their EPR properties, all VFe preparations retained their capacity to form high-quality crystals that diffracted to near-atomic or true atomic resolution. For the preparations with an altered EPR signal, a 1.2 Å crystal structure was obtained and revealed an unexpected change at the FeV cofactor [123]. At the center of the active site cluster, the carbonate ligand, as well as

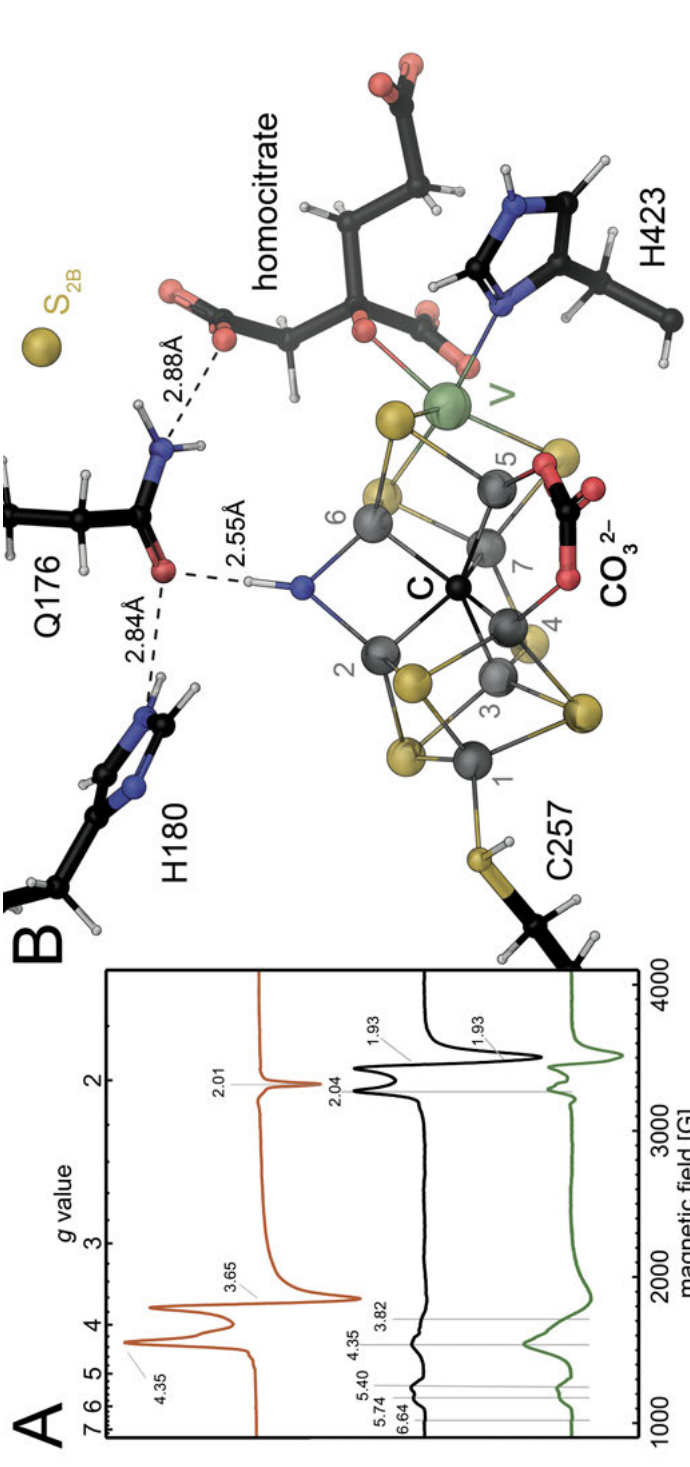


Figure 13. Structural analysis of the E_6 state of the nitrogenase catalytic cycle. (**A**) Continuous-wave, X-band EPR spectra of the resting states E_0 of the MoFe protein (orange) and the VFe protein (black); preparations of the VFe protein with an altered intensity ratio of the individual features indicate a different functional state. (**B**) The structural analysis of this state shows sulfide S2B replaced by a protonated light atom, N or O; an NH adduct would represent a late-stage intermediate of the catalytic cycle, E_6, replacement of S2B in this structure leads to a conformational change of the side chain of glutamine 176, forming a strong H bond to histidine 180, the proton donor during catalysis, and opening up a holding site for sulfide S2B, rationalizing its observed re-binding upon returning to the E_0 state.

the bridging sulfide S5A were in place exactly as observed in the resting state. In the third bridging position, however, sulfide S2B was replaced just as in the CO complex of the MoFe protein. In the VFe protein, the moiety replacing S2B was a light atom rather than the elongated CO ligand seen previously. The light atom is bound to Fe2 and Fe6 in a μ-bridging arrangement with bond distances of 2.0 Å, creating a sufficiently large distance to residue H180 for the adjacent Q176 to rotate its entire side chain by more than 130°, and to form a short hydrogen bond with its amide oxygen to the imidazole moiety of H180. This swing of Q176 in turns opened a pocket previously occupied by its amide group that in the ligand-bound structure now serves as a holding site for the displaced sulfide S2B, most likely as the hydrosulfide anion (HS^-), that could be unambiguously identified by its anomalous scattering properties (Figure 13B) [123]. Hereby, the anion interacted with the backbone amide groups of the protein chain in a way that was strongly reminiscent of the stabilization of a negative charge in the oxyanion hole of serine proteinases. The bound ligand itself must be an atom with a mass less than that of sulfur, as evidenced be the electron density map and the short bond distances to the adjacent iron ions. Based on the magnitude of the electron density map peak, candidate atoms are nitrogen or oxygen, and the geometry of the surrounding residues, in particular the distance of only 2.55 Å to the amide oxygen of residue Gln176, implies that the ligand must be protonated. Based on structural data, the remaining candidates thus were NH and OH, with obvious consequences for understanding the reactivity of the enzyme in either case.

5.4. Implications for the Mechanism of Biological Nitrogen Fixation

Nitrogenases catalyze the stepwise 6-electron reduction of dinitrogen to ammonia in biological nitrogen fixation. They operate at ambient conditions and pressure, but require the hydrolysis of two molecules of ATP per transferred electron and evolve at least one molecule of H_2 during the reaction. In addition, Lowe and Thorneley found that the enzymes must accumulate four electrons prior to N_2 binding, and that they are likely to produce substantial amounts of additional H_2 as an unwanted side product if this accumulation process fails (see Section 2.3 ff). These individual data points, among others, were accumulated over the course of several decades in different laboratories, and yet only in recent years the available pieces of information begin to fall in place. Two major obstacles for understanding nitrogenase catalysis in the past were the lack of a comprehensive model of the electronic structure of the active site cofactor, and the fact that no clear binding site for substrate molecules could be defined. Both points have been extensively addressed in recent times, integrating into an increasingly detailed mechanism at atomic resolution.

5.4.1. A Substrate Binding Site on the Cofactor

With the structures of the nitrogenase cofactor in its resting state at hand, numerous attempts were undertaken to elucidate the position and mode of substrate binding to a moiety that lacked any obvious coordination site (see Section 4.3). The complete removal of sulfide S2B could not be anticipated in such studies, and while its mechanistic relevance still remains under debate, the correlation of the CO adduct of Mo-nitrogenase and the ligand-bound state of V-nitrogenase are quite suggestive. In the latter, the light atom was interpreted as an NH species, making it a bound reaction intermediate, most likely at the E_6 state of the catalytic cycle [123]. The available ligand-bound structures break with the predominant concept that binding of substrate to nitrogenase cofactors essentially occurs at a single coordination site. Instead, Fe2 and Fe6 form a dinuclear center that allows for the bridging binding of small molecule substrates [126]. This motif is not unprecedented, or even uncommon, in either inorganic complexes or bioinorganic chemistry, but was not considered due to the consistent presence of sulfide S2B in all structures. In particular, the availability of two metal sites facilitates the possibility for a concerted two-electron transfer from reduced centers to a bridging ligand. For a substrate such as N_2, where the transfer of the first electron is the critical and energetically most challenging step [127], this may well be relevant for catalysis.

Nitrogenases are generally isolated with the bridging sulfide S2B in place, indicating that the resting state E_0 of the Lowe-Thorneley cycle represents a stable and thus preferred form of the cluster. Notably, as there are no protocols available as yet for the recombinant production of nitrogenases, all these enzymes were isolated from diazotrophic bacteria grown under nitrogen-fixing conditions. They will therefore already have turned over N_2 and will have gone through the catalytic cycle multiple times. The same holds true for the ligand-bound structure of VFe protein, and indeed preparations in the resting state and the ligand-bound turnover state exhibit the same catalytic activities, showing that the latter does not represent an inhibited off-pathway product [123]. However, it remains entirely unclear how a ligand-bound structure was obtained for vanadium nitrogenase, while attempts to generate a similar adduct of MoFe protein remained unsuccessful so far (unpublished data). Both enzymes exhibit an extraordinary degree of structural similarity in the vicinity of their cofactors including the position of coordinated water molecules, providing no obvious rationale for a difference in function other than the electronic properties of the cofactors themselves. Furthermore, the analysis of the electron density maximum observed at 1.2 Å resolution could not conclusively distinguish between a bound NH and an OH species. While the protonation of the ligand was derived from the short distance of 2.55 Å to the amide oxygen of Gln176, either light atom would satisfy the observed density peak. The assignment as an NH species was based on two facts. First, the enzyme as isolated had turned over N_2 previously and could be returned to the resting state E_0 by addition of the strong reductant dithionite; the enzymatic activity of preparations showing E_6 was unaltered from those showing E_0. Second, a bridging OH^- ligand would be at the oxidation state of H_2O, imply-

ing that either a single protonation was required to release H_2O (see Section 5.4.3), or that H_2O would bind to the FeV cofactor, which in an aqueous medium should lead to an inhibitory effect that was not observed. The question for the nature of the bound ligand will be eventually answered through labeling experiments, but given the unaltered reactivity of the enzyme, even a OH^- ligand would highlight a binding position for a monoatomic, protonated species concomitant with the flip of Gln176 that plays a central role in nitrogenase catalysis.

5.4.2. Hydrogen Evolution and the Mechanism of Nitrogenases

The discovery – and repeated confirmation – of a ligand-binding site on both the FeMo and FeV cofactors represents a crucial step in understanding the activation of highly inert N_2 under mild reaction conditions. It furthermore ties in seamlessly with a prevalent hypothesis that rationalizes one of the most enigmatic postulates of the Lowe-Thorneley cycle. The stable resting state of nitrogenases, E_0, is well understood with regard to its electronic properties (see Section 4.2.3). The molybdenum ion of the FeMo cofactor is in the Mo(III) state, and if localized charges are considered, three Fe ions are Fe(II), while four are Fe(III). With its central, electron-rich carbide, the large cluster is already highly reduced in the resting state, but according to the Lowe-Thorneley scheme it must accumulate four additional electrons before N_2 is bound and activated. Even if the accumulation of charge is balanced by concomitant protonation of the cofactor, it is very difficult to conceive how these four reduction steps can be accommodated with the required electrons arriving sequentially and at constant midpoint potential from Fe protein. The reduction of the cluster should also lead to changes in bond lengths and geometries that were not observed, although it may well be that the formation of such states for spectroscopic characterization had not been successful.

In 2007, Seefeldt, Hoffman, Dean, and associates presented a cryo-annealing protocol to trap intermediate states of the catalytic cycle of Mo-nitrogenase [128]. They observed distinct spectroscopic signatures along the reaction coordinate, and in a subsequent series of papers concluded that these observations are best explained by assuming that electrons are not accumulated in the cluster itself, but are rather used to form metal hydrides at its surface (Figure 14) [129–132]. Accordingly, a proton is required to stabilize two electrons at the cluster surface, whereby theoretical studies favor the generation of hydrides bridging two Fe sites over the less stable terminal hydrides at a single site [133]. A bridging hydride relieves the need for the cluster to internalize the additional charge, but it creates the risk of losing H^- to an undesired protonation event that yields H_2 and leaves the cluster in a two-electron oxidized state, effectively turning back the position in the Lowe-Thorneley cycle. This is exactly what was observed for all nitrogenases: H_2 is formed as an unwanted by-product, in particular under conditions of low electron flux to the cluster [134]. The hydride hypothesis helps to understand that this is indeed a futile side reaction, but rationalizes its existence with the requirement for the accumulation of hydrides during the progression of the enzyme through the catalytic cycle. And indeed, Lowe and Thorneley

have observed the release of H_2 at stage E_2 (returning the enzyme to E_0), but also at E_3 and until E_4, where N_2 can be bound and activated [57]. Hydrogen formation by nitrogenases thus reflects the challenge of accumulating sufficient reduction power for the breaking of N_2, and the danger of losing costly electrons to accidental protonation of bridging hydrides seems to be a price that nature has not achieved to avoid. It underlines the significance of controlling electron flow to the active site, further highlighting the role of the conserved histidine residue near the S2B position of the FeMo and the FeV cofactors.

However, the classic experiment of Simpson and Burris that seemingly addresses this exact process holds a much more essential clue [8]. High pressure of N_2 limits the release of H_2 by nitrogenase, which is in line with a direct competition of both processes. Nevertheless, even at 50 atm of N_2 the release of H_2 did not cease entirely. Instead, the molar ratio of H_2 released/N_2 fixed approached a value of 1.13 ± 0.13. Seefeldt and colleagues understood this remaining hydrogen gas molecule to be of a fundamentally different nature than all others released on the way to the state E_4. With the enzyme reaching E_4, the four electrons provided by the Fe protein upon ATP hydrolysis have generated two bridging hydrides at the cluster surface. Rather than reacting with further protons, these hydrides could then recombine to eliminate H_2, leaving the cluster with the remaining electrons [135, 136]. The effect of this reductive elimination of H_2 is that the nitrogenase cofactor itself now is retained in a highly reactive, two-electron reduced form (Figure 14). The sequential electron transfer from the Fe protein will never reach this state, as the midpoint potential of the delivered electrons is likely sufficient for the one-electron reduction of the site, but not for the transfer of a second electron to the metal core. In this sequential process, the alternative chosen by the enzyme is the formation of surface hydrides, which in turn do not allow the stored electrons to react with N_2. Thus, the effect of the reductive elimination of H_2 is to rapidly produce a super-reduced cofactor and it is this state that holds the reducing power to react with N_2 [126]. The reactivity of this E_4^* state requires a high degree of coordination within the enzyme. For once, the release of H_2 is likely reversible, as indicated by the formation of HD upon the addition of D_2 under turnover conditions [133]. The enzyme can generate two bridging D^- anions from state E_4^*, with HD as product. Also, a simple protonation of the super-reduced state might lead back to a single bridging H^-, taking the enzyme once more to E_2. It seems imperative that the elimination of H_2 from state E_4 is mechanistically coupled to the binding of N_2, possibly in the form of a transient and less stable association to one of the adjacent Fe centers, Fe2 or Fe6.

Prior to the recognition of the reversible dissociation of sulfide S2B, Seefeldt, Hoffman, Dean, and coworkers [124] hypothesized a binding of the two bridging H^- in E_4 on one of the four Fe faces of the FeMo cofactor. Combining both lines of information now gives rise to a chemically reasonable model for hydride formation on the nitrogenase cofactor that is compatible with the available experimental evidence [126]. The first reduction from E_0 to E_1 would lead to an additional electron being stored within the cofactor itself, most likely at one of the most oxidized Fe sites, Fe2 or Fe6. In parallel, the transfer of a proton would

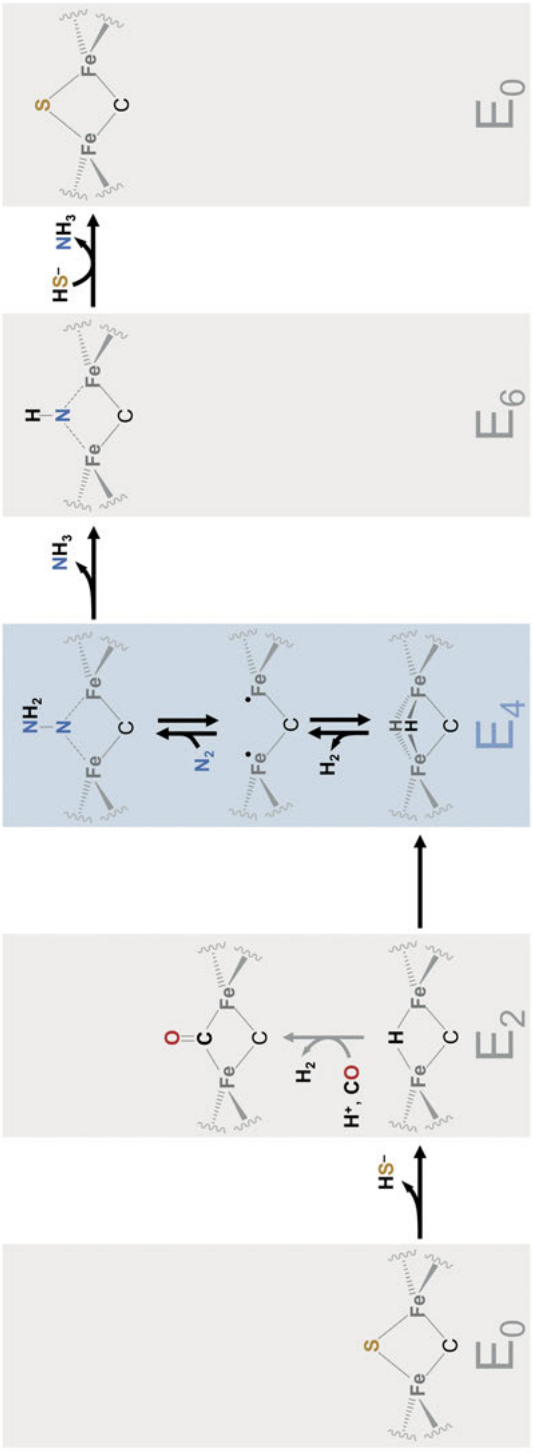


Figure 14. Mechanistic proposal for nitrogenase. The catalytic cofactor continuously oscillates between the paramagnetic ($S = 3/2$) even E-states and the diamagnetic odd E-states (omitted in the scheme). A one-electron transfer leads to reduction of the cluster itself, but every second electron can no longer be retained in the cofactor, leading to the formation of a bridging surface H^- anion in E_2 and a second one in E_4. Unwanted accidental protonation of these hydrides leads to a futile cycle of H_2 release and regression of the enzyme in the catalytic cycle by two steps, as shown in the Lowe-Thorneley scheme (Figure 4). In E_4, however, the two bridging H^- can recombine to release H_2 and leave a two-electron-reduced cofactor able to bind and break N_2. This crucial step of the cycle is the only two-electron reduction involved, after transferring two electrons to N_2, the remaining 4 single-electron steps are uncritical, and the lack of further H_2 evolution indicates that the formation of H^- is no longer required.

balance the charge and is proposed to occur at one of the bridging sulfides. Considering proton transfer to occur from His180, the most likely target for this protonation is S2B itself. In the following round of ATP hydrolysis and electron transfer, the second electron can no longer be stored in the cluster, and together with the next proton it will form a bridging H^- at the cluster surface. Its replacement is under debate, but note that the most oxidized site on the cofactor is still either Fe2 or Fe6, depending on where the first electron goes. According to the electronic coupling scheme by Noodleman and coworkers, Fe2 and Fe6 couple antiferromagnetically [63], so that a H^- bridge forming at Fe2 and Fe6 seems the most likely option [126]. Combined with the finding of a labile S2B, corroborated by three ligand-bound structures, this necessarily raises the question whether the sulfide bridge is already partly or fully released during these early stages of the catalytic cycle. A critical question in this context is the lifetime of the newly generated surface H^-. In order to proceed along the Lowe-Thorneley scheme the enzyme must stabilize this labile intermediate and avoid its accidental protonation and concomitant loss of the two accumulated electrons in the form of H_2. One implication of the ligand-bound structure of the E_6 state of the V-nitrogenase is that the observed swing of the side chain of Gln176 leads to the formation of a short hydrogen bond of 2.86 Å to the $N_{\varepsilon 2}$ nitrogen of His180, the putative proton donor for the reduction reactions taking place at the cofactor [123]. A prerequisite for this, however, is the dissociation of S2B and its relocation into the binding pocket provided by the movement of the side chain of Gln176. Thus, at the point when a blockage of proton access through H-bond formation is required, the dissociation of S2B that drives the swing of Gln176, ties in with the formation of the first bridging H^- precisely at the now open binding site at Fe2 and Fe6. A current working model for the activation of the cofactor in the course of moving from the resting state E_0 to E_4, where the reductive elimination of H_2 can occur, thus predicts the dissociation of S2B upon reaching E_2 [126], when the conformational change of Gln176 hinders proton access and thus helps to extend the lifetime of the H^- adduct. With the next reduction to the E_3 state, Fe2 or Fe6 can accept another electron, but at least through the entry point of His180, proton access remains restricted. In the E_3 state, N_2 can already associate with the cofactor, but its binding is reversible and does not yet lead the enzyme to its reductive conversion [30]. Reaching E_4 upon further reduction, a second bridging H^- is formed, requiring an additional proton. We suggested that this hydride might bind between Fe2 and Fe6 as well, forming a *bis*-μ^2-dihydride that allows for two protons and two electrons to recombine and eliminate H_2, leaving both Fe2 and Fe6 in the super-reduced state that could not be reached through single-electron transfer from the Fe protein [126]. As seen from the HD exchange experiments, this elimination is reversible, but nevertheless the two-electron reduced cofactor will be a highly reactive entity, and the enzyme likely couples the provision of the substrate N_2 to the release of H_2, aiming to suppress unwanted side reactions. Interestingly, a recent study by Peters and coworkers reported a new ligand, $[SiP_2O]H_2$, that was able to stabilize a dinuclear Fe site with one or two μ-bridging hydrides in an arrangement reminiscent of the proposed structures for E_2 and E_4 [137]. This ligand facilitates the end-on coordina-

tion of one or two molecules of N_2 to the Fe sites, and although no reductive conversion was observed in the model complexes, these results could reflect the events at the nitrogenase cofactor prior to the reductive elimination of H_2.

5.4.3. Distal or Alternating: Substrate Reductions at the Cofactor

As a dinuclear active site, the S2B-depleted Fe2-Fe6 edge of nitrogenase cofactors seems well set to bind substrates, including even inert molecules such as N_2. The two electrons remaining on the cofactor after H_2 elimination are expected to reside on different Fe ions, and consequently a bridging binding mode of a bound ligand will be favored over a terminal binding. It is not obvious, how this type of enzyme-substrate complex is formed. Diatomic molecules can bind in a μ^2-bridging fashion, as in the case of CO, but a $\mu{:}\eta^2\text{-}\eta^2$ coordination is conceivable as well. For CO, the negative partial charge on the carbon atom promotes the observed binding mode, but in the case of the isoelectronic N_2 no such partial charge exists. The structure of an E_6 state of the Lowe-Thorneley cycle obtained for the VFe protein depicts the situation after cleavage of the N≡N bond and consequently does not reveal the initial N_2 binding mode. This ambiguity is also reflected in the current mechanistic models for nitrogenase catalysis [138, 139]. Although these do not yet consider a diatomic binding site, they propose two lines along which the multi-step reduction reaction may occur. In the *distal mechanism*, the electron delivered from the enzyme to the substrate will accumulate on the distal nitrogen atom, leading to the cleavage of the N≡N bond of the substrate and dissociation of the first molecule of product NH_4^+ after three electrons are obtained. As two electrons are transferred upon binding to E_4, this would occur already in E_5, leaving behind a formal nitrido (N^{3-}) species that must be stabilized by the enzyme and is successively reduced to product in the following three electron transfer steps that close the catalytic cycle. In the *alternating mechanism*, electrons and protons will be transferred to the two nitrogen atoms alternatingly, leading through the stages of diazene (HN=NH) at E_4, hydrazide ($H_2N\text{-}NH^-$) at E_5 and hydrazine ($H_2N\text{-}NH_2$) at E_6, so that the N–N bond is only cleaved in the penultimate reduction step, upon reaching E_7. A point in favor for the latter mechanism is that all these intermediates are alternative substrates of the enzyme [140]. The observation of the cofactor to bind a monoatomic, singly protonated species seems to favor the *distal mechanism* in this scheme [123], but at least for model complexes it was already demonstrated that a reaction mechanism it not required to strictly follow either route, but can instead proceed along a hybrid pathway [141]. With the new data on a dinuclear Fe2-Fe6 edge as the site of substrate binding, the two different pathways of reduction are well reflected in different binding modes for a diatomic ligand: (i) μ^2-binding, as observed for CO, with H180 as a proton source, would favor the *distal mechanism*, while (ii) a more symmetric $\mu{:}\eta^2\text{–}\eta^2$ coordination of N_2 would be in line with the *alternating mechanism*. The available data does not yet allow to draw a final conclusion, but the question highlights how in recent

years, nitrogenase research starts to increasingly integrate theoretical, spectroscopic, biochemical and structural data, moving towards a comprehensive understanding of the unique catalytic capabilities of the enzymatic machinery that underlies the essential process of biological nitrogen fixation.

6. FUTURE DIRECTIONS

Catalyzing what arguably represents the most challenging multi-electron redox reaction known in nature, nitrogenase is an enzyme system of striking complexity on multiple levels. Its three known variants, Mo-, V-, and Fe-nitrogenase, differ in some highly interesting properties, but are generally considered to follow the very same mechanistic route for N_2 activation and reduction. The mechanism outlined by Lowe and Thorneley thus appears to reflect general principles of this reaction and remains the gold standard for a kinetic understanding of the system. Beyond kinetics, however, the expectation for a complete mechanistic understanding of an enzyme also includes a detailed structural description of each reaction intermediate, defined here as the eight E-states of the Lowe-Thorneley cycle. This refers to the atomic as well as to the electronic structure of at least the cofactors of nitrogenases, and as we discussed above, the trapping of individual intermediates was partly achieved for spectroscopic investigations, but remains highly challenging for direct structural analyses that rely largely on the slow process of protein crystallization and will only be sufficiently informative if true atomic resolution can be achieved.

In the past, the broad range of theoretical approaches towards understanding nitrogenase catalysis has led to several proposals that were, not too surprisingly, partly contradictory in their explanations and predictions. With a substantially increased basis of experimental data and the constant development of methods and increase of computing power, we can expect these various approaches to eventually converge onto a generally accepted model for the cofactor that will then advance the analytic and predictive power that is indispensable to complement experimental work on the way to completing and refining a model of the catalytic reaction. Among the remaining open questions is not only the one for the nature of all intermediates and the binding mode of substrates (see Section 5.4.2), but also the coupling of electron and proton transfer and the sites of protonation in the individual steps of the reaction. The postulate that S2B is released at E_2 remains to be confirmed, and it will be relevant to elucidate whether reduction of the cluster and the suggested protonation of S2B in E_1 already breaks one of the ligand's bonds to a cluster iron. For CO reduction by V-nitrogenase, the steps of the reduction leading to the point of C–C bond formation remain largely obscure, as this in particular almost certainly requires the binding of a second substrate molecule to the cluster, which in turn does not occur during N_2 reduction. For the reduction itself, the mechanism of electron transfer from Fe protein to the enzymatic component has been intensively studied in particular by Seefeldt and coworkers in recent years [117, 142–144], but important details regarding the coupling of ATP hydrolysis and the triggering

of an electron being ejected from the P-cluster at low potential justify further scrutiny.

Finally, a vast area hardly touched in the present overview, concerns the assembly and biogenesis of the components of the nitrogenases, their accessory factors and their corresponding electron donor systems. Decade-long efforts by multiple teams have provided an excellent and generally accepted outline of the process, and we can hope and expect that the results of the coming years will allow to fill in the gaps and understand metal cluster assembly, maturation, transport, and insertion into the apo-enzyme to generate the unique machineries of biological nitrogen fixation.

ACKNOWLEDGMENTS

The authors thank Doug Rees, Jim Howard, Thomas Spatzal, Daniel Sippel, Laure Decamps, Serena DeBeer, Ragnar Björnsson, Pat Holland, Holger Dobbek, Franc Meyer, Jonas Peters, Brian Hoffman, Dennis Dean, and Patricia dos Santos for insights and stimulating discussions. This work was supported by the Deutsche Forschungsgemeinschaft (RTG 1976 and PP 1927 to O.E.) and the European Research Council (grant no. 310656 to O.E.).

ABBREVIATIONS AND DEFINITIONS

AMPPCP	adenosine 5′-[β,γ-methylene]triphosphate
ADP	adenosine 5′-diphosphate
ATP	adenosine 5′-triphosphate
BS-DFT	broken symmetry density functional theory
COR	chlorophyllide oxidoreductase
DFT	density functional theory
DMSO	dimethylsulfoxide
DPOR	dark operative protochlorophyllide reductase
EPR	electronic paramagnetic resonance
ENDOR	electron nuclear double resonance
ESEEM	electron spin echo envelope modulation
flx	flavodoxin
HERFD-XAS	high energy resolution fluorescence-detected X-ray absorption spectroscopy
NDP	nucleoside diphosphate
NTP	nucleoside triphosphate
P_i	inorganic phosphate (PO_4^{3-})
SAM	*S*-adenosylmethionine
SpReAD	spatially resolved anomalous dispersion
XAS	X-ray absorption spectroscopy

REFERENCES

1. J. Chatt, G. J. Leigh, *Chem. Soc. Rev.* **1972**, *1*, 121–136.
2. V. Smil, *Global Biogeochem. Cycles* **1999**, *13*, 647–662.
3. S. J. Ferguson, *Curr. Opin. Chem. Biol.* **1998**, *2*, 182–193.
4. D. E. Canfield, A. N. Glazer, P. G. Falkowski, *Science* **2010**, *330*, 192–196.
5. V. Smil, *Popul. Dev. Rev.* **1991**, *17*, 569–601.
6. J. B. Howard, D. C. Rees, *Chem. Rev.* **1996**, *96*, 2965–2982.
7. G. Ertl, *Z. Anorg. Allg. Chem.* **2012**, *638*, 487–489.
8. F. B. Simpson, R. H. Burris, *Science* **1984**, *224*, 1095–1097.
9. J. R. Gallon, L. J. Stal, *Nato Adv. Sci. I C-Mat* **1992**, *362*, 115–139.
10. S. Schubert, *Fert. Res.* **1995**, *42*, 99–107.
11. R. D. Joerger, P. E. Bishop, *CRC Crit. Rev. Microbiol.* **1988**, *16*, 1–14.
12. R. R. Eady, *Chem. Rev.* **1996**, *96*, 3013–3030.
13. R. V. Hageman, W. H. Orme-Johnson, R. H. Burris, *Biochemistry* **1980**, *19*, 2333–2342.
14. M. M. Georgiadis, H. Komiya, P. Chakrabarti, D. Woo, J. J. Kornuc, D. C. Rees, *Science* **1992**, *257*, 1653–1659.
15. C. P. Owens, F. E. Katz, C. H. Carter, M. A. Luca, F. A. Tezcan, *J. Am. Chem. Soc.* **2015**, *137*, 12704–12712.
16. F. A. Tezcan, J. T. Kaiser, D. Mustafi, M. Y. Walton, J. B. Howard, D. C. Rees, *Science* **2005**, *309*, 1377–1380.
17. Z. Y. Yang, R. Ledbetter, S. Shaw, N. K. Pence, M. Tokmina-Lukaszewska, B. J. Eilers, Q. Guo, N. Pokhrel, V. L. Cash, D. R. Dean, E. Antony, B. Bothner, J. W. Peters, L. C. Seefeldt, *Biochemistry* **2016**, *55*, 3625–3635.
18. H. Schindelin, C. Kisker, J. L. Sehlessman, J. B. Howard, D. C. Rees, *Nature* **1997**, *387*, 370–376.
19. J. B. Howard, K. J. Kechris, D. C. Rees, A. N. Glazer, *Plos One* **2013**, *8*, e72751.
20. D. C. Rees, J. B. Howard, *Curr. Opin. Chem. Biol.* **2000**, *4*, 559–566.
21. G. Layer, J. Krausze, J. Moser, *Adv. Exp. Med. Biol.* **2017**, *925*, 147–161.
22. K. Y. Zheng, P. D. Ngo, V. L. Owens, X. P. Yang, S. O. Mansoorabadi, *Science* **2016**, *354*, 339–342.
23. S. J. Moore, S. T. Sowa, C. Schuchardt, E. Deery, A. D. Lawrence, J. V. Ramos, S. Billig, C. Birkemeyer, P. T. Chivers, M. J. Howard, S. E. J. Rigby, G. Layer, M. J. Warren, *Nature* **2017**, *543*, 78–82.
24. E. C. Raulfs, I. P. O'Carroll, P. C. Dos Santos, M. C. Unciuleac, D. R. Dean, *Proc. Natl. Acad. Sci. USA* **2008**, *105*, 8591–8596.
25. P. C. Dos Santos, D. R. Dean, Y. L. Hu, M. W. Ribbe, *Chem. Rev.* **2004**, *104*, 1159–1173.
26. Y. Hu, M. W. Ribbe, *Ann. Rev.Biochem.* **2016**, *85*, 455–483.
27. R. V. Hageman, R. H. Burris, *Proc. Natl. Acad. Sci. USA* **1978**, *75*, 2699–2702.
28. D. J. Lowe, R. N. F. Thorneley, *Biochem. J.* **1984**, *224*, 877–886.
29. D. J. Lowe, R. N. F. Thorneley, *Biochem. J.* **1984**, *224*, 895–901.
30. R. N. F. Thorneley, D. J. Lowe, *Biochem. J.* **1984**, *224*, 887–894.
31. R. N. F. Thorneley, D. J. Lowe, *Biochem. J.* **1984**, *224*, 903–909.
32. B. K. Burgess, S. Wherland, W. E. Newton, E. I. Stiefel, *Biochemistry* **1981**, *20*, 5140–5146.
33. J. H. Guth, R. H. Burris, *Biochemistry* **1983**, *22*, 5111–5122.
34. D. F. Harris, D. A. Lukoyanov, S. Shaw, P. Compton, M. Tokmina-Lukaszewska, B. Bothner, N. Kelleher, D. R. Dean, B. M. Hoffman, L. C. Seefeldt, *Biochemistry* **2018**, *57*, 701–710.
35. J. S. Kim, D. C. Rees, *Nature* **1992**, *360*, 553–560.

36. J. S. Kim, D. C. Rees, *Science* **1992**, *257*, 1677–1682.
37. S. M. Mayer, D. M. Lawson, C. A. Gormal, S. M. Roe, B. E. Smith, *J. Mol. Biol.* **1999**, *292*, 871–891.
38. C. P. Owens, F. E. H. Katz, C. H. Carter, V. F. Oswald, F. A. Tezcan, *J. Am. Chem. Soc.* **2016**, *138*, 10124–10127.
39. L. M. Zhang, C. N. Morrison, J. T. Kaiser, D. C. Rees, *Acta Crystallogr. D* **2015**, *71*, 274–282.
40. H. Beinert, R. H. Holm, E. Münck, *Science* **1997**, *277*, 653–659.
41. G. D. Watt, K. R. N. Reddy, *J. Inorg. Biochem.* **1994**, *53*, 281–294.
42. H. C. Angove, S. J. Yoo, E. Munck, B. K. Burgess, *J. Biol. Chem.* **1998**, *273*, 26330–26337.
43. P. Strop, P. M. Takahara, H. J. Chiu, H. C. Angove, B. K. Burgess, D. C. Rees, *Biochemistry* **2001**, *40*, 651–656.
44. J. W. Peters, M. H. B. Stowell, S. M. Soltis, M. G. Finnegan, M. K. Johnson, D. C. Rees, *Biochemistry* **1997**, *36*, 1181–1187.
45. S. M. Keable, O. A. Zadvornyy, L. E. Johnson, B. Ginovska, A. J. Rasmussen, K. Danyal, B. J. Eilers, G. A. Prussia, A. X. LeVan, S. Raugei, L. C. Seefeldt, J. W. Peters, *J. Biol. Chem.* **2018**, *293*, 9629–9635.
46. K. Danyal, D. R. Dean, B. M. Hoffman, L. C. Seefeldt, *Biochemistry* **2011**, *50*, 9255–9263.
47. S. Duval, K. Danyal, S. Shaw, A. K. Lytle, D. R. Dean, B. M. Hoffman, E. Antony, L. C. Seefeldt, *Proc. Natl. Acad. Sci. USA* **2013**, *110*, 16414–16419.
48. F. A. Tezcan, J. T. Kaiser, J. B. Howard, D. C. Rees, *J. Am. Chem. Soc.* **2015**, *137*, 146–149.
49. K. L. C. Gronberg, C. A. Gormal, M. C. Durrant, B. E. Smith, R. A. Henderson, *J. Am. Chem. Soc.* **1998**, *120*, 10613–10621.
50. C. Kennedy, D. Dean, *Mol. Gen. Genet.* **1992**, *231*, 494–498.
51. V. K. Shah, W. J. Brill, *Proc. Natl. Acad. Sci. USA* **1977**, *74*, 3249–3253.
52. R. H. Holm, *Chem. Soc. Rev.* **1981**, *10*, 455–490.
53. D. Coucouvanis, *Acc. Chem. Res.* **1981**, *14*, 201–209.
54. B. E. Smith, R. R. Eady, *Eur. J. Biochem.* **1992**, *205*, 1–15.
55. D. M. Kurtz, R. S. Mcmillan, B. K. Burgess, L. E. Mortenson, R. H. Holm, *Proc. Natl. Acad. Sci. USA* **1979**, *76*, 4986–4989.
56. M. K. Chan, J. S. Kim, D. C. Rees, *Science* **1993**, *260*, 792–794.
57. R. N. F. Thorneley, D. J. Lowe, in *Molybdenum Enzymes*, Ed. T. G. Spiro, Vol. 1, Wiley-Interscience, New York, 1985, pp. 221–284.
58. O. Einsle, F. A. Tezcan, S. L. A. Andrade, B. Schmid, M. Yoshida, J. B. Howard, D. C. Rees, *Science* **2002**, *297*, 1696–1700.
59. O. Einsle, *J. Biol. Inorg. Chem.* **2014**, *19*, 737–745.
60. H. I. Lee, P. M. C. Benton, M. Laryukhin, R. Y. Igarashi, D. R. Dean, L. C. Seefeldt, B. M. Hoffman, *J. Am. Chem. Soc.* **2003**, *125*, 5604–5605.
61. T. C. Yang, N. K. Maeser, M. Laryukhin, H. I. Lee, D. R. Dean, L. C. Seefeldt, B. M. Hoffman, *J. Am. Chem. Soc.* **2005**, *127*, 12804–12805.
62. T. Lovell, J. Li, D. A. Case, L. Noodleman, *J. Biol. Inorg. Chem.* **2002**, *7*, 735–749.
63. T. Lovell, J. Li, T. Q. Liu, D. A. Case, L. Noodleman, *J. Am. Chem. Soc.* **2001**, *123*, 12392–12410.
64. T. Lovell, T. Q. Liu, D. A. Case, L. Noodleman, *J. Am. Chem. Soc.* **2003**, *125*, 8377–8383.
65. D. Lukoyanov, V. Pelmenschikov, N. Maeser, M. Laryukhin, T. C. Yang, L. Noodleman, D. R. Dean, D. A. Case, L. C. Seefeldt, B. M. Hoffman, *Inorg. Chem.* **2007**, *46*, 11437–11449.

66. K. M. Lancaster, M. Roemelt, P. Ettenhuber, Y. L. Hu, M. W. Ribbe, F. Neese, U. Bergmann, S. DeBeer, *Science* **2011**, *334*, 974–977.
67. T. Spatzal, M. Aksoyoğlu, L. M. Zhang, S. L. A. Andrade, E. Schleicher, S. Weber, D. C. Rees, O. Einsle, *Science* **2011**, *334*, 940.
68. J. A. Wiig, Y. L. Hu, C. C. Lee, M. W. Ribbe, *Science* **2012**, *337*, 1672–1675.
69. B. Hinnemann, J. K. Norskov, *J. Am. Chem. Soc.* **2003**, *125*, 1466–1467.
70. J. B. Varley, Y. Wang, K. Chan, F. Studt, J. K. Norskov, *Phys. Chem. Chem. Phys.* **2015**, *17*, 29541–29547.
71. J. Kästner, P. E. Blöchl, *ChemPhysChem* **2005**, *6*, 1724–1726.
72. P. E. M. Siegbahn, *J. Am. Chem. Soc.* **2016**, *138*, 10485–10495.
73. H. I. Lee, B. J. Hales, B. M. Hoffman, *J. Am. Chem. Soc.* **1997**, *119*, 11395–11400.
74. S. J. Yoo, H. C. Angove, V. Papaefthymiou, B. K. Burgess, E. Münck, *J. Am. Chem. Soc.* **2000**, *122*, 4926–4936.
75. T. Spatzal, O. Einsle, S. L. Andrade, *Angew. Chem.* **2013**, *52*, 10116–10119.
76. D. V. Fomitchev, C. C. McLauchlan, R. H. Holm, *Inorg. Chem.* **2002**, *41*, 958–966.
77. R. Björnsson, F. A. Lima, T. Spatzal, T. Weyhermüller, P. Glatzel, E. Bill, O. Einsle, F. Neese, S. DeBeer, *Chem. Sci.* **2014**, *5*, 3096–3103.
78. G. Friedel, *Comptes Rendus de l'Academie des Sciences* **1913**, *157*, 1533–1536.
79. O. Einsle, S. L. Andrade, H. Dobbek, J. Meyer, D. C. Rees, *J. Am. Chem. Soc.* **2007**, *129*, 2210–2211.
80. T. Spatzal, J. Schlesier, E. M. Burger, D. Sippel, L. M. Zhang, S. L. A. Andrade, D. C. Rees, O. Einsle, *Nat. Commun.* **2016**, *7*, 10902.
81. R. Björnsson, F. Neese, S. DeBeer, *Inorg. Chem.* **2017**, *56*, 1470–1477.
82. R. Björnsson, F. Neese, R. R. Schrock, O. Einsle, S. DeBeer, *J. Biol. Inorg. Chem.* **2015**, *20*, 447–460.
83. Y. G. Zhang, J. L. Zuo, H. C. Zhou, R. H. Holm, *J. Am. Chem. Soc.* **2002**, *124*, 14292–14293.
84. K. S. Hagen, A. D. Watson, R. H. Holm, *J. Am. Chem. Soc.* **1983**, *105*, 3905–3913.
85. Y. Ohki, Y. Ikagawa, K. Tatsumi, *J. Am. Chem. Soc.* **2007**, *129*, 10457–10465.
86. Y. G. Zhang, R. H. Holm, *J. Am. Chem. Soc.* **2003**, *125*, 3910–3920.
87. J. Rittle, J. C. Peters, *Proc. Natl. Acad. Sci. USA* **2013**, *110*, 15898–15903.
88. J. S. Anderson, J. Rittle, J. C. Peters, *Nature* **2013**, *501*, 84–88.
89. A. Takaoka, N. P. Mankad, J. C. Peters, *J. Am. Chem. Soc.* **2011**, *133*, 8440–8443.
90. K. D. Demadis, S. M. Malinak, D. Coucouvanis, *Inorg. Chem.* **1996**, *35*, 4038–4046.
91. D. V. Yandulov, R. R. Schrock, *Science* **2003**, *301*, 76–78.
92. K. Arashiba, Y. Miyake, Y. Nishibayashi, *Nat. Chem.* **2011**, *3*, 120–125.
93. W. W. Weare, X. Dai, M. J. Byrnes, J. M. Chin, R. R. Schrock, P. Muller, *Proc. Natl. Acad. Sci. USA* **2006**, *103*, 17099–17106.
94. R. R. Schrock, *Philos. Trans. Royal Society a-Math. Phys. Eng. Sci.***2005**, *363*, 959–969.
95. J. C. Hwang, C. H. Chen, R. H. Burris, *Biochim. Biophys. Acta* **1973**, *292*, 256–270.
96. L. M. Cameron, B. J. Hales, *Biochemistry* **1998**, *37*, 9449–9456.
97. L. F. Yan, C. H. Dapper, S. J. George, H. X. Wang, D. Mitra, W. B. Dong, W. E. Newton, S. P. Cramer, *Eur. J. Inorg. Chem.* **2011**, 2064–2074.
98. L. F. Yan, V. Pelmenschikov, C. H. Dapper, A. D. Scott, W. E. Newton, S. P. Cramer, *Chem. Eur. J.* **2012**, *18*, 16349–16357.
99. T. Spatzal, K. A. Perez, O. Einsle, J. B. Howard, D. C. Rees, *Science* **2014**, *345*, 1620–1623.
100. I. Coric, B. Q. Mercado, E. Bill, D. J. Vinyard, P. L. Holland, *Nature* **2015**, *526*, 96–99.
101. I. Coric, P. L. Holland, *J. Am. Chem. Soc.* **2016**, *138*, 7200–7211.
102. T. Spatzal, K. A. Perez, J. B. Howard, D. C. Rees, *Elife* **2015**, *4*, e11620.

103. D. C. Rees, *Ann. Rev. Biochem.* **2002**, *71*, 221–246.
104. C. C. Lee, Y. L. Hu, M. W. Ribbe, *Science* **2010**, *329*, 642–642.
105. C. C. Lee, K. Tanifuji, M. Newcomb, J. Liedtke, Y. L. Hu, M. W. Ribbe, *ChemBioChem* **2018**, *19*, 649–653.
106. C. C. Lee, J. Wilcoxen, C. J. Hiller, R. D. Britt, Y. L. Hu, *Angew. Chem. Int. Ed.* **2018**, *57*, 3411–3414.
107. C. C. Lee, Y. L. Hu, M. W. Ribbe, *Angew. Chem. Int. Ed.* **2012**, *51*, 1947–1949.
108. Y. L. Hu, C. C. Lee, M. W. Ribbe, *Science* **2011**, *333*, 753–755.
109. J. C. Setubal, P. dos Santos, B. S. Goldman, H. Ertesvag, G. Espin, L. M. Rubio, S. Valla, N. F. Almeida, D. Balasubramanian, L. Cromes, L. Curatti, Z. J. Du, E. Godsy, B. Goodner, K. Hellner-Burris, J. A. Hernandez, K. Houmiel, J. Imperial, C. Kennedy, T. J. Larson, P. Latreille, L. S. Ligon, J. Lu, M. Maerk, N. M. Miller, S. Norton, I. P. O'Carroll, I. Paulsen, E. C. Raulfs, R. Roemer, J. Rosser, D. Segura, S. Slater, S. L. Stricklin, D. J. Studholme, J. Sun, C. J. Viana, E. Wallin, B. M. Wang, C. Wheeler, H. J. Zhu, D. R. Dean, R. Dixon, D. Wood, *J. Bacteriol.* **2009**, *191*, 4534–4545.
110. D. Sippel, J. Schlesier, M. Rohde, C. Trncik, L. Decamps, I. Djurdjevic, T. Spatzal, S. L. A. Andrade, O. Einsle, *J. Biol. Inorg. Chem.* **2017**, *22*, 161–168.
111. Y. N. Zheng, D. F. Harris, Z. Yu, Y. F. Fu, S. Poudel, R. N. Ledbetter, K. R. Fixen, Z. Y. Yang, E. S. Boyd, M. E. Lidstrom, L. C. Seefeldt, C. S. Harwood, *Nature Microbiol.* **2018**, *3*, 281–286.
112. M. W. Ribbe, Y. L. Hu, K. O. Hodgson, B. Hedman, *Chem. Rev.* **2014**, *114*, 4063–4080.
113. M. Rohde, C. Trncik, D. Sippel, S. Gerhardt, O. Einsle, *J. Biol. Inorg. Chem.* **2018**, *23*, 1049–1056.
114. D. Sippel, O. Einsle, *Nat. Chem. Biol.* **2017**, *13*, 956–960.
115. L. M. Zhang, J. T. Kaiser, G. Meloni, K. Y. Yang, T. Spatzal, S. L. A. Andrade, O. Einsle, J. B. Howard, D. C. Rees, *Angew. Chem. Int. Ed.* **2013**, *52*, 10529–10532.
116. T. A. Clarke, S. Fairhurst, D. J. Lowe, N. J. Watmough, R. R. Eady, *Biochem. Soc. Trans.* **2011**, *39*, 201–206.
117. K. Danyal, S. Shaw, T. R. Page, S. Duval, M. Horitani, A. R. Marts, D. Lukoyanov, D. R. Dean, S. Raugei, B. M. Hoffman, L. C. Seefeldt, E. Antony, *Proc. Natl. Acad. Sci. USA* **2016**, *113*, E5783–E5791.
118. J. T. Kaiser, Y. L. Hu, J. A. Wiig, D. C. Rees, M. W. Ribbe, *Science* **2011**, *331*, 91–94.
119. J. A. Rees, R. Bjornsson, J. Schlesier, D. Sippel, O. Einsle, S. DeBeer, *Angew. Chem. Int. Ed.* **2015**, *54*, 13249–13252.
120. J. A. Rees, R. Björnsson, J. K. Kowalska, F. A. Lima, J. Schlesier, D. Sippel, T. Weyhermüller, O. Einsle, J. A. Kovacs, S. DeBeer, *Dalton T* **2017**, *46*, 2445–2455.
121. K. Fisher, M. J. Dilworth, C. H. Kim, W. E. Newton, *Biochemistry* **2000**, *39*, 2970–2979.
122. K. Fisher, M. J. Dilworth, W. E. Newton, *Biochemistry* **2000**, *39*, 15570–15577.
123. D. Sippel, M. Rohde, J. Netzer, C. Trncik, J. Gies, K. Grunau, I. Djurdjevic, L. Decamps, S. L. A. Andrade, O. Einsle, *Science* **2018**, *359*, 1484–1489.
124. I. Dance, *Dalton Trans.* **2012**, *41*, 7647–7659.
125. E. Münck, H. Rhodes, W. H. Orme-Johnson, L. C. Davis, W. J. Brill, V. K. Shah, *Biochim. Biophys. Acta* **1975**, *400*, 32–53.
126. M. Rohde, D. Sippel, C. Trncik, S. L. A. Andrade, O. Einsle, *Biochemistry* **2018**, *57*, 5497–5504.
127. F. Studt, F. Tuczek, *J. Comput. Chem.* **2006**, *27*, 1278–1291.
128. D. Lukoyanov, B. M. Barney, D. R. Dean, L. C. Seefeldt, B. M. Hoffman, *Proc. Natl. Acad. Sci. USA* **2007**, *104*, 1451–1455.

129. D. Lukoyanov, N. Khadka, Z. Y. Yang, D. R. Dean, L. C. Seefeldt, B. M. Hoffman, *J. Am. Chem. Soc.* **2016**, *138*, 10674–10683.
130. D. Lukoyanov, Z. Y. Yang, B. M. Barney, D. R. Dean, L. C. Seefeldt, B. M. Hoffman, *Proc. Natl. Acad. Sci. USA* **2012**, *109*, 5583–5587.
131. D. Lukoyanov, Z. Y. Yang, N. Khadka, D. R. Dean, L. C. Seefeldt, B. M. Hoffman, *J. Am. Chem. Soc.* **2015**, *137*, 3610–3615.
132. Z. Y. Yang, N. Khadka, D. Lukoyanov, B. M. Hoffman, D. R. Dean, L. C. Seefeldt, *Proc. Natl. Acad. Sci. USA* **2013**, *110*, 16327–16332.
133. N. Khadka, R. D. Milton, S. Shaw, D. Lukoyanov, D. R. Dean, S. D. Minteer, S. Raugei, B. M. Hoffrnan, L. C. Seefeldt, *J. Am. Chem. Soc.* **2017**, *139*, 13518–13524.
134. R. Broach, M. Sorlie, B. Hales, *J. Inorg. Biochem.* **1999**, *74*, 83–83.
135. L. C. Seefeldt, D. R. Dean, B. M. Hoffman, *Biochemistry* **2017**, *5*, 274–296.
136. L. C. Seefeldt, B. M. Hoffman, D. R. Dean, *Annu. Rev. Biochem.* **2009**, *78*, 701–722.
137. J. Rittle, C. C. L. McCrory, J. C. Peters, *J. Am. Chem. Soc.* **2014**, *136*, 13853–13862.
138. B. M. Hoffman, D. Lukoyanov, D. R. Dean, L. C. Seefeldt, *Acc. Chem. Res.* **2013**, *46*, 587–595.
139. I. Djurdjevic, O. Einsle, L. Decamps, *Chemistry – An Asian Journal* **2017**, *12*, 1447–1455.
140. B. M. Barney, J. McClead, D. Lukoyanov, M. Laryukhin, T. C. Yang, D. R. Dean, B. M. Hoffman, L. C. Seefeldt, *Biochemistry* **2007**, *46*, 6784–6794.
141. J. Rittle, J. C. Peters, *J. Am. Chem. Soc.* **2016**, *138*, 4243–4248.
142. R. Davydov, N. Khadka, Z. Y. Yang, A. J. Fielding, D. Lukoyanov, D. R. Dean, L. C. Seefeldt, B. M. Hoffman, *Isr. J. Chem.* **2016**, *56*, 841–851.
143. L. C. Seefeldt, B. M. Hoffman, J. W. Peters, S. Raugei, D. N. Beratan, E. Antony, D. R. Dean, *Acc. Chem. Res.* **2018**, *51*, 2179–2186.
144. L. C. Seefeldt, J. W. Peters, D. N. Beratan, B. Bothner, S. D. Minteer, S. Raugei, B. M. Hoffman, *Curr. Opin. Chem. Biol.* **2018**, *47*, 54–59.

9

Molybdenum and Tungsten Cofactors and the Reactions They Catalyze

Khadanand KC and Martin L. Kirk*

Department of Chemistry and Chemical Biology, The University of New Mexico, MSC03 2060, 1 University of New Mexico, Albuquerque, NM 87131-0001, USA
<knkc@unm.edu>
<mkirk@unm.edu>

Metal Ions in Life Sciences, Volume 20 Guest Editors: Martha E. Sosa Torres and Peter M. H. Kroneck
Series Editors: Astrid Sigel, Eva Freisinger, and Roland K. O. Sigel

https://doi.org/10.1515/9783110589757-009

Abstract: The last 20 years have seen a dramatic increase in our mechanistic understanding of the reactions catalyzed by pyranopterin Mo and W enzymes. These enzymes possess a unique cofactor (Moco) that contains a novel ligand in bioinorganic chemistry, the pyranopterin ene-1,2-dithiolate. A synopsis of Moco biosynthesis and structure is presented, along with our current understanding of the role Moco plays in enzymatic catalysis. Oxygen atom transfer (OAT) reactivity is discussed in terms of breaking strong metal-oxo bonds and the mechanism of OAT catalyzed by enzymes of the sulfite oxidase (SO) family that possess dioxo Mo(VI) active sites. OAT reactivity is also discussed in members of the dimethyl sulfoxide (DMSO) reductase family, which possess *des*-oxo Mo(IV) sites. Finally, we reveal what is known about hydride transfer reactivity in xanthine oxidase (XO) family enzymes and the formate dehydrogenases. The formal hydride transfer reactivity catalyzed by xanthine oxidase family enzymes is complex and cleaves substrate C–H bonds using a mechanism that is distinct from monooxygenases. The chapter primarily highlights developments in the field that have occurred since ~2000, which have contributed to our collective structural and mechanistic understanding of the three canonical pyranopterin Mo enzymes families: XO, SO, and DMSO reductase.

Keywords: dimethyl sulfoxide reductase · hydride transfer · hydroxylation · molybdenum cofactor · molybdenum enzymes · oxygen atom transfer · pyranopterin dithiolene · sulfite oxidase · xanthine oxidase

1. INTRODUCTION

1.1. Molybdenum and Tungsten Enzyme Families and the Reactions They Catalyze

The canonical pyranopterin Mo enzymes include sulfite oxidase (SO), xanthine oxidase (XO), and dimethyl sulfoxide reductase (DMSO reductase), and the

names of these three enzymes are used to represent the three common enzyme families. These three families are broadly distinguished from one another by the nature of the reactions they catalyze, their protein fold, and the structure of their active sites [1–10]. The pyranopterin tungsten enzymes most closely resemble those of the DMSO reductase family and, in some cases, the two metals can be interchanged in the active sites [11]. Typically, SO and DMSO reductase family enzymes catalyze oxygen atom transfer (OAT) reactions, where the oxygen atom being transferred either derives from, or is converted into, a water molecule. XO family enzymes catalyze a more complicated hydroxylation reaction [3, 5]. This unique type of hydroxylation reaction generates, rather than consumes, reducing equivalents and activates water instead of dioxygen as the source of the oxygen atom formally inserted into the substrate C–H bond [12]. Different types of transformations are catalyzed by some of the non-canonical members of the three enzyme families [2, 5]. Although two-electron chemistry dominates in the catalytic cycles of these enzymes, one-electron chemistry has been observed [13–15] in the generation of NO from nitrite. The active site structures and the nature of the ligands bound to the metal center appear to be fine-tuned so that the reactions catalyzed by pyranopterin Mo enzymes are close to thermoneutral [16]. The pyranopterin tungsten enzymes catalyze both redox [10, 17–21] and non-redox reactions [22]. With respect to their relationship to the Mo enzymes, the W formate dehydrogenases are most similar to enzymes of the Mo DMSO reductase family, while the W aldehyde oxidoreductases are more distinct [17].

1.2. The Global Importance of Molybdenum and Tungsten Enzymes

Molybdenum and tungsten are the only second and third row transition metal ions that are utilized in biological systems. The pyranopterin Mo enzymes [2, 4–6, 9] are essential to human health and life processes and catalyze important chemical transformations that occur in the metabolic pathways of carbon, nitrogen, and sulfur compounds [7, 9, 23, 24]. A properly synthesized molybdenum cofactor (Moco) is essential in humans, since Moco deficiency results in severe neurological disorders and infant mortality [25, 26]. XO family enzymes catalyze important oxidative transformations of N-heterocycles and aldehydes, and are involved in catabolic processes [27], xenobiotic detoxification [27], drug metabolism [27–30], prodrug activation and conversion [31–35], oxidative stress leading to postischemic reperfusion injury [36], and NO synthase activity [37]. Carbon monoxide dehydrogenase (CODH) is a unique heterobimetallic XO family enzyme, which possesses a Mo-Cu active site and catalyzes the oxidation of CO to CO_2 [3, 38, 39].

Vertebrate SO is located in the mitochondrial intermembrane space [5, 40] and catalyzes the physiologically important oxidation of sulfite to sulfate, a reaction that represents the terminal step in the oxidative degradation of cysteine and methionine [7]. Isolated sulfite oxidase deficiency derives from specific mutations in the SO gene, and leads to severe neurological abnormalities, dislocation

of the ocular lens, and attenuated brain growth [41]. In plants, SO is localized in peroxisomes [40, 42–45] and removes the toxic metabolite sulfite to protect the cell [42]. Bacterial methionine sulfoxide reductase (MsrP, formerly YedY) [46–48] is an SO family enzyme that functions to repair oxidized surface methionines that have been damaged by antimicrobials [49].

The DMSO reductase family enzymes are the most structurally and catalytically diverse of the three pyranopterin Mo enzyme families [4, 5]. Bacterial DMSO reductase and trimethylamine-N-oxide reductase (TMAO reductase) are of increasing environmental importance since they catalyze the oxidation of marine osmolytes and facilitate cloud formation and albedo [8]. TMA/TMAO regulation is of human health importance since TMA can be converted to TMAO by human gut microbiota and TMAO has been implicated in cardiovascular disease, glucose and lipid homeostasis, and reverse cholesterol transport [50]. Although TMAO may function as a pathogenic molecule in humans, its accumulation may derive from cellular adaptation to stress, with TMAO being a biomarker and not a mediator of disease [51].

1.3. Scope of the Contribution

This contribution primarily focuses on developments since ~2000 that are related to our structural and mechanistic understanding of the three canonical pyranopterin Mo enzymes XO, SO, and DMSO reductase. Highlighted are the enzyme structures, mechanisms, key spectroscopic studies, and important electronic structure contributions to their reactivity. Although not explicitly covered in this contribution, the role of model chemistry [2] in enhancing our understanding of pyranopterin Mo and W enzyme structure and mechanism cannot be understated [2]. This derives from the fact that most of these enzymes possess additional strongly absorbing chromophores that impede their spectroscopic investigation.

2. THE MOLYBDENUM COFACTOR – Moco

2.1. Molybdopterin Biosynthesis

In humans, the biosynthesis of Moco is a four-step process that involves six gene products. The proposed mechanism for mature Moco formation is presented in Figure 1. The *MOCS* and *GEPHYRIN* genes are involved in human Moco biosynthesis with analogous cognate proteins being found in bacteria and plants [52]. Starting from guanosine 5′-triphosphate (GTP) the first step in this process involves GTP conversion to a cyclic pyranopterin monophosphate (cPMP) intermediate. The initial conversion of 5′-GTP to 3′,8-cH$_2$GTP is catalyzed by MOCS1A, while MOCS1B catalyzes the conversion of the 3′,8-cH$_2$GTP intermediate to cPMP. In Figure 1, cPMP is shown as the *gem* diol, consistent with the solution NMR structure [53]. However, in the MoaE X-ray structure cPMP possesses a keto oxygen on the pyran ring [54]. In the second step, cPMP is sulfurat-

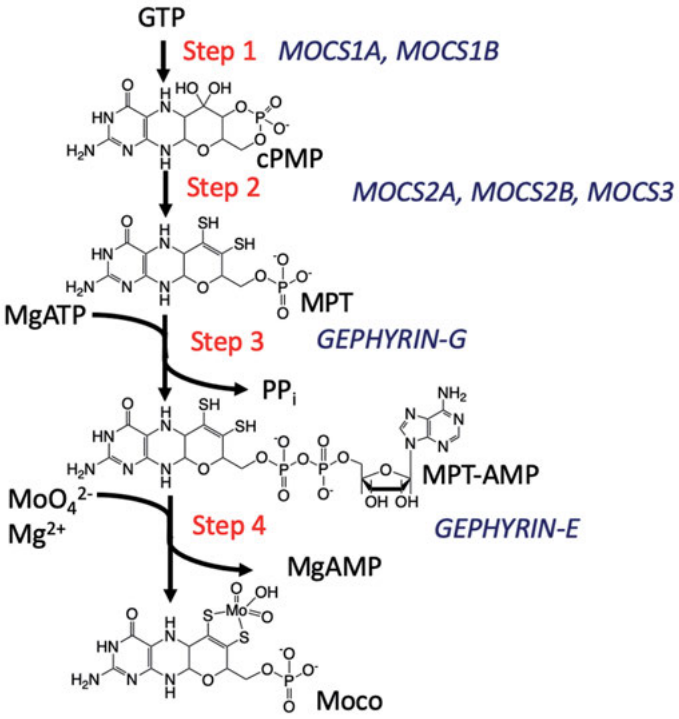


Figure 1. The Moco biosynthesis pathway. Relevant intermediates are shown, along with the human proteins that catalyze the individual reactions [52, 56, 58].

ed and converted to the mature pyranopterin, commonly referred to as molybdopterin (MPT; or the pyranopterin dithiolene, PDT), by an MPT synthase. Here, the sulfur insertion into cPMP results in the formation of a *cis*-dithiolene chelate via sequential sulfuration processes [55]. MPT is adenylated in the third step to yield MPT-AMP, which is catalyzed by GEPHYRIN G prior to MPT-AMP binding to GEPHYRIN-E, which is the locus of Mo incorporation [52]. It is currently believed that the AMP component of MPT-AMP may function as an anchor [56, 57]. An anchoring role for AMP would allow for proper positioning of the MPT dithiolene to enhance MPT-AMP binding to enzyme-bound molybdate via the dithiolene side chain of MPT [56, 57].

2.2. Molybdate Insertion, Molybdenum Cofactor Sulfuration, and Additional Molybdenum Cofactor Modifications

Although the steps required for the formation of MPT-AMP are well-understood, the final Moco maturation step is not. This final Moco formation step involves Mo insertion from molybdate into MPT, and this is catalyzed by insertase proteins [56–58]. Recently, high resolution X-ray data sets of the plant Mo-

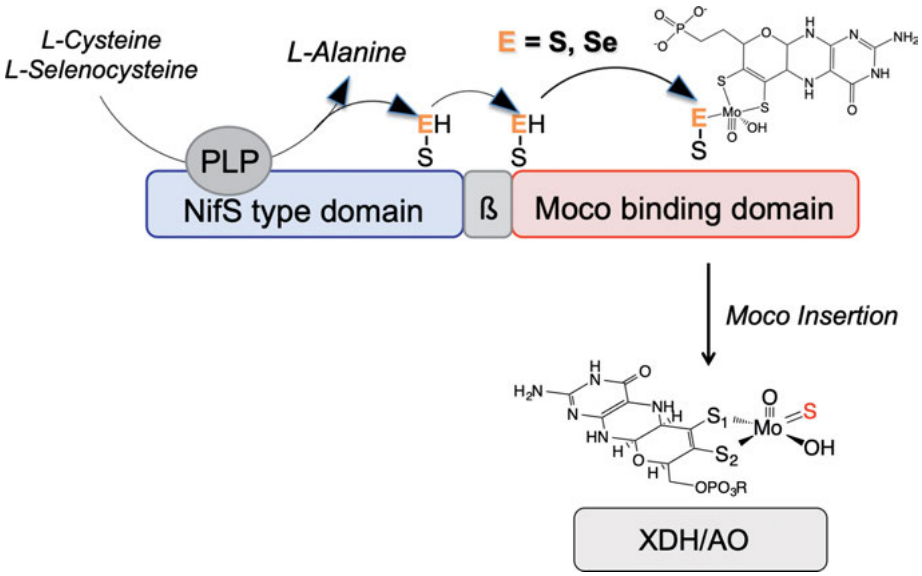


Figure 2. Proposed mechanism for sulfur incorporation (as sulfido) in the human Mo cofactor sulfurase (HMCS). The HMCS modified cofactor can then be incorporated into *apo* enzyme forms of the xanthine oxidase family.

insertase Cnx1E were used to show that there are two molybdate binding sites in the active site [56]. The presence of molybdate anions in a specific site appears to be tied to a distinct backbone conformation, and this may be essential for molybdate selectivity as well as efficiency of insertion [56].

Moco as shown in Figure 1 can be further modified for use in XO and DMSO reductase family enzymes. For XO family enzymes, molybdenum sulfurase proteins (e.g., HMCS-CT/ABA3-CT) catalyze the insertion of the catalytically essential terminal sulfido ligand into Moco (Figure 2) [59–61]. The sulfido S derives from cysteine and is extracted in the sulfurase NifS-type domain of HMCS and eventually transferred to Moco via a cysteine persulfide in the sulfurase domain [59–66]. Other Moco modifications are found in archaeal and bacterial enzymes. In these enzymes, a nucleotide is attached to the MPT phosphate to form a dinucleotide version of Moco. Bacterial DMSO reductase family enzymes possess a bis-molybdopterin guanine dinucleotide (bis-MGD) cofactor that is obtained by adding GMP to the MPT terminal phosphates [5].

3. MOLYBDOPTERIN – THE PYRANOPTERIN DITHIOLENE

3.1. Structure of Molybdopterin

The basic structure of MPT has been elucidated by numerous X-ray crystallographic studies on SO [67–75], XO [12, 76–84], and DMSO reductase [85–90]

family enzymes as well as pyranopterin W enzymes [91, 92]. MPT is a tri-cyclic ligand that possesses a dithiolene chelate fused to a heterocyclic pterin ring system by a bridging pyran ring. An alternative form of MPT may also exist as a pyran ring-opened structure in some enzymes [93–95]. MPT is a remarkably elaborate molecule and represents one of the most electronically flexible ligands in biology [4, 21, 96]. This electronic flexibility derives from that fact that both the dithiolene and pterin components of MPT are redox active moieties. Generally, the structural resolution has not been sufficiently high enough to differentiate between C–C and C–N single- and double-bonds in MPT, and to accurately determine the oxidation and protonation state of MPT in the enzymes. The MPT is known to be extensively hydrogen-bonded to the polypeptide but is not covalently attached to the protein. For example, X-ray crystallographic studies of XO family enzymes provide evidence for fifteen highly conserved hydrogen bonding interactions [76, 82–84] that tightly bind Moco to the protein.

In order to obtain greater insight into the electronic and geometric structure of the MPT, a novel study investigated the conformations of 319 MPTs found in 102 different pyranopterin molybdenum and tungsten enzyme X-ray structures [21]. Remarkably, different degrees of MPT non-planarity were observed that correlated with protein function. Namely, the nature of the MPT conformations were shown to correlate with a given enzyme family. The most distorted MPT structures were found in XO family enzymes, with lesser degrees of MPT non-planarity correlating with SO family enzymes. Enzymes that belong to the DMSO reductase family possess two coordinated MPTs, with one of the MPTs adopting a SO family configuration and the other a more distorted XO enzyme family structure. Geometry optimizations of MPT show that the magnitude of the out-of-plane distortion increases as the MPT is reduced from the oxidized, to the dihydro, to the tetrahydro form. However, this correlation is merely suggestive of the MPT adopting different oxidation states in different protein families and does not constitute proof that different MPT redox states exist in the enzymes.

3.2. Role of Molybdopterin in Catalysis

Although MPT has been suggested to play various roles in the catalytic cycles of pyranopterin Mo and W enzymes [2, 21, 97–99], the precise function of MPT is not known with certainty. However, enzyme and MPT structures provide critical insight into MPT function. XO family enzymes and the arsenite oxidase from the DMSOR enzyme family do not possess a Mo ligand that derives from the polypeptide. Thus, MPT functions as an anchor to properly position the cofactor in the active sites of these enzymes using an extensive network of hydrogen bonds. Such MPT hydrogen bonding networks may also control the degree of MPT non-planarity and the magnitude of the MoS_2C_2 chelate ring fold angle [21]. This would allow for static and dynamic Mo-dithiolene charge redistribution, also known as the electronic buffer effect [100], along the electron and

atom transfer reaction coordinates during catalysis [101]. In this way, enzyme-MPT hydrogen bonding would provide a means of converting protein vibrational energy into specific electronic structure changes at the Mo center to modulate metal-dithiolene covalency and the metal ion reduction potential [4, 102–106]. MPT is also likely to function as an electron transfer conduit in some enzymes to couple the metal ion with specific electron transfer partners [107, 108]. For example, an electron transfer chain comprised of the Mo ion, two [2Fe-2S] clusters, and FAD is found in XO family enzymes that provides a pathway for electron egress from Mo to the ultimate electron acceptors NAD^+ or dioxygen.

Evidence supporting a role for MPT as an electron transfer conduit [107, 109, 110] derives, in part, from the structure of bovine xanthine dehydrogenase (XDH) [12], in addition to electron transfer kinetics [9, 108, 111, 112] and resonance Raman (rR) spectroscopic probes of the pyranopterin dithiolene [113]. In DMSO reductase family enzymes, which possess two coordinated MPT ligands, the more distorted proximal MPT may function as an electron transfer conduit while the distal MPT may function to control the redox potential [21]. Recently, Fourier-transformed alternating-current voltammetry experiments were performed and used to suggest that the MPT pterin is redox active in Mo-containing methionine sulfoxide reductase (MsrP) catalysis, with the pterin providing the electron equivalents necessary to perform two-electron OAT reactivity [114]. Although the potential for ligand-redox reactions in the catalytic cycles of pyranopterin-containing Mo and W enzymes is apparent from the inherently redox-active nature of the ligand, the two-electron chemistry in all of the other enzymes can be accounted for by redox cycling through the metal M(IV), M(V), and M(VI) oxidation states [9]. As such, a redox-active MPT in the catalytic cycle of MsrP would represent a new paradigm in Mo and W pyranopterin enzyme reactivity. Synergistic interactions that occur between the pterin ring and the dithiolene may be important in the catalytic cycles of some enzymes [115]. This idea is supported by recent enzyme crystal structures that show the pyran ring of MPT can exist in a ring-closed pyrano form or a ring-opened non-cyclized form [93–95].

4. OXO TRANSFER REACTIVITY

4.1. Making and Breaking Strong Metal-Oxo Bonds

Over 30 years ago, Holm and coworkers placed OAT on a thermodynamic reactivity scale [116–118]. This effort has allowed for a greater conceptual understanding of both reductive $M^{x+2}O + Y^y \rightarrow M^x + Y^{y+2}O$ (M = Mo, W; Y = substrate) and oxidative $M^x + Y^{y+2}O \rightarrow M^{x+2}O + Y^y$ OAT reactions that involve monooxo and dioxo Mo and W centers with a variety of different substrates. The OAT reaction couples for the tungsten systems are found to be more negative than the corresponding values for molybdenum at structural parity. This is con-

sistent with the difference between the strength of the W-oxo bond in $[WOCl_4]$ and that of the Mo-oxo bond in $[MoOCl_4]$. The $M^{VI}{\equiv}O$ binding energy for $[MoOCl_4]$ is 101 kcal/mol and the corresponding BDE for $[WOCl_4]$ is 26 % higher at 127 kcal/mol [119]. Thus, W^{VI}-oxo systems are expected to be poorer oxo donors and better oxo acceptors when compared with their Mo counterparts. These experimental bond strengths parallel the corresponding reduction potentials of Mo and W centers, with Mo complexes being markedly easier to reduce compared to W complexes at structural parity.

The two-electron chemistry common to pyranopterin Mo and W enzymes may be utilized to activate strong M-oxo bonds by transferring an electron pair to a metal-oxo antibonding orbital along the reaction coordinate in order to weaken the bond. The enzymatic OAT systems discussed here possess either one (DMSO reductase family enzymes) or two (SO family enzymes) terminal oxo ligands. In monooxo sites, the terminal Mo- and W-oxo bonds are formally triple bonds (M≡O; bond order = 3), while the formal metal-oxo bond order is reduced to 2.5 for symmetry equivalent dioxo sites. Thus, for efficient oxo transfer to substrates (i.e., substrate oxidation), the two electrons should be transferred in a manner that will allow maximum activation of a single oxo ligand, and this will be controlled by how the substrate accesses the metal center in the active site in addition to the geometric and electronic structure of the Mo site in the enzymes.

4.2. The Mechanism of Oxo Transfer Reactivity Catalyzed by Dioxo Sites

4.2.1. Sulfite Oxidase as an Exemplar of Dioxo Reactivity: Enzyme Structure

There are at least three subfamilies of sulfite oxidase enzymes [5]. The first includes the sulfite dehydrogenases [120–123], the eukaryotic assimilatory nitrate reductases [124–126], and the canonical sulfite-oxidizing enzymes (i.e., the sulfite oxidases) [42, 44, 45, 67, 68, 71, 122, 127–129]. Enzymes such as the bacterial methionine sulfoxide reductase [49, 114, 130–134] and the mammalian mitochondrial amidoxime-reducing component (mARC) [13, 15, 28, 33, 66, 135–140] make up the second subfamily, with the third subfamily being comprised of MOSC domain proteins (e.g., HMCS-CT) that are involved in Moco sulfuration [60, 66]. In general, we will restrict our discussion to the sulfite-oxidizing enzymes [141–143], which convert toxic sulfite to sulfate via 2-electron OAT.

A number of X-ray crystal structures have been determined for SOs from a variety of organisms [67, 72, 74, 144, 145], and the coordination geometry about the Mo center has been probed by X-ray absorption spectroscopy (e.g., EXAFS) [146, 147]. In the oxidized Mo(VI) state, the Mo ion adopts a square pyramidal geometry, and is coordinated by MPT, two terminal oxo ligands, and a cysteine S donor (Figure 3). The reduced enzyme possesses a similar square pyramidal geometry with the equatorial oxo ligand being replaced by solvent water. Numer-

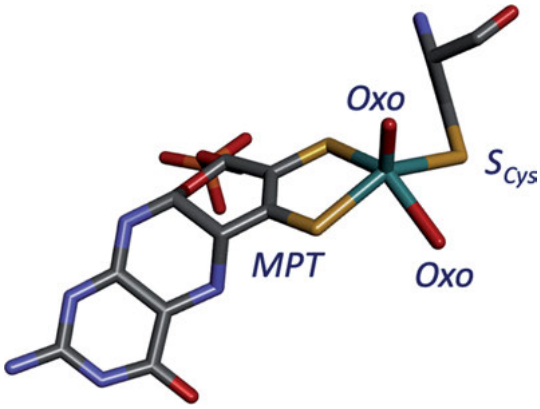


Figure 3. Structure of chicken liver sulfite oxidase (PDB 1SOX) [74].

ous spectroscopic studies have been performed on the paramagnetic Mo(V) form of SO [127, 148–152], which possesses a $[(MPT)MoO(S_{Cys})(OH)]^{1-}$ structure. An oxo ligand occupies the apical position of the square pyramidal structure in all three oxidation states. The crystal structures of SO show that the substrate access channel points toward the equatorial oxo ligand, and possesses a positively charged Arg-rich binding pocket for the stabilization and proper positioning of substrate in the active site.

4.2.2. The Mechanism of Sulfite Oxidase

Studies on dioxo-molybdenum model complexes indicate that OAT occurs via an associative mechanism. By using triethylphosphine (PEt_3) as a sulfite mimic, a Mo-O-PEt_3 intermediate was observed that displayed Mo–O bond lengthening (2.18 Å) with a decrease in the length of the remaining Mo≡O bond (1.67 Å) [153], confirming the "spectator oxo" effect detailed by Rappé and Goddard [154]. Early computational work on high symmetry dioxo-molybdenum models, where the oxo ligands are symmetry-equivalent, showed that the two-electron oxidation of sulfite to sulfate was initialized by an attack of the sulfite S lone pair on an oxo ligand to produce sulfate bound to a reduced Mo(IV) center [155, 156]. Enzyme studies [45, 157] also support this mechanism and indicate that the sulfite lone pair attacks the oxo ligand that is directed toward the substrate access channel. The pH dependence on the steady-state kinetics are similar for both the chicken and *Arabidopsis thaliana* enzymes, as are the $K_m^{sulfite}$ values [45, 157]. In contrast, the rate of substrate oxidation is observed to be markedly faster in the plant enzyme [45, 157]. The enzyme mechanism with sulfite as substrate has been evaluated computationally using quantum mechanical (QM) and QM molecular mechanics approaches (QMMM) [158]. The results of the computational studies show that substrate lone pair attack on an oxo ligand to form the E-P complex is rate limiting, with a computed activation barrier of ~9–12 kcal/mol. The Mo-

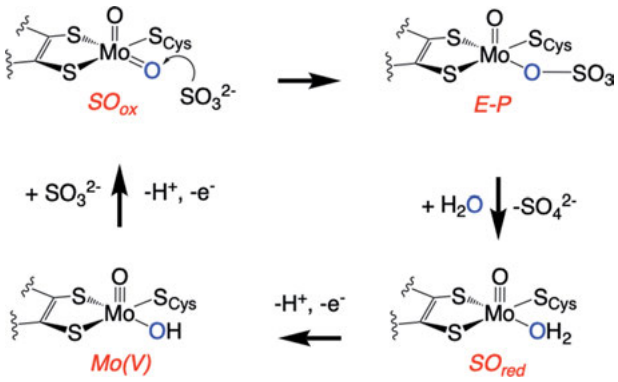


Figure 4. Proposed mechanism for sulfite oxidase consistent with spectroscopic and computational studies [45].

bound sulfate intermediate (E-P) subsequently breaks down in a second step by product dissociation (Figure 4).

Electron transfer regeneration of the oxidized Mo(VI) state following OAT in the oxidative half reaction of SO occurs via two sequential $1e^{-}/1H^{+}$ transfers [159]. In vertebrate enzymes, the two electron equivalents are shuttled out of the Mo site via a *b*-type cytochrome and are then used to reduce the physiological oxidant, cytochrome *c*. In contrast, the plant SO from *A. thaliana* lacks the cytochrome b_5 domain that is found in the mammalian and avian sulfite oxidases. In vertebrate enzymes, the heme and the Mo domains are linked via a flexible polypeptide tether, with the Fe and Mo centers being separated by 32 Å in the X-ray structure of chicken SO [74]. This distance is inconsistent with the observed rate of electron transfer, indicating that the X-ray structure does not reflect the productive electron transfer geometry. Intramolecular electron transfer kinetics that probe tether length and flexibility in human SO suggest that the observed k_{et}, driving force, and reorganizational energy are a function of numerous averaged geometric conformations [159]. Electron transfer has also been evaluated in the sulfite dehydrogenase from *Starkeya novella*, which is a heterodimeric complex that consists of a Mo subunit and a *c*-type cytochrome subunit with a Mo–Fe distance of 16.6 Å [144]. Here, rapid through-bond electron transfer has been suggested to occur via a direct link involving Arg-55A in addition to a pathway involving aromatic residues [144, 160].

4.2.3. Electronic Structure Contributions to Reactivity

The concept of dioxomolybdenum symmetry breaking leading to activation of a specific oxo ligand being transferred to substrate was first suggested by Izumi et al. [161], who showed that the O_{apical}-Mo-$S_{thiolate}$-C torsion angle in a $[(L\text{-}N_3)MoO_2(SCH_2Ph)]$ model can effectively select for a specific oxo. Thus, symmetry breaking coupled with the pre-transition state initial attack of the

substrate lone pair functions to activate a strong Mo-oxo bond for OAT. Computational studies performed at the density functional level of theory on a $[Mo^{VI}O_2(S_2C_2Me_2)(SMe)]^{1-}$ model for oxidized SO argue against the O_{apical}-Mo-$S_{thiolate}$-C torsion angle controlling OAT in the enzymes [162]. However, this torsion angle may contribute to the reduction potential of the reduced monooxo molybdenum SO site [163, 164] to control electron transfer regeneration of the oxidized active site.

A rR vibrational analysis of *A. thaliana* SO using ^{18}O labeling has allowed for a detailed interpretation of the Mo-oxo stretching region [45], and the vibrational frequency shifts that occur upon ^{18}O labeling are consistent with an active site that possesses geometrically inequivalent oxo ligands. This data supports a redox orbital (LUMO) with a high degree of Mo(xy) – $O_{equatorial}$ d-p π^* character. Electron occupancy of this orbital occurs with substrate lone pair attack on the equatorial oxo to activate the Mo=O_{eq} bond and lower the activation barrier for OAT. A very recent study has investigated the effects of charge on OAT reactivity [165]. Here, a combination of kinetic and computational studies were performed on dioxo-molybdenum systems at structural parity, but varying by a single unit of charge. The results show that the more positively charged complex possesses a greater thermodynamic driving force, earlier transition state, and lowered activation barrier for OAT, consistent with the Bell-Evans-Polanyi principle [166]. Remarkably, OAT rate enhancement was observed to increase ~500-fold as the charge of the complex increased by +1, indicating that charge effects in the enzymes may play a dramatic role in catalysis.

4.3. The Mechanism of Oxo Transfer Reactivity Catalyzed by Monooxo Sites

4.3.1. Dimethylsulfoxide Reductase as an Exemplar of Monooxo Reactivity: Enzyme Structure

The DMSO reductases from *Rhodobacter capsulatus* and *Rhodobacter sphaeroides* represent the canonical enzymes of the DMSO reductase family. The DMSO reductase family enzymes have been divided into three classes (Types I, II, and III) [4, 5] that are distinguished from each other by their active site structure and the nature of the donor ligand that is provided by the polypeptide. However, DMSO reductase family enzymes are quite diverse and not all of the enzymes in this family adhere to this general classification scheme [85, 167]. DMSO reductases are Type III enzymes and a combination of EXAFS [168] and high resolution X-ray crystallography [169] shows that the oxidized active site possesses a distorted six-coordinate trigonal prismatic $[(MPT)_2MoO(O_{Ser})]^{1-}$ coordination geometry (Figure 5). EXAFS of the reduced Mo(IV) state reveals a *des*-oxo $[(MPT)_2Mo(OH_2)(O_{ser})]^{1-}$ structure [170]. The Type III enzymes are thus distinguished by having a coordinating serine residue bound to Mo, and

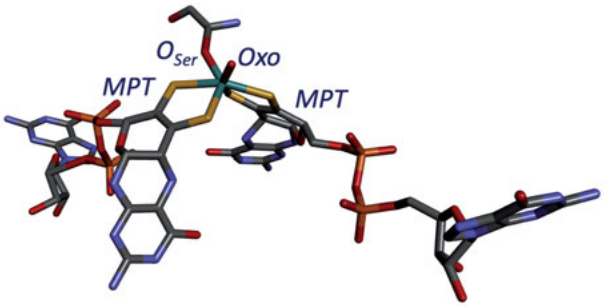


Figure 5. Structure of oxidized dimethylsulfoxide reductase showing apical oxo, serine oxygen, and bis-MGD dithiolene coordination (PDB 1eu1 [169]).

include trimethylamine-N-oxide (TMAO) reductase and the bacterial DMSO reductases. A coordinated S-Cys or Se-Cys ligand is observed for the Type I enzymes, which include formate dehydrogenases and prokaryotic periplasmic nitrate reductases (NAP). The Type II enzymes possess a coordinated aspartate and include ethylbenzene dehydrogenase and the prokaryotic respiratory nitrate reductase (NarGHI). Our focus here will be on the DMSOR from *Rhodobacter sphaeroides*, which was the first enzyme in this family to be structurally characterized by X-ray crystallography [90].

4.3.2. The Mechanism of Dimethylsulfoxide Reductase

Electronic absorption spectral changes observed in *R. sphaeroides* DMSO reductase obtained under turnover conditions using DMSO or TMAO as oxidizing substrate have been used to obtain remarkably detailed information regarding the reaction mechanism of this enzyme [171]. Reactions of reduced enzyme with DMSO are biphasic in the pH range 6–10, and reveal a fast initial substrate-binding phase followed by a catalytic phase that is independent of the substrate concentration. Spectral deconvolution of the absorption spectrum over the time course of the reaction reveals contributions from four distinct species that are catalytically relevant. These include an oxidized Mo(VI) state, a one-electron reduced Mo(V) *high-g split* intermediate, the fully reduced enzyme in the Mo(IV) state, and a reduced enzyme form with DMSO bound to the Mo site. Although the reduced DMSO-bound enzyme form accumulates when DMSO is used as the substrate, the situation with TMAO as oxidizing substrate is different. Using TMAO, rate-limiting Mo(V) → Mo(IV) reduction leads to near quantitative accumulation of the paramagnetic *high-g split* intermediate, which has been studied in detail by EXAFS and electronic absorption, magnetic circular dichroism (MCD), and EPR spectroscopies [172, 173]. A number of computational studies employing various methods [172, 174–184] have been performed to address the nature of the DMSO reductase reaction coordinate [185] and they are in general agreement with the reaction mechanism provided in Figure 6.

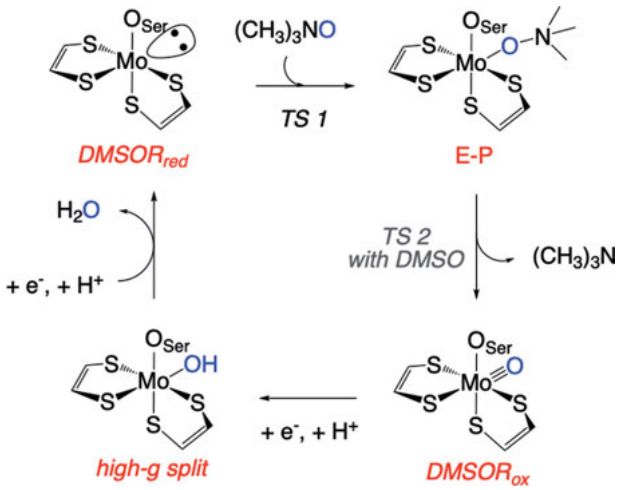


Figure 6. Proposed mechanism for DMSO reductase. Note that the *high-g split* intermediate is located here on the catalytic pathway in the reductive half reaction of the catalytic cycle [172].

In this mechanism, the catalytic cycle begins with the oxidized Mo(VI) state of the enzyme, which is converted to the *high-g split* Mo(V) state via a coupled e^-/H^+ transfer. A second coupled e^-/H^+ transfer, followed by loss of water, leads to a fully reduced five-coordinate *des*-oxo Mo(IV) resting state. This *des*-oxo form adopts a six-coordinate distorted trigonal prismatic geometry upon binding of substrates [172]. OAT followed by the release of product leads to the formation of the fully oxidized monooxo Mo(VI) enzyme form.

4.3.3. Electronic Structure Contributions to Reactivity

The distorted trigonal prismatic active site structure in DMSO reductase imparts significant electronic structure contributions to enzymatic catalysis. The early computational studies of Webster and Hall [185] noted that the low-symmetry twisting of the MPT dithiolenes away from a trigonal prismatic geometry raised the energy of the reduced Mo(IV) state [172, 178, 180, 181]. This effectively contributes to lowering the activation barrier for the OAT half reaction [172, 178, 180]. With DMSO as substrate, two transition states are observed. The first represents substrate binding to Mo(IV) leading to the formation of a weakly bound Mo(IV)-substrate species [180, 181], while the second leads to the largest barrier along the reaction coordinate and corresponds to the S–O bond cleavage step.

A detailed EPR, electronic absorption, and MCD spectroscopic study has been performed on the *high-g split* intermediate [172], which builds up to ~100 % under enzyme turnover conditions with TMAO as substrate. The experimentally determined spin-Hamiltonian parameters derived from the EPR data reveal a

rare rhombic ^{95,97}Mo hyperfine tensor, which correlates with a low-symmetry coordination geometry for this intermediate. The experimental rhombic ^{95,97}Mo hyperfine interaction is reproduced by a fully relaxed six-coordinate $[(\text{dithiolene})_2\text{Mo(OMe)(OH)}]^{1-}$ computational model for *high-g split* that has been interpreted in the context of an energetically stable distorted trigonal prismatic coordination geometry for this intermediate [172]. The relaxed geometry of the *high-g split* computational model also reproduces the observed MCD and electronic absorption transition energies and intensities, as well as the correct sign of the MCD bands. The SOMO redox orbital of *high-g split* possesses a low degree of Mo-$S_{dithiolene}$ covalency that has been proposed to reduce the magnitude of the electronic coupling for electron transfer [172], which when coupled with the energetic stability of *high-g split* provides a key contribution to the slow kinetics of the Mo(V) → Mo(IV) step and buildup of the intermediate. The use of TMAO as substrate results in a computed reaction coordinate for the oxidative half-reaction that possesses a single transition state, which arises from nascent Mo–O_{TMAO} bond formation leading to an effectively activationless N–O bond cleavage step [172].

5. HYDRIDE TRANSFER REACTIVITY

5.1. Xanthine Oxidase as an Exemplar of Formal Hydride Transfer Reactivity

5.1.1. The Structure of Xanthine Oxidase

X-ray crystallography and EXAFS have been used to structurally characterize XO family enzymes. These studies reveal a 5-coordinate square pyramidal coordination geometry for XO enzymes and the related aldehyde oxidases (AO) [12, 76, 78, 79, 83, 186–189]. The Mo ion in the oxidized enzyme is coordinated by a single MPT, terminal oxo and sulfido donors, and a water-based ligand that is oriented toward the substrate access channel. The initial X-ray crystallographic data for XO were interpreted in the context of the terminal sulfido ligand residing in the apical position, with the terminal oxo ligand being coordinated to Mo in the equatorial plane. In contrast, MCD studies of the *very rapid* Mo(V) intermediate were used to show that the terminal oxo ligand was located in the apical position [107]. This spectroscopically derived coordination environment was later observed in a high resolution X-ray study of a reduced Mo(IV) mechanism-based inhibitor complex [12] (Figure 7). Determining the orientation of the sulfido ligand in the structure was extremely important since this is a catalytically essential ligand. In the reduced Mo(IV) state the terminal sulfido ligand is protonated and coordinated to Mo as a sulfhydryl (-SH) group.

Also essential for catalysis is a conserved glutamate that functions as a base to activate the Mo-OH hydroxide for nucleophilic attack on substrates. Mo K-edge EXAFS of oxidized XO shows that the Mo-OH bond decreases from 1.97 Å at pH 6 to 1.75 Å at pH 10 [190], and this is consistent with metal activat-

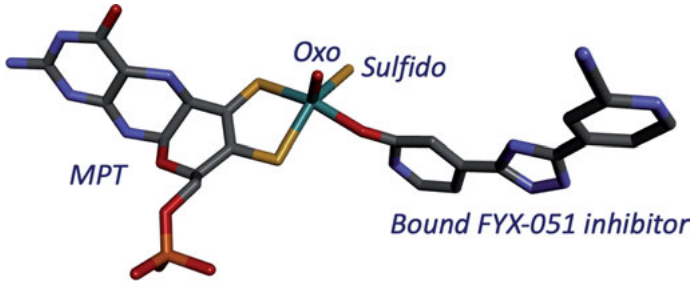


Figure 7. Structure of product-bound bovine xanthine dehydrogenase (PDB 1V97). The structure shows the coordinated MPT, apical oxo, and equatorial sulfhydryl ligands. The inhibitor FYX-051 is bound in the equatorial position that would normally be occupied by hydroxide in oxidized enzyme [12].

ed water being present as a hydroxide ligand at physiological pH. The EXAFS work provided strong evidence that this ligand was not a water molecule, as had been previously suggested by earlier X-ray crystallographic studies, and supports a reaction mechanism that begins with base-assisted nucleophilic attack of the substrate by Mo-OH. A second glutamate (Glu) is found in the substrate binding pocket that may be involved in controlling the tautomerization state of the substrate, while an active site arginine (Arg) has been proposed to affect the orientation of substrate molecules [5, 191, 192]. These Glu and Arg residues located in the substrate binding pocket of XOs are not conserved in AOs [5].

5.1.2. Spectroscopic Studies Related to Xanthine Oxidase Structure and Reactivity

In addition to the MCD studies that were used to determine that the oxo ligand was apical in the *cis*-MoOS active site, XO has been probed by electronic absorption, rR, EPR, and ENDOR spectroscopies. As is the case for most pyranopterin Mo enzymes, the use of optical spectroscopies (e.g., MCD, electronic absorption, rR) for interrogating the Mo center is limited due to the presence of other strongly absorbing chromophores. In XO family enzymes, these complications arise from the presence of [2Fe-2S] and FAD redox chromophores in the proteins. As a result, paramagnetic resonance studies dominate the spectroscopic literature of the enzymes. Important EPR signals that have been observed include *very rapid*, *inhibited*, *rapid*, and *slow*. Structures associated with these EPR signals are provided in Figure 8.

XO *very rapid* possesses a high g_{max} around 2.025 that derives from the presence of the terminal sulfido ligand. The *very rapid* intermediate is of mechanistic importance due to it being a Mo-product species formed under specific conditions during the electron transfer (oxidative) half-reaction. Although the magnitude of g_{max} is not a function of enzyme source, (e.g., *Pp* QuinOr, *bovine milk* XO, *Pd* IsoOr, *As* QualOx) g_{mid} and g_{min} appear to be correlated with the en-

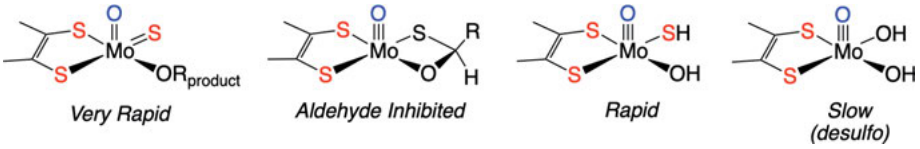


Figure 8. Paramagnetic Mo(V) forms of xanthine oxidase. The *very rapid* species is a true enzyme intermediate observed under specific reaction conditions and is a product-bound species. *Aldehyde-inhibited* (R = H, formaldehyde) xanthine oxidase possesses a tetrahedral carbon center and serves as a mimic of the Mo-O-R interaction present in the catalytic cycle of the enzyme. *Rapid* is a paramagnetic analog of the Michaelis complex and the *slow* form is a paramagnetic desulfo form of the enzyme that lacks the terminal sulfido-based ligand.

zyme source [193]. The nicotinic acid hydroxylase from *Clostridium barkeri* possesses a terminal selenido ligand in place of the terminal sulfido. As a result, a very high $g_{max} = 2.067$ is observed, and contributions to the magnitude of the g-shift reflect lower energy charge transfers, increased spin-orbit coupling, and increased covalency [194–198]. EPR spectra of *very rapid* using 2-oxo-6-methylpurine [199, 200] or xanthine [199–201] as the reducing substrate show a large ^{33}S hyperfine coupling constant (A_2 ~ 77–82 MHz; $A_{average}$ ~ 41 MHz) [199, 200] indicating the unpaired electron is highly delocalized with a considerable spin population on the terminal sulfido ligand (~38 %). In 1-methylxanthine *rapid type 1*, the ^{33}S hyperfine coupling is dramatically reduced ($A_{average}$ ~ 10 MHz) [200, 201] consistent with protonation of the sulfido and a concomitant reduction in Mo-S covalency. EPR and ENDOR spectroscopies of XO *very rapid* show strong and isotropic hyperfine coupling of the unpaired electron to ^{17}O [202, 203]. When this data is analyzed in the context of the ENDOR determined ^{13}C hyperfine tensor [204] and single turnover ^{17}O experiments [205] a viable structure for XO *very rapid* emerges as a Mo^{V}-$OR_{product}$ bound species. The Mo–C8 distance determined by ENDOR is consistent with an Mo–O–C bonding scheme and not a direct Mo–C bond, resulting in the product being bound to Mo as the enolate tautomer. Thus, the combination of ^{13}C and ^{17}O EPR and ENDOR data support a mechanism that involves nucleophilic attack of metal-activated water on the substrate carbon center in the catalytic sequence.

Other EPR signals that have been observed include various types of *inhibited* signals [193, 201, 202, 206–208]. The structure of the formaldehyde-*inhibited* species has recently been determined by ENDOR [209] and used to support Mo-O-R through-bond contributions to C–H bond cleavage in the reaction mechanism [206]. *Rapid type 1* and *rapid type 2* EPR signals arise from dead-end Mo(V) species with unbound substrate in the active site and have been suggested to be paramagnetic analogs of the Michaelis complex [210]. The *slow* signal derives from a non-functional desulfo form of the enzyme [9].

Bovine XO and *R. capsulatus* xanthine dehydrogenase (XDH) have been studied by rR and electronic absorption spectroscopies [113]. These enzymes are able to oxidize lumazine to violapterin [78, 211, 212], resulting in the formation of a reduced Mo(IV) enzyme-product complex that displays an intense charge

transfer band [213] assigned as a Mo(IV) → violapterin charge transfer transition [113, 214]. Resonance Raman data collected on resonance with the Mo(IV) → violapterin transfer transition have been used to assign specific vibrational modes for the violapterin product complex [211, 215]. More recently, new XO/XDH reducing substrates (4-thiolumazine and 2,4-dithiolumazine) have been used to produce enzyme-product complexes that display intense metal → ligand charge transfer (MLCT) bands which have been shifted to the near infrared region of the electronic absorption spectrum [113, 191, 214]. At least nine low-frequency Raman bands are observed between 200 – 600 cm^{-1} [113, 214], which have been assigned as in-plane bending modes of the bound product complexes in addition to key low-frequency Mo-MPT vibrations. The asymmetric and symmetric S-Mo-S dithiolene stretching vibrations were assigned at 351 cm^{-1} and 326 cm^{-1}, respectively. The large charge redistribution that occurs as a result of the MLCT transition results in resonance enhancement of vibrations that possess motions that extend to the amino terminus of MPT, supporting a hypothesis that MPT is directly involved in enzymatic electron transfer, functioning as a conduit for shuttling electrons from Mo to the proximal [2F-e2S] I cluster in XO. Other rR studies on these MLCT enzyme-product complexes indicate how substrate orientation in the binding pocket is critical for efficient catalytic throughput and support the role of the Mo-O-C covalent linkage in the $2e^-$ oxidation of heterocyclic substrates [191].

5.1.3. The Mechanism of Xanthine Oxidase and Electronic Structure Contributions to Reactivity

The catalytically essential sulfido ligand in XO and AO enzymes plays a critical role in the enzyme mechanism by functioning as a *formal* hydride acceptor in the heterolytic cleavage of substrate C–H bonds. Much of what we know regarding the electronic structure of the *cis*-$[MoOS]^{2+}$ unit is derived from detailed spectroscopic and electronic structure studies of well-defined synthetic model systems that possess this structural moiety. Wedd and coworkers [200] used a combination of enzyme and model compound studies to show that the terminal oxo ligand in *cis*-$[MoOS]^{1+/2+}$ systems dominates the ligand field due to the oxo being a stronger π-donor ligand than sulfido. Their work showed that a highly covalent Mo(xy)-S(p) π interaction was present in XO *very rapid* and, by inference, present also in the oxidized form of the enzymes. This work was followed more than 15 years later by a more detailed electronic structure description of the *cis*-[MoOS] unit in model complexes [216]. This study used a combination of S K-edge XAS and vibrational spectroscopies, evaluated in the context of detailed bonding and spectroscopic calculations, to evaluate Mo-$S_{sulfido}$ bond covalency and the ligand field splitting of the *cis*-[MoOS] moiety, develop a valence bond description of the Mo-sulfido bond order, and show that electron occupation of the Mo(xy)-S(p) π* redox orbital modulates the relative electro/nucleophilicity of the terminal sulfido ligand. Together with the earlier work of Wedd [200], this *cis*-[MoOS] bonding description indicates that the enzyme reaction mechanism is under orbital control.

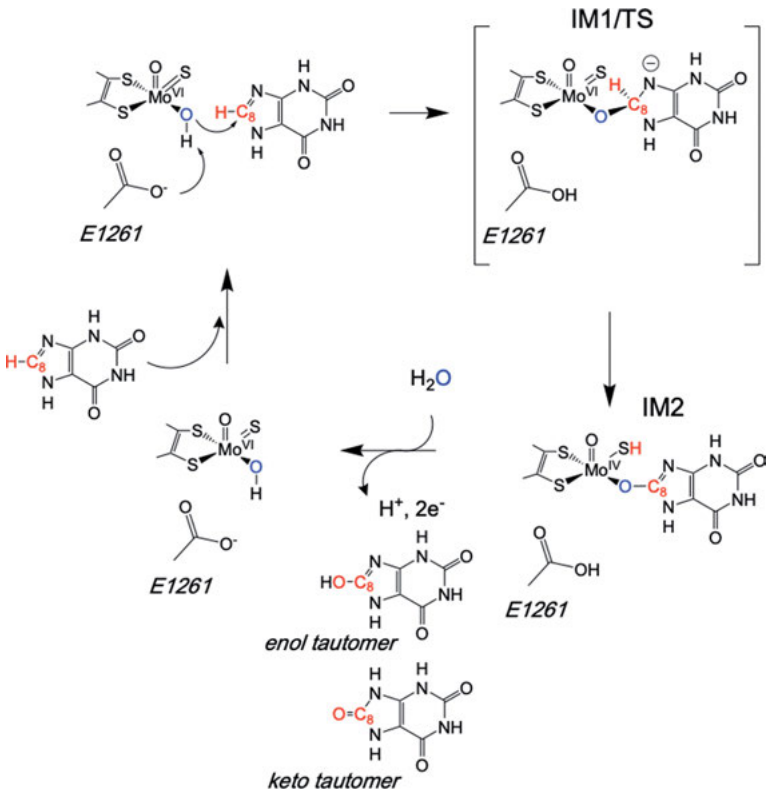


Figure 9. Proposed mechanism for xanthine oxidase. Note that IM1/TS possesses the tetrahedral-type carbon found in *formaldehyde-inhibited*. IM2 is the reduced Mo(IV) analog of *very rapid*. The role of the active site glutamate, E1261, is shown.

The XO reaction coordinate has been investigated by DFT [3, 217–219] and QMMM [218, 220–222] methods. The mechanism of XO presented in Figure 9 involves nucleophilic attack of a Mo-bound hydroxide on a substrate carbon atom (e.g., the C-8 carbon of xanthine), resulting in the formation of a tetrahedral intermediate or transition state [5, 78, 222–224]. The formation of a Mo-O_{eq}-$C_{product}$ linkage in the mechanism is supported by both spectroscopic and structural studies (*vide supra*) [78, 191, 211]. Consistent with $C_{substrate}$–H bond weakening at the transition state geometry [225], a large kinetic isotope effect has been observed when the C-8 H of xanthine is deuterated. The *formal* hydride transfer from the substrate to the *cis*-[MoOS] unit results in the two-electron reduction of Mo and protonation the terminal sulfido to yield a Mo(IV)-SH center. Thus, the closed shell nature of the hydroxylation mechanism is distinctly different from the radical mechanisms proposed for monooxygenase enzymes [7, 226].

A detailed EPR and computational study of XO addressed the importance of electron delocalizations present within the *cis*-[MoOS] and Mo-O_{eq}-C fragments

along the reaction coordinate and how they contribute to C–H bond activation, bond cleavage, and transition state stabilization [206]. A natural bond orbital analysis of the reaction coordinate noted the importance of Mo=S $\pi \rightarrow$ C-H π^* and C-H $\pi \rightarrow$ Mo=S π^* back donations for activating the substrate C–H bond [206]. Additional important contributions were shown to derive from O_{eq} lone pair $\rightarrow$ C-H σ^* and S lone pair $\rightarrow$ C-H σ^* interactions that contribute to energy stabilization at the transition state. The covalent Mo-O-R interaction is important, since the O_{eq} donor plays an even more dramatic role in stabilizing the transition state through $O_{eq} \rightarrow (Mo + C_{substrate})$ charge transfer. Importantly, recognition of this key Mo-O_{eq}-C delocalization pathway has allowed for a description of the transition state as a resonance hybrid of valence bond wave functions involving the reactant and the product [206].

5.2. The Formate Dehydrogenases

The Mo bacterial formate dehydrogenases (FDHs) are Type I DMSO reductase family enzymes that catalyze the oxidation of formate to CO_2. Interestingly, several bacterial formate dehydrogenases can also catalyze CO_2 reduction to formate. Three different genes in *E. coli* encode for three distinct formate dehydrogenases: formate dehydrogenase H [227], formate dehydrogenase N [228], and formate dehydrogenase O [5]. Additionally, many aerobic bacteria possess a Mo-containing NAD^+-dependent formate dehydrogenase [229–231] that is distinct from the $NAD(P)^+$-dependent formate dehydrogenases found in higher plants and yeasts. The general consensus from X-ray crystallographic studies and EXAFS is that the coordination geometry in the oxidized enzyme forms is distorted trigonal prismatic, with two coordinated MPTs, a terminal sulfido, and either a S-Cys or Se-Cys provided by the protein polypeptide [91, 232–235]. In spite of this apparent structural consensus, at least 5 different mechanisms have been proposed for formate dehydrogenase [236].

A spectroscopic and kinetic study of the cysteine-containing *R. eutropha* FdsABG [237] enzyme revealed the presence of a strongly coupled proton in the EPR spectrum of a Mo(V) enzyme form that is consistent with what has been observed in the selenocysteine containing *E. coli* FdhF formate dehydrogenase [238]. This proton was previously assigned as originating from a His residue that is relatively distant (6.1 Å) from the Mo ion. Using deuterated formate (^{2}H-COO^-) with partially reduced enzyme a Mo(V) EPR signal was observed that displayed increased proton coupling over time. This results from the Cα deuteron being solvent-exchangeable after transfer to the sulfido ligand, providing evidence for a direct hydride transfer mechanism that does not require coordination of the substrate to the Mo ion. A primary [^{2}H]-formate kinetic isotope effect of 2.1 was observed, consistent with the isotope sensitive step in the reductive half-reaction being partially rate-limiting. It has also been determined that an oxygen from solvent water is not incorporated into the CO_2 product during the reaction sequence [238], suggesting that formate oxidation is not an oxidative

hydroxylation similar to that observed in XO family enzymes. Collectively, the data discussed here point toward a mechanism that involves direct hydride transfer to the sulfido ligand leading to two-electron reduction of Mo and formation of the CO_2 product.

The five proposed mechanisms for formate dehydrogenase have recently been evaluated using large-scale QM and QM/MM approaches [236]. The computational results are consistent with recent experimental studies [237, 239] and suggest that the most plausible reaction mechanism involves direct hydride transfer. Here, formate is positioned in the second coordination sphere and does not bind as a ligand to Mo. The most favored mechanism involves the transfer of hydride from formate to the terminal sulfido ligand, leading to protonation of the sulfido ligand, reduction of Mo(VI) to Mo(IV), and the product potentially binding to the coordinated S-Cys as a thiocarbonate. The catalytically competent $[(MPT)_2MoS(S\text{-}Cys)]^{1-}$ site is then regenerated via two-electron oxidation of the Mo ion and release of the product. It was noted that a similar direct hydride transfer mechanism is also likely to occur in the W-containing formate dehydrogenases [237].

6. GENERAL CONCLUSIONS

We have presented here a brief overview of our current understanding regarding the interrelationships between structure, mechanism, and function for canonical members of the pyranopterin Mo and W enzyme families. This has included work that has advanced our understanding of the MPT and Moco biosynthetic pathways, and the potential roles of MPT in enzymatic catalysis. Oxygen atom transfer reactivity has been highlighted in terms of making and breaking the metal-oxo bond in enzymes that require dioxo (SO family), and monooxo (DMSO reductase Type III) oxidized active sites. For SO family and DMSO reductase Type III enzymes, the unique geometric structures of the active sites contribute to the different reactions that they catalyze.

Formal hydride transfer reactivity in XO family enzymes is complex and requires a unique $[MoOS]^{2+}$ oxidized active site for the insertion of an O atom from water into the substrate C–H bond. Remarkable electron delocalizations have been shown to minimize charge buildup along the XO reaction coordinate. A direct hydride transfer mechanism involving heterolytic substrate C–H bond scission and hydride transfer to the terminal sulfido is currently favored for DMSO reductase Type I formate dehydrogenases, and supported by structural, kinetic, and computational studies.

The future appears bright, and there is much left to be accomplished. Among the future fertile areas of study will be obtaining higher resolution crystal structures in order to obtain a more complete understanding of geometric structure contributions to catalysis. Researchers will develop new spectroscopic methods for studying the Mo and W centers of enzymes that possess other highly absorbing chromophores, and employ advanced computational, biochemical, molecular biology, theoretical, and synthetic strategies to obtain new and greater insight

into electronic and geometric structure contributions to catalysis. These enzymes include non-canonical members of the SO enzyme family, W enzymes, DMSO reductase Type II enzymes, and the nitrate reductases and formate dehydrogenases of the DMSO reductase family. We can expect to define the role of MPT in catalysis and determine if its potential for redox activity is utilized in the catalytic mechanisms of specific enzymes, in addition to understanding the final steps in Moco biosynthesis and maturation.

ACKNOWLEDGMENTS

M. L. K. would like to thank all of the collaborators, undergraduate students, graduate students, and postdoctoral associates who have contributed to the studies described in this chapter. M. L. K. also acknowledges the National Institutes of Health (GM-057378) for continued generous support of the author's work that is detailed in this review.

ABBREVIATIONS AND DEFINITIONS

AO	aldehyde oxidase
AMP	adenosine monophosphate
Bis-MGD	bis-molybdopterin guanine dinucleotide
CODH	carbon monoxide dehydrogenase
cPMP	cyclic pyranopterin monophosphate
DFT	density functional theory
DMSO	dimethylsulfoxide
DMSOR	diemthylsulfoxide reductase
EPR	electron paramagnetic resonance
EXAFS	extended x-ray absorption fine structure
FAD	flavin adenine dinucleotide
FDH	formate dehydrogenase
GTP	guanosine-5′-triphosphate
HMCS-CT	human molybdenum cofactor sulfurase C-terminal domain
LUMO	lowest unoccupied molecular orbital
mARC	mitochondrial amidoxime reducing component
MCD	magnetic circular dichroism
MLCT	metal → ligand charge transfer
Moco	molybdenum cofactor
MOSC	Moco sulfurase C-terminal
MPT	molybdopterin
MsrP	methionine sulfoxide reductase
$NAD(P)^+$	nicotinamide adenine dinucleotide (phosphate), oxidized
NAP	prokaryotic periplasmic nitrate reductase
NarGHI	respiratory nitrate reductase

NMR	nuclear magnetic resonance
OAT	oxygen atom transfer
PDT	pyranopterin dithiolene
PEt_3	triethylphosphine
PLP	pyridoxal-5′-phosphate
QM	quantum mechanics
QMMM	QM molecular mechanics
rR	resonance Raman
SO	sulfite oxidase
SOMO	singly occupied molecular orbital
TMA	trimethylamine
TMAO	trimethylamine oxide
XDH	xanthine dehydrogenase
XO	xanthine oxidase

REFERENCES

1. M. L. Kirk, *Spectroscopic and Electronic Structure Studies Probing Mechanism*, in *Molybdenum and Tungsten Enzymes: Spectroscopic and Theoretical Investigations*, Eds R. Hille, C. Schulzke, M. L. Kirk, The Royal Society of Chemistry, Cambridge, UK, **2016**, pp. 1–12.
2. M. L. Kirk, *Spectroscopic and Electronic Structure Studies of Mo Model Compounds and Enzymes*, in *Molybdenum and Tungsten Enzymes: Spectroscopic and Theoretical Investigations*, Eds R. Hille, C. Schulzke, M. L. Kirk, The Royal Society of Chemistry, Cambridge, UK, **2016**, pp. 13–67.
3. B. W. Stein, M. L. Kirk, *J. Biol. Inorg. Chem.* **2015**, *20*, 183–194.
4. M. L. Kirk, B. Stein, in *Comprehensive Inorganic Chemistry II,* 2nd ed., Eds J. Reedijk, K. R. Poeppelmeier, Elsevier, Amsterdam, **2013**, pp. 263–293.
5. R. Hille, J. Hall, P. Basu, *Chem. Rev.* **2014**, *114*, 3963–4038.
6. R. Hille, T. Nishino, F. Bittner, *Coord. Chem. Rev.* **2011**, *255*, 1179–1205.
7. R. Hille, *Arch. Biochem. Biophys.* **2005**, *433*, 107–116.
8. R. Hille, *Molybdenum Enzymes Containing the Ppyranopterin Cofactor: An overview*, Vol. 39 of *Metal Ions in Biological Systems*, Eds A. Sigel, H. Sigel, Marcel Dekker, Inc., New York, **2002**, pp. 187–226.
9. R. Hille, *Chem. Rev.* **1996**, *96*, 2757–2816.
10. L. B. Maia, I. Moura, J. J. G. Moura, *Molybdenum and Tungsten-Containing Enzymes: An Overview*, in *Molybdenum and Tungsten Enzymes: Biochemistry*, Eds R. Hille, C. Schulzke, M. L. Kirk, *RSC Metallobiology Series* No 5, Vol. 28, The Royal Society of Chemistry, Cambridge, UK, **2017**, pp. 1–80.
11. J. Pacheco, D. Niks, R. Hille, *J. Biol. Inorg. Chem.* **2018**, *23*, 295–301.
12. K. Okamoto, K. Matsumoto, R. Hille, B. T. Eger, E. F. Pai, T. Nishino, *Proc. Nat. Acad. Sci. USA* **2004**, *101*, 7931–7936.
13. J. Yang, L. J. Giles, C. Ruppelt, R. R. Mendel, F. Bittner, M. L. Kirk, *J. Am. Chem. Soc.* **2015**, *137*, 5276–5279.
14. L. B. Maia, J. J. G. Moura, *J. Biol. Inorg. Chem.* **2011**, *16*, 443–460.
15. C. E. Sparacino-Watkins, J. Tejero, B. Sun, M. C. Gauthier, J. Thomas, V. Ragireddy, B. A. Merchant, J. Wang, I. Azarov, P. Basu, M. T. Gladwin, *J. Biol. Chem.* **2014**, *289*, 10345–10358.

16. J. L. Li, U. Ryde, *Inorg. Chem.* **2014**, *53*, 11913–11924.
17. L. E. Bevers, P. L. Hagedoorn, W. R. Hagen, *Coord. Chem. Rev.* **2009**, *253*, 269–290.
18. W. R. Hagen, A. F. Arendsen, *Structure and Bonding* **1998**, *90*, 161–192.
19. D. Niks, R. Hille, *Prot. Sci.* **2018**, *28*, 111–122.
20. L. B. Maia, J. J. G. Moura, I. Moura, *J. Biol. Inorg. Chem.* **2015**, *20*, 287–309.
21. R. A. Rothery, B. Stein, M. Solomonson, M. L. Kirk, J. H. Weiner, *Proc. Natl. Acad. Sci. USA* **2012**, *109*, 14773–14778.
22. G. B. Seiffert, G. M. Ullmann, A. Messerschmidt, B. Schink, P. M. H. Kroneck, O. Einsle, *Proc. Natl. Acad. Sci. USA* **2007**, *104*, 3073–3077.
23. R. Hille, J. Retey, U. Bartlewski-Hof, W. Reichenbecher, B. Schink, *FEMS Microbiol. Rev.* **1998**, *22*, 489–501.
24. R. Hille, *J. Biol. Inorg. Chem.* **1996**, *1*, 397–404.
25. J. Reiss, *Hum. Genet.* **2000**, *106*, 157–163.
26. G. Schwarz, *Curr. Op. Chem. Biol.* **2016**, *31*, 179–187.
27. C. A. Pritsos, *Chem.-Biol. Interact.* **2000**, *129*, 195–208.
28. A. Havemeyer, J. A. Lang, B. Clement, *Drug Metab. Rev.* **2011**, *43*, 524–539.
29. A. Havemeyer, S. Grunewald, B. Wahl, F. Bittner, R. Mendel, P. Erdelyi, J. Fischer, B. Clement, *Drug Metab. Dispos.* **2010**, *38*, 1917–1921.
30. R. S. Obach, P. Huynh, M. C. Allen, C. Beedham, *J. Clin. Pharmacol.* **2004**, *44*, 7–19.
31. M. Rooseboom, J. N. M. Commandeur, N. P. E. Vermeulen, *Pharmacol. Rev.* **2004**, *56*, 53–102.
32. J. Kotthaus, D. Schade, U. Schwering, H. Hungeling, H. Muller-Fielitz, W. Raasch, B. Clement, *Chemmedchem* **2011**, *6*, 2233–2242.
33. A. Havemeyer, S. Grünewald, B. Wahl, F. Bittner, R. Mende, P. Erdélyi, J. Fischer, B. Clement, *Drug Metab. Dispos.* **2010**, *38*, 1917–1921.
34. S. Clarke, A. Harrell, R. Chenery, *Drug Metab. Dispos.* **1995**, *23*, 251–254.
35. M. Rashidi, J. Smith, S. Clarke, C. Beedham, *Drug Metab. Dispos.* **1997**, *25*, 805–813.
36. C. E. Berry, J. M. Hare, *J. Physiol.-London* **2004**, *555*, 589–606.
37. H. T. Li, A. Samouilov, X. P. Liu, J. L. Zweier, *J. Biol. Chem.* **2001**, *276*, 24482–24489.
38. B. Zhang, C. F. Hemann, R. Hille, *J. Biol. Chem.* **2010**, *285*, 12571–12578.
39. M. Gnida, R. Ferner, L. Gremer, O. Meyer, W. Meyer-Klaucke, *Biochemistry* **2003**, *42*, 222–230.
40. R. R. Mendel, T. Kruse, *Biochim. Biophys. Acta* **2012**, *1823*, 1568–1579.
41. W. H. Tan, F. S. Eichler, S. Hoda, M. S. Lee, H. Baris, C. A. Hanley, E. Grant, K. S. Krishnamoorthy, V. E. Shih, *Pediatrics* **2005**, *116*, 757–766.
42. R. Hansch, C. Lang, E. Riebeseel, R. Lindigkeit, A. Gessler, H. Rennenberg, R. R. Mendel, *J. Biol. Chem.* **2006**, *281*, 6884–6888.
43. D. Kaufholdt, C. K. Baillie, T. Wille, C. Lang, S. Hallier, C. Herschbach, H. Rennenberg, R. Mendel, R. Haensch, in *Molecular Physiology and Ecophysiology of Sulfur*, Eds L. J. DeKok, M. J. Hawkesford, H. Rennenberg, K. Saito, E. Schnug, Springer International Publishing, Cham, Switzerland, **2015**, pp. 179–187.
44. D. Randewig, D. Hamisch, C. Herschbach, M. Eiblmeier, C. Gehl, J. Jurgeleit, J. Skerra, R. R. Mendel, H. Rennenberg, R. Hansch, *Plant Cell Environ.* **2012**, *35*, 100–115.
45. C. Hemann, B. L. Hood, M. Fulton, R. Hansch, G. Schwarz, R. R. Mendel, M. L. Kirk, R. Hille, *J. Am. Chem. Soc.* **2005**, *127*, 16567–16577.
46. C. Juillan-Binard, A. Picciocchi, J. P. Andrieu, J. Dupuy, I. Petit-Hartlein, C. Caux-Thang, C. Vives, V. Niviere, F. Fieschi, *J. Biol. Chem.* **2017**, *292*, 2485–2494.
47. R. Dhouib, D. Othman, V. Lin, X. J. Lai, H. G. S. Wijesinghe, A. T. Essilfie, A. Davis, M. Nasreen, P. V. Bernhardt, P. M. Hansbro, A. G. McEwan, U. Kappler, *Frontiers in Microbiology* **2016**, *7*, e49638.

48. R. A. Melnyk, M. D. Youngblut, I. C. Clark, H. K. Carlson, K. M. Wetmore, M. N. Price, A. T. Iavarone, A. M. Deutschbauer, A. P. Arkin, J. D. Coates, *Mbio* **2015**, *6*, e00233–15.
49. A. Gennaris, B. Ezraty, C. Henry, R. Agrebi, A. Vergnes, E. Oheix, J. Bos, P. Leverrier, L. Espinosa, J. Szewczyk, D. Vertommen, O. Iranzo, J.-F. Collet, F. Barras, *Nature* **2015**, *528*, 409–412.
50. D. Fennema, I. R. Phillips, E. A. Shephard, *Drug Metab. Dispos.* **2016**, *44*, 1839–1850.
51. M. T. Velasquez, A. Ramezani, A. Manal, D. S. Raj, *Toxins* **2016**, *8*, 326.
52. R. R. Mendel, S. Leimkühler, *J. Biol. Inorg. Chem.* **2015**, *20*, 337–347.
53. J. A. Santamaria-Araujo, V. Wray, G. Schwarz, *J. Biol. Inorg. Chem.* **2012**, *17*, 113–122.
54. J. N. Daniels, M. M. Wuebbens, K. V. Rajagopalan, H. Schindelin, *Biochemistry* **2008**, *47*, 615–626.
55. M. M. Wuebbens, K. V. Rajagopalan, *J. Biol. Chem.* **2003**, *278*, 14523–14532.
56. J. Krausze, T. W. Hercher, D. Zwerschke, M. L. Kirk, W. Blankenfeldt, R. R. Mendel, T. Kruse, *Biochem. J.* **2018**, *475*, 1739–1753.
57. J. Krausze, C. Probst, U. Curth, J. Reichelt, S. Saha, D. Schafflick, D. W. Heinz, R. R. Mendel, T. Kruse, *Biochem. J.* **2017**, *474*, 163–178.
58. G. Schwarz, R. R. Mendel, *Ann. Rev. Plant Biol.* **2006**, *57*, 623–647.
59. M. Lehrke, S. Rump, T. Heidenreich, J. Wissing, R. R. Mendel, F. Bittner, *Biochem. J.* **2012**, *441*, 823–832.
60. V. Anantharaman, L. Aravind, *FEMS Microbiol. Lett.* **2002**, *207*, 55–61.
61. F. Bittner, M. Oreb, R. R. Mendel, *J. Biol. Chem.* **2001**, *276*, 40381–40384.
62. H. Peretz, M. S. Naamati, D. Levartovsky, A. Lagziel, E. Shani, I. Horn, H. Shalev, D. Landau, *Mol. Gen. Metabol.* **2007**, *91*, 23–29.
63. S. Wollers, T. Heidenreich, M. Zarepour, D. Zachmann, C. Kraft, Y. D. Zhao, R. R. Mendel, F. Bittner, *J. Biol. Chem.* **2008**, *283*, 9642–9650.
64. T. Heidenreich, S. Wollers, R. R. Mendel, F. Bittner, *J. Biol. Chem.* **2005**, *280*, 4213–4218.
65. R. R. Mendel, R. Hansch, *J. Exp. Bot.* **2002**, *53*, 1689–1698.
66. L. J. Giles, C. Ruppelt, J. Yang, R. R. Mendel, F. Bittner, M. L. Kirk, *Inorg. Chem.* **2014**, *53*, 9460–9462.
67. N. Schrader, K. Fischer, K. Theis, R. R. Mendel, G. Schwarz, C. Kisker, *Structure* **2003**, *11*, 1251–1263.
68. R. Hille, *Structure* **2003**, *11*, 1189–1190.
69. K. V. Rajagopalan, *J. Biol. Inorg. Chem.* **1997**, *2*, 786–789.
70. C. Kisker, H. Schindelin, D. C. Rees, *Annu. Rev. Biochem.* **1997**, *66*, 233–267.
71. J. A. Qiu, H. L. Wilson, M. J. Pushie, C. Kisker, G. N. George, K. V. Rajagopalan, *Biochemistry* **2010**, *49*, 3989–4000.
72. E. Karakas, H. L. Wilson, T. N. Graf, S. Xiang, S. Jaramillo-Busquets, K. V. Rajagopalan, C. Kisker, *J. Biol. Chem.* **2005**, *280*, 33506–33515.
73. H. Schindelin, C. Kisker, D. C. Rees, *J. Biol. Inorg. Chem.* **1997**, *2*, 773–781.
74. C. Kisker, H. Schindelin, A. Pacheco, W. A. Wehbi, R. M. Garrett, K. V. Rajagopalan, J. H. Enemark, D. C. Rees, *Cell* **1997**, *91*, 973–983.
75. C. Kisker, H. Schindelin, D. C. Rees, *Annu. Rev. Biochem.* **1997**, *66*, 233–267.
76. J. Truglio, K. Theis, S. Leimkuhler, R. Rappa, K. Rajagopalan, C. Kisker, *Structure* **2002**, *10*, 115–125.
77. K. Okamoto, Y. Kawaguchi, B. T. Eger, E. F. Pai, T. Nishino, *J. Am. Chem. Soc.* **2010**, *132*, 17080–17083.
78. J. M. Pauff, H. Cao, R. Hille, *J. Biol. Chem.* **2009**, *284*, 8751–8758.
79. J. M. Pauff, J. J. Zhang, C. E. Bell, R. Hille, *J. Biol. Chem.* **2008**, *283*, 4818–4824.

80. C. Enroth, B. T. Eger, K. Okamoto, T. Nishino, T. Nishino, E. F. Pai, *Proc. Natl. Acad. Sci. USA* **2000**, *97*, 10723–10728.
81. H. N. Cao, J. Hall, R. Hille, *Biochemistry* **2014**, *53*, 533–541.
82. U. Dietzel, J. Kuper, J. A. Doebbler, A. Schulte, J. J. Truglio, S. Leimkuhler, C. Kisker, *J. Biol. Chem.* **2009**, *284*, 8759–8767.
83. C. Enroth, B. Eger, K. Okamoto, T. Nishino, T. Nishino, E. Pai, *Proc. Nat. Acad. Sci. USA* **2000**, *97*, 10723–10728.
84. S. Leimkühler, R. Hodson, G. N. George, K. V. Rajagopalan, *J. Biol. Chem.* **2003**, *278*, 20802–20811.
85. P. J. Ellis, T. Conrads, R. Hille, P. Kuhn, *Structure* **2001**, *9*, 125–132.
86. A. McAlpine, A. McEwan, S. Bailey, *J. Mol. Biol.* **1998**, *275*, 613–623.
87. M. Czjzek, J.-P. D. Santos, J. Pomier, G. Giordano, V. Méjean, R. Haser, *J. Mol. Biol.* **1998**, *284*, 435–447.
88. A. McAlpine, A. McEwan, A. Shaw, S. Bailey, *J. Biol. Inorg. Chem.* **1997**, *2*, 690–701.
89. F. Schneider, J. Lowe, R. Huber, H. Schindelin, C. Kisker, J. Knablein, *J. Mol. Biol.* **1996**, *263*, 53–69.
90. H. Schindelin, C. Kisker, J. Hilton, K. Rajagopalan, D. Rees, *Science* **1996**, *272*, 1615–1621.
91. H. Raaijmakers, S. Macieira, J. M. Dias, S. Teixeira, S. Bursakov, R. Huber, J. J. G. Moura, I. Moura, M. J. Romão, *Structure* **2002**, *10*, 1261–1272.
92. Y. Hu, S. Faham, R. Roy, M. W. W. Adams, D. C. Rees, *J. Mol. Biol.* **1999**, *286*, 899–914.
93. M. G. Bertero, R. A. Rothery, M. Palak, C. Hou, D. Lim, F. Blasco, J. H. Weiner, N. C. J. Strynadka, *Nature Structural Biology* **2003**, *10*, 681–687.
94. D. P. Kloer, C. Hagel, J. Heider, G. E. Schulz, *Structure* **2006**, *14*, 1377–1388.
95. M. Jormakka, D. Richardson, B. Byrne, S. Iwata, *Structure* **2004**, *12*, 95–104.
96. R. A. Rothery, J. H. Weiner, *J. Biol. Inorg. Chem.* **2015**, *20*, 349–372.
97. S. P. Greatbanks, I. H. Hillier, C. D. Garner, J. A. Joule, *J. Chem. Soc. Perkin Trans.2* **1997**, 1529–1534.
98. J. H. Enemark, C. D. Garner, *J. Biol. Inorg. Chem.* **1997**, *2*, 817–822.
99. S. J. N. Burgmayer, D. L. Pearsall, S. M. Blaney, E. M. Moore, C. Sauk-Schubert, *J. Biol. Inorg. Chem.* **2004**, *9*, 59–66.
100. B. L. Westcott, N. E. Gruhn, J. H. Enemark, *J. Am. Chem. Soc.* **1998**, *120*, 3382–3386.
101. F. E. Inscore, S. Z. Knottenbelt, N. D. Rubie, H. K. Joshi, M. L. Kirk, J. H. Enemark, *Inorg. Chem.* **2006**, *45*, 967–976.
102. P. Basu, S. J. N. Burgmayer, *Coord. Chem. Rev.* **2011**, *255*, 1016–1038.
103. K. G. Matz, R. P. Mtei, R. Rothstein, M. L. Kirk, S. J. N. Burgmayer, *Inorg. Chem.* **2011**, *50*, 9804–9815.
104. K. G. Matz, R. P. Mtei, B. Leung, S. J. N. Burgmayer, M. L. Kirk, *J. Am. Chem. Soc.* **2010**, *132*, 7830–7831.
105. S. J. N. Burgmayer, M. Kim, R. Petit, A. Rothkopf, A. Kim, S. BelHamdounia, Y. Hou, A. Somogyi, D. Habel-Rodriguez, A. Williams, M. L. Kirk, *J. Inorg. Biochem.* **2007**, *101*, 1601–1616.
106. B. W. Stein, J. Yang, R. Mtei, N. J. Wiebelhaus, D. K. Kersi, J. LePluart, D. L. Lichtenberger, J. H. Enemark, M. L. Kirk, *J. Am. Chem. Soc.* **2018**, *140*, 14777–14788.
107. R. M. Jones, F. E. Inscore, R. Hille, M. L. Kirk, *Inorg. Chem.* **1999**, *38*, 4963–4970.
108. R. Hille, R. F. Anderson, *J. Biol. Chem.* **1991**, *266*, 5608–5615.
109. H. B. Gray, J. R. Winkler, *Proc. Natl. Acad. Sci. USA* **2005**, *102*, 3534–3539.
110. M. L. Kirk, D. A. Shultz, D. E. Stasiw, G. F. Lewis, G. B. Wang, C. L. Brannen, R. D. Sommer, P. D. Boyle, *J. Am. Chem. Soc.* **2013**, *135*, 17144–17154.
111. R. Hille, *Biochemistry* **1991**, *30*, 8522–8529.
112. R. Hille, V. Massey, *J. Biol. Chem.* **1981**, *256*, 9090–9095.

113. C. Dong, J. Yang, S. Leimkühler, M. L. Kirk, *Inorg. Chem.* **2014**, *53*, 7077–7079.
114. H. Adamson, A. N. Simonov, M. Kierzek, R. A. Rothery, J. H. Weiner, A. M. Bond, A. Parkin, *Proc. Natl. Acad. Sci. USA* **2015**, *112*, 14506–14511.
115. D. R. Gisewhite, J. Yang, B. R. Williams, A. Esmail, B. Stein, M. L. Kirk, S. J. N. Burgmayer, *J. Am. Chem. Soc.* **2018**, *140*, 12808–12818.
116. R. H. Holm, *Chem. Rev.* **1987**, *87*, 1401–1449.
117. R. H. Holm, J. P. Donahue, *Polyhedron* **1993**, *12*, 571–589.
118. S. C. Lee, R. H. Holm, *Inorg. Chim. Acta* **2008**, *361*, 1166–1176.
119. R. H. Holm, E. I. Solomon, A. Majumdar, A. Tenderholt, *Coord. Chem. Rev.* **2011**, *255*, 993–1015.
120. A. V. Astashkin, E. L. Klein, D. Ganyushin, K. Johnson-Winters, F. Neese, U. Kappler, J. H. Enemark, *Phys. Chem. Chem. Phys.* **2009**, *11*, 6733–6742.
121. G. J. Workun, K. Moquin, R. A. Rothery, J. H. Weiner, *Microbiol. Mol. Biol. Rev.* **2008**, *72*, 228–248.
122. C. Feng, G. Tollin, J. H. Enemark, *Biochim. Biophys. Acta* **2007**, *1774*, 527–539.
123. C. Wodara, F. Bardischewsky, C. G. Friedrich, *J. Bacteriol.* **1997**, *179*, 5014–5023.
124. C. Sparacino-Watkins, J. F. Stolz, P. Basu, *Chem. Soc. Rev.* **2014**, *43*, 676–706.
125. N. A. Eckardt, *Plant Cell* **2005**, *17*, 1029–1031.
126. R. R. Mendel, G. Schwarz, *Critical Reviews in Plant Sciences* **1999**, *18*, 33–69.
127. A. V. Astashkin, K. Johnson-Winters, E. L. Klein, C. J. Feng, H. L. Wilson, K. V. Rajagopalan, A. K. Raitsimring, J. H. Enemark, *J. Am. Chem. Soc.* **2008**, *130*, 8471–8480.
128. C. J. Doonan, H. L. Wilson, K. V. Rajagopalan, R. M. Garrett, B. Bennett, R. C. Prince, G. N. George, *J. Am. Chem. Soc.* **2007**, *129*, 9421–9428.
129. J. H. Enemark, M. M. Cosper, *Met. Ions Biol. Syst.* **2002**, *39*, 621–654.
130. S. J. Brokx, R. A. Rothery, G. J. Zhang, D. P. Ng, J. H. Weiner, *Biochemistry* **2005**, *44*, 10339–10348.
131. K. G. V. Havelius, S. Reschke, S. Horn, A. Döring, D. Niks, R. Hille, C. Schulzke, S. Leimkühler, M. Haumann, *Inorg. Chem.* **2011**, *50*, 741–748.
132. L. Loschi, S. J. Brokx, T. L. Hills, G. Zhang, M. G. Bertero, A. L. Lovering, J. H. Weiner, N. C. J. Strynadka, *J. Biol. Chem.* **2004**, *279*, 50391–50400.
133. M. J. Pushie, C. J. Doonan, K. Moquin, J. H. Weiner, R. Rothery, G. N. George, *Inorg. Chem.* **2011**, *50*, 732–740.
134. J. Yang, R. Rothery, J. Sempombe, J. H. Weiner, M. L. Kirk, *J. Am. Chem. Soc.* **2009**, *131*, 15612–15614.
135. G. Ott, A. Havemeyer, B. Clement, *J. Biol. Inorg. Chem.* **2015**, *20*, 265–275.
136. N. Krompholz, C. Krischkowski, D. Reichmann, D. Garbe-Schonberg, R. R. Mendel, F. Bittner, B. Clement, A. Havemeyer, *Chem. Res. Toxicol.* **2012**, *25*, 2443–2450.
137. A. Rajapakshe, A. V. Astashkin, E. L. Klein, D. Reichmann, R. R. Mendel, F. Bittner, J. H. Enemark, *Biochemistry* **2011**, *50*, 8813–8822.
138. B. Wahl, D. Reichmann, D. Niks, N. Krompholz, A. Havemeyer, B. Clement, T. Messerschmidt, M. Rothkegel, H. Biester, R. Hille, R. R. Mendel, F. Bittner, *J. Biol. Chem.* **2010**, *285*, 37847–37859.
139. C. Kubitza, F. Bittner, C. Ginsel, A. Havemeyer, B. Clement, A. J. Scheidig, *Proc. Nat. Acad. Sci.* **2018**, *115*, 11958–11963.
140. S. Gruenewald, B. Wahl, F. Bittner, H. Hungeling, S. Kanzow, J. Kotthaus, U. Schwering, R. R. Mendel, B. Clement, *J. Med. Chem.* **2008**, *51*, 8173–8177.
141. S. Reschke, D. Niks, H. Wilson, K. G. V. Sigfridsson, M. Haumann, K. V. Rajagopalan, R. Hille, S. Leimkühler, *Biochemistry* **2013**, *52*, 8295–8303.
142. A. Rajapakshe, K. Johnson-Winters, A. R. Nordstrom, K. T. Meyers, S. Emesh, A. V. Astashkin, J. H. Enemark, *Biochemistry* **2010**, *49*, 5154–5159.

143. A. M. Raitsimring, A. V. Astashkin, C. Feng, H. L. Wilson, K. V. Rajagopalan, J. H. Enemark, *Inorg. Chim. Acta* **2008**, *361*, 941–946.
144. U. Kappler, S. Bailey, *J. Biol. Chem.* **2005**, *280*, 24999–25007.
145. U. Kappler, S. Bailey, C. J. Feng, M. J. Honeychurch, G. R. Hanson, P. V. Bernhardt, G. Tollin, J. H. Enemark, *Biochemistry* **2006**, *45*, 9696–9705.
146. H. H. Harris, G. N. George, K. V. Rajagopalan, *Inorg. Chem.* **2006**, *45*, 493–495.
147. G. N. George, R. M. Garrett, R. C. Prince, K. V. Rajagopalan, *Inorg. Chem.* **2004**, *43*, 8456–8460.
148. E. L. Klein, A. V. Astashkin, D. Ganyushin, C. Riplinger, K. Johnson-Winters, F. Neese, J. H. Enemark, *Inorg. Chem.* **2009**, *48*, 4743–4752.
149. A. V. Astashkin, E. L. Klein, J. H. Enemark, *J. Inorg. Biochem.* **2007**, *101*, 1623–1629.
150. A. V. Astashkin, K. Johnson-Winters, E. L. Klein, R. S. Byrne, R. Hille, A. M. Raitsimring, J. H. Enemark, *J. Am. Chem. Soc.* **2007**, *129*, 14800–14810.
151. M. L. Kirk, in *Paramagnetic Resonance of Metallobiomolecules*, Ed. J. Telser, *ACS Symposium Series, Vol. 858*, Washington, DC, **2003**, pp. 340–357.
152. M. E. Helton, A. Pacheco, J. McMaster, J. H. Enemark, M. L. Kirk, *J. Inorg. Biochem.* **2000**, *80*, 227–233.
153. P. D. Smith, A. J. Millar, C. G. Young, A. Ghosh, P. Basu, *J. Am. Chem. Soc.* **2000**, *122*, 9298–9299.
154. A. K. Rappé, W. A. Goddard, *Nature* **1980**, *285*, 311–312.
155. M. A. Pietsch, M. B. Hall, *Inorg. Chem.* **1996**, *35*, 1273–1278.
156. L. M. Thomson, M. B. Hall, *J. Am. Chem. Soc.* **2001**, *123*, 3995–4002.
157. M. S. Brody, R. Hille, *Biochemistry* **1999**, *38*, 6668–6677.
158. O. Caldararu, M. Feldt, D. Cioloboc, M. C. van Severen, K. Starke, R. A. Mata, E. Nordlander, U. Ryde, *Sci. Rep.* **2018**, *8(1)*: 4684.
159. K. Johnson-Winters, A. R. Nordstrom, S. Emesh, A. V. Astashkin, A. Rajapakshe, R. E. Berry, G. Tollin, J. H. Enemark, *Biochemistry* **2010**, *49*, 1290–1296.
160. A. Rajapakshe, K. T. Meyers, R. E. Berry, G. Tollin, J. H. Enemark, *J. Biol. Inorg. Chem.* **2012**, *17*, 345–352.
161. Y. Izumi, T. Glaser, K. Rose, J. McMaster, P. Basu, J. H. Enemark, B. Hedman, K. O. Hodgson, E. I. Solomon, *J. Am. Chem. Soc.* **1999**, *121*, 10035–10046.
162. K. Peariso, R. L. McNaughton, M. L. Kirk, *J. Am. Chem. Soc.* **2002**, *124*, 9006–9007.
163. R. L. McNaughton, A. A. Tipton, N. D. Rubie, R. R. Conry, M. L. Kirk, *Inorg. Chem.* **2000**, *39*, 5697–5706.
164. K. Peariso, M. E. Helton, E. N. Duesler, S. E. Shadle, M. L. Kirk, *Inorg. Chem.* **2007**, *46*, 1259–1267.
165. J. Paudel, A. Pokhrel, M. L. Kirk, F. Li, *Inorg. Chem.* **2019**, *58*, 2054–2068.
166. M. G. Evans, M. Polanyi, *Trans. Farad. Soc.* **1938**, *34*, 11–23.
167. T. Conrads, C. Hemann, G. N. George, I. J. Pickering, R. C. Prince, R. Hille, *J. Am. Chem. Soc.* **2002**, *124*, 11276–11277.
168. M. J. Pushie, G. N. George, *Coord. Chem. Rev.* **2011**, *255*, 1055–1084.
169. H.-K. Li, C. Temple, K. V. Rajagopalan, H. Schindelin, *J. Am. Chem. Soc.* **2000**, *122*, 7673–7680.
170. G. N. George, J. Hilton, C. Temple, R. C. Prince, K. V. Rajagopalan, *J. Am. Chem. Soc.* **1999**, *121*, 1256–1266.
171. N. Cobb, T. Conrads, R. Hille, *J. Biol. Chem.* **2005**, *280*, 11007–11017.
172. R. P. Mtei, G. Lyashenko, B. Stein, N. Rubie, R. Hille, M. L. Kirk, *J. Am. Chem. Soc.* **2011**, *133*, 9762–9774.
173. M. J. Pushie, J. J. H. Cotelesage, G. Lyashenko, R. Hille, G. N. George, *Inorg. Chem.* **2013**, *52*, 2830–2837.
174. A. Thapper, R. J. Deeth, E. Nordlander, *Inorg. Chem.* **2002**, *41*, 6695–6702.

175. J. P. McNamara, I. H. Hillier, T. S. Bhachu, C. D. Garner, *Dalton Trans.* **2005**, 3572–3579.
176. M. Hofmann, *J. Mol. Struct. Theochem.* **2006**, *773*, 59–70.
177. M. Hofmann, *Inorg. Chem.* **2008**, *47*, 5546–5548.
178. R. L. McNaughton, B. S. Lim, S. Z. Knottenbelt, R. H. Holm, M. L. Kirk, *J. Am. Chem. Soc.* **2008**, *130*, 4628–4636.
179. E. Hernandez-Marin, T. Ziegler, *Can. J. Chem.-Rev. Can. Chimie* **2010**, *88*, 683–693.
180. A. L. Tenderholt, J. J. Wang, R. K. Szilagyi, R. H. Holm, K. O. Hodgson, B. Hedman, E. I. Solomon, *J. Am. Chem. Soc.* **2010**, *132*, 8359–8371.
181. A. L. Tenderholt, K. O. Hodgson, B. Hedman, R. H. Holm, E. I. Solomon, *Inorg. Chem.* **2012**, *51*, 3436–3442.
182. G. Dong, U. Ryde, *J. Inorg. Biochem.* **2017**, *171*, 45–51.
183. J. L. Li, M. Andrejic, R. A. Mata, U. Ryde, *Eur. J. Inorg. Chem.* **2015**, *2015*, 3580–3589.
184. J. L. Li, R. A. Mata, U. Ryde, *J. Chem. Theory Comput.* **2013**, *9*, 1799–1807.
185. C. E. Webster, M. B. Hall, *J. Am. Chem. Soc.* **2001**, *123*, 5820–5821.
186. R. Huber, P. Hof, R. Duarte, J. Moura, I. Moura, M. Liu, J. LeGall, R. Hille, M. Archer, M. Romão, *Proc. Natl. Acad. Sci. USA* **1996**, *93*, 8846–8851.
187. D. R. Boer, A. Thapper, C. D. Brondino, M. J. Romão, J. J. G. Moura, *J. Am. Chem. Soc.* **2004**, *126*, 8614–8615.
188. M. Romão, M. Archer, I. Moura, J. Moura, J. LeGall, R. Engh, M. Schneider, P. Hof, R. Huber, *Science* **1995**, *270*, 1170–1176.
189. M. J. Romão, *Dalton Trans.* **2009**, 4053–4068.
190. C. J. Doonan, A. Stockert, R. Hille, G. N. George, *J. Am. Chem. Soc.* **2005**, *127*, 4518–4522.
191. J. Yang, C. Dong, M. L. Kirk, *Dalton Trans.* **2017**, *46*, 13242–13250.
192. E. Y. Choi, A. L. Stockert, S. Leimkühler, R. Hille, *J. Inorg. Biochem.* **2004**, *98*, 841–848.
193. S. Grimaldi, F. Biaso, B. Burlat, B. Guigliarelli, Electron Paramagnetic Resonance Studies of Molybdenum Enzymes, in *Molybdenum and Tungsten Enzymes: Spectroscopic and Theoretical Investigations*, Eds R. Hille, C. Schulzke, M. L. Kirk, The Royal Society of Chemistry, Cambridge, UK, **2016**, pp. 68–120.
194. J. Swann, T. D. Westmoreland, *Inorg. Chem.* **1997**, *36*, 5348–5357.
195. C. Balagopalakrishna, J. T. Kimbrough, T. D. Westmoreland, *Inorg. Chem.* **1996**, *35*, 7758–7768.
196. V. N. Gladyshev, S. V. Khangulov, T. C. Stadtman, *Proc. Nat. Acad. Sci.* **1994**, *91*, 232–236.
197. F. E. Mabbs, D. Collison, *Electron Paramagnetic Resonance of d Transition Metal Compounds*, Elsevier, Amsterdam, **1992**.
198. A. J. Stone, *Proc. Roy. Soc. London, Math. Phys. Sci.* **1963**, *271*, 424–434.
199. G. George, R. Bray, *Biochemistry* **1988**, *27*, 3603–3609.
200. G. L. Wilson, R. J. Greenwood, J. R. Pilbrow, J. T. Spence, A. G. Wedd, *J. Am. Chem. Soc.* **1991**, *113*, 6803–6812.
201. J. Malthouse, G. George, D. Lowe, R. Bray, *Biochem J.* **1981**, *199*, 629–637.
202. B. D. Howes, R. C. Bray, R. L. Richards, N. A. Turner, B. Bennett, D. J. Lowe, *Biochemistry* **1996**, *35*, 1432–1443.
203. S. Gutteridge, R. C. Bray, *Biochem. J.* **1980**, *189*, 615–623.
204. P. Manikandan, E. Choi, R. Hille, B. Hoffman, *J. Am. Chem. Soc.* **2001**, *123*, 2658–2663.
205. M. Xia, R. Dempski, R. Hille, *J. Biol. Chem.* **1999**, *274*, 3323–3330.
206. J. Sempombe, B. Stein, M. L. Kirk, *Inorg. Chem.* **2011**, *50*, 10919–10928.
207. M. R. Bray, R. J. Deeth, *J. Chem. Soc. Dalton Trans.* **1997**, 4005–4010.

208. G. A. Lorigan, R. D. Britt, J. H. Kim, R. Hille, *Biochim. Biophys. Acta* **1994**, *1185*, 284–294.
209. M. Shanmugam, B. Zhang, R. L. McNaughton, R. A. Kinney, R. Hille, B. M. Hoffman, *J. Am. Chem. Soc.* **2010**, *132*, 14015–14017.
210. R. Hille, *Biochim. Biophys. Acta* **1994**, *1184*, 143–169.
211. C. Hemann, P. Ilich, A. L. Stockert, E. Y. Choi, R. Hille, *J. Phys. Chem. B* **2005**, *109*, 3023–3031.
212. C. Hemann, P. Ilich, R. Hille, *J. Phys. Chem. B* **2003**, *107*, 2139–2155.
213. M. Davis, J. Olson, G. Palmer, *J. Biol. Chem.* **1984**, *259*, 3526–3533.
214. C. Dong, J. Yang, S. Reschke, S. Leimkühler, M. L. Kirk, *Inorg. Chem.* **2017**, *56*, 6830–6837.
215. N. C. Maiti, T. Tomita, T. Kitagawa, K. Okamoto, T. Nishino, *J. Biol. Inorg. Chem.* **2003**, *8*, 327–333.
216. C. J. Doonan, N. D. Rubie, K. Peariso, H. H. Harris, S. Z. Knottenbelt, G. N. George, C. G. Young, M. L. Kirk, *J. Am. Chem. Soc.* **2008**, *130*, 55–65.
217. S. Metz, W. Thiel, *Coord. Chem. Rev.* **2011**, *255*, 1085–1103.
218. J. M. Dieterich, H. J. Werner, R. A. Mata, S. Metz, W. Thiel, *J. Chem. Phys.* **2010**, *132*, 035101; https://doi.org/10.1063/1.3280164.
219. A. Voityuk, K. Albert, M. Romão, R. Huber, N. Rosch, *Inorg. Chem.* **1998**, *37*, 176–180.
220. S. Metz, W. Thiel, *J. Phys. Chem. B* **2010**, *114*, 1506–1517.
221. S. Metz, D. Q. Wang, W. Thiel, *J. Am. Chem. Soc.* **2009**, *131*, 4628–4640.
222. S. Metz, W. Thiel, *J. Am. Chem. Soc.* **2009**, *131*, 14885–14902.
223. M. L. Kirk, S. Knottenbelt, A. Habtegabre, in *Computational Inorganic and Bioinorganic Chemistry*, Eds E. I. Solomon, R. A. Scott, B. R. King, John Wiley & Sons, Hoboken, **2009**, pp. 614.
224. X. H. Zhang, Y. D. Wu, *Inorg. Chem.* **2005**, *44*, 1466–1471.
225. S. D'Ardenne, D. Edmondson, *Biochemistry* **1990**, *29*, 9046–9052.
226. J. T. Groves, *J. Inorg. Biochem.* **2006**, *100*, 434–447.
227. K. Bagramyan, A. Trchounian, *Biochemistry-Moscow* **2003**, *68*, 1159–1170.
228. M. Jormakka, B. Byrne, S. Iwata, *Curr. Opin. Struct. Biol.* **2003**, *13*, 418–423.
229. T. Reda, C. M. Plugge, N. J. Abram, J. Hirst, *Proc. Natl. Acad. Sci. USA* **2008**, *105*, 10654–10658.
230. U. Ruschig, U. Muller, P. Willnow, T. Hopner, *Eur. J. Biochem.* **1976**, *70*, 325–330.
231. A. Alissandratos, H. K. Kim, H. Matthews, J. E. Hennessy, A. Philbrook, C. J. Easton, *Appl. Environ. Microbiol.* **2013**, *79*, 741–744.
232. J. C. Boyington, V. N. Gladyshev, S. V. Khangulov, T. C. Stadtman, P. D. Sun, *Science* **1997**, *275*, 1305–1308.
233. H. C. A. Raaijmakers, M. J. Romão, *J. Biol. Inorg. Chem.* **2006**, *11*, 849–854.
234. M. Jormakka, S. Tornroth, B. Byrne, S. Iwata, *Science* **2002**, *295*, 1863–1868.
235. P. Schrapers, T. Hartmann, R. Kositzki, H. Dau, S. Reschke, C. Schulzke, S. Leimkühler, M. Haumann, *Inorg. Chem.* **2015**, *54*, 3260–3271.
236. G. Dong, U. Ryde, *J. Biol. Inorg. Chem.* **2018**, *23*, 221–229.
237. D. Niks, J. Duvvuru, M. Escalona, R. Hille, *J. Biol. Chem.* **2016**, *291*, 1162–1174.
238. S. V. Khangulov, V. N. Gladyshev, G. C. Dismukes, T. C. Stadtman, *Biochemistry* **1998**, *37*, 3518–3528.
239. L. B. Maia, L. Fonseca, I. Moura, J. J. G. Moura, *J. Am. Chem. Soc.* **2016**, *138*, 8834–8846.

10

The Siroheme-[4Fe-4S] Coupled Center

Isabel Askenasy [1,2] ***and M. Elizabeth Stroupe*** [1]

[1]Department of Biological Science and Institute of Molecular Biophysics, 91 Chieftain Way, Tallahassee, FL 32306
<mestroupe@bio.fsu.edu>

[2]Current Address: Department of Biomolecular Chemistry, University of Wisconsin School of Medicine and Public Health, 440 Henry Mall, Biochemical Sciences Building, Room 4206C, Madison, WI 53706, USA

Metal Ions in Life Sciences, Volume 20 Guest Editors: Martha E. Sosa Torres and Peter M. H. Kroneck
Series Editors: Astrid Sigel, Eva Freisinger, and Roland K. O. Sigel

https://doi.org/10.1515/9783110589757-010

Abstract: In nature, sulfur exists in a range of oxidation states and the two-electron reduced form is the most commonly found in biomolecules like the sulfur-containing amino acids cysteine and methionine, some cofactors, and polysaccharides. Sulfur is reduced through two pathways: dissimilation, where sulfite (SO_3^{2-}) is used as terminal electron acceptor; and assimilation, where sulfite is reduced to sulfide (S^{2-}) for incorporation into biomass. The pathways are independent, but share the sulfite reductase function, in which a single enzyme reduces sulfite by six electrons to make sulfide. With few exceptions, sulfite reductases from either pathway are iron metalloenzymes with structurally diverse configurations that range from monomers to tetramers. The hallmark of sulfite reductase is its catalytic center made of an iron-containing porphyrinoid called siroheme that is covalently coupled to a [4Fe-4S] cluster through a shared cysteine ligand. The substrate evolves through a push-pull mechanism, where electron transfer is coupled to three dehydration steps. Siroheme is an isobacteriochlorin that is more readily oxidized than protoporphyin IX-derived hemes. It is synthesized from uroporphyrinogen III in three steps (methylation, a dehydrogenation, and ferrochelation) that are performed by enzymes with homology to those involved in cobalamin synthesis. Future research will need to address how the siroheme-[4Fe-4S] clusters are assembled into apo-sulfite and nitrite reductases. The chapter will discuss how environmental microbes use sulfite reductase to survive in a range of ecosystems;

how atomic-resolution structures of dissimilatory and assimilatory sulfite reductases reveal their ancient homology; how the siroheme-[4Fe-4S] cluster active site catalyzes the six-electron reduction of sulfite to sulfide; and how siroheme is synthesized across diverse microrganisms.

Keywords: assimilatory sulfur reduction · dissimilatory sulfur reduction · iron-sulfur cluster · siroheme · sulfite reductase

1. INTRODUCTION: THE BIOGEOCHEMICAL SULFUR CYCLE

1.1. The Varied Redox States of Sulfur

Sulfur sits in a special place in the periodic table with a flexible electron distribution ($1s^2 2s^2 2p^6 3s^2 3p^4$) that allows it to access a wide range of oxidation states,

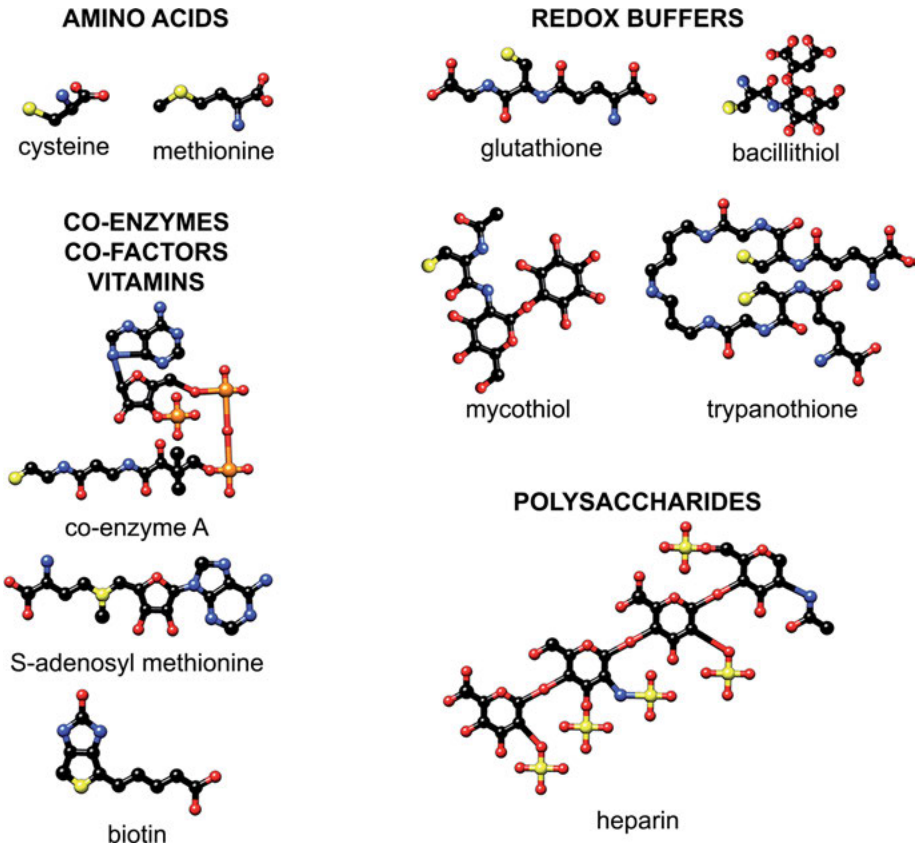


Figure 1. Sulfur (yellow) containing biomolecules fall into several broad categories including amino acids, redox buffers, cofactors/coenzyme/vitamins, and polysaccharides.

from S^{VI}, for example as in sulfate (SO_4^{2-}), to S^{IV}, for example as in sulfite (SO_3^{2-}), to S^{-II}, for example as in sulfide (S^{2-}). The two-electron reduced form of sulfur is found in biomolecules like the sulfur-containing amino acids cysteine and methionine as well as in cofactors and polysaccharides like *S*-adenosyl-L-methionine (SAM), coenzyme A, glutathione, bacillithiol, trypanothione, mycothiol, biotin, or heparin (Figure 1). H_2S is also a potent signaling molecule in mammals, playing a role in inflammation and central nervous system signaling [1]. In plants, H_2S plays a role in seedling development from germination, through root morphogenesis, basic photosynthesis, and flower senescence (see Chapter 2) [2].

Environmental sulfur is also found in multiple redox states. Oxidized sulfur is fairly non-reactive because it forms stable anions like SO_4^{2-} with a full second valence shell. Partially reduced sulfur-containing anions, including SO_3^{2-}, thiosulfate ($S_2O_3^{2-}$), and trithionate ($S_3O_6^{2-}$), are environmental hazards because they can disrupt heavy metal absorption by inhibiting enzymes that can bioremediate such pollutants [3, 4]. Sulfur dioxide (SO_2), also S^{IV}, is a byproduct of coal burning; sometimes used as a pesticide; and the cause of acid rain [5–8]. S^{2-} is the bioavailable form of sulfur but also is an environmental pollutant as H_2S, the compound that smells like rotten eggs (see Chapter 2).

1.2. The Cycling of Sulfur Redox States

Biological sulfur reduction occurs in two independent, but related, pathways (Figure 2). First, sulfur reduction can serve as the terminal electron transfer event during sulfur-based anaerobic respiration, driving oxidative phosphorylation to build up the proton gradient that allows chemiosmosis-driven ATP synthesis. Respiratory sulfur reduction is called "dissimilatory" sulfur reduction. Second, sulfur reduction can be performed to build biomass. This form of sulfur reduction is called "assimilatory" sulfur reduction.

Proteins regulate the cycling of sulfur through its biogeochemical cycle (Figure 3). Sulfur is imported into sulfur-reducing cells as SO_4^{2-} or $S_2O_3^{2-}$ via a sulfate

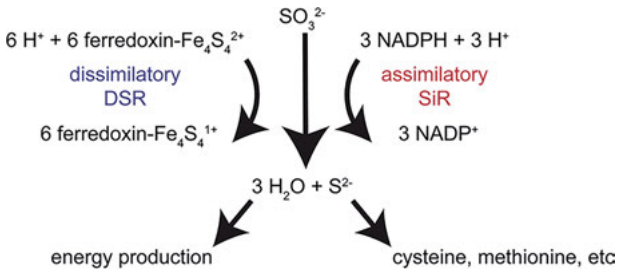


Figure 2. Two parallel pathways exist for sulfur reduction across biology, reducing sulfite by six electrons for energy conservation (dissimilatory sulfur reduction) or biomass formation (assimilatory sulfur reduction).

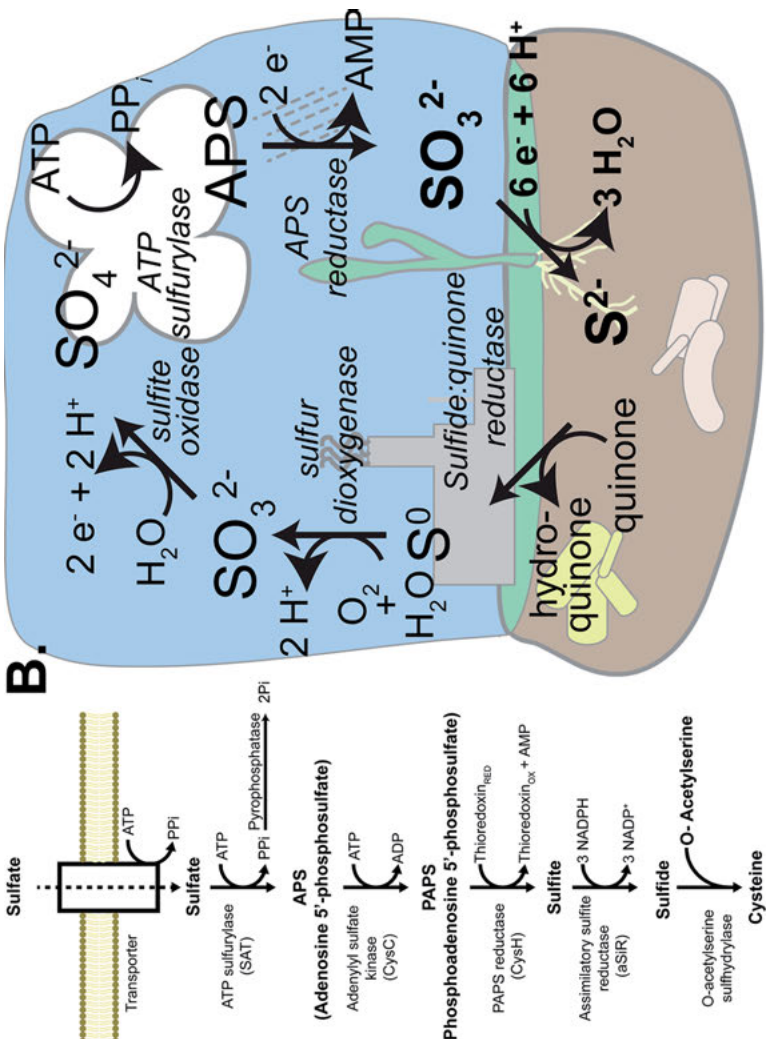


Figure 3. Sulfur metabolism and environmental cycling is a multi-step process. **A.** Sulfur import and metabolism in the dissimilatory (left) or assimilatory (right) pathways. **B.** Environmentally, sulfur is reduced and re-oxidized by sulfur-reducing and sulfur-oxidizing microorganisms, respectively, that inhabit diverse niches.

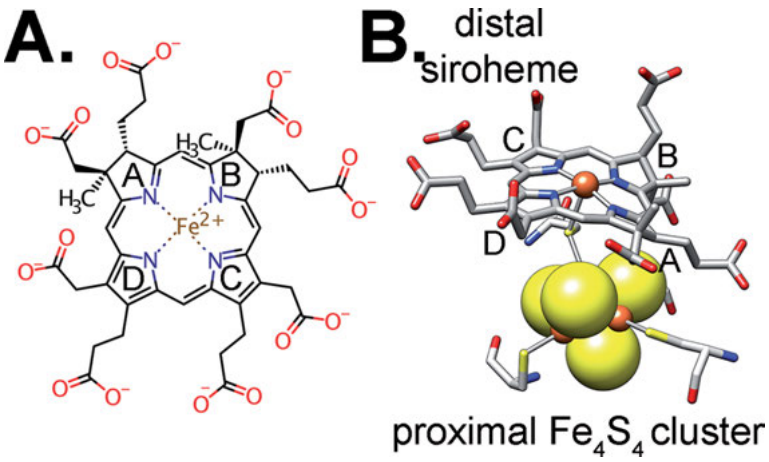


Figure 4. Siroheme is a unique porphyrinoid that works with a [4Fe-4S] cluster in most sulfite reductases. **A.** Siroheme's chemical structure. **B.** DSR and SiRHP use a covalently-coupled siroheme/[4Fe-4S] cluster through a shared cysteine ligand proximal to the active site, which sits on the distal face of the siroheme [22], first identified in the *E. coli* SiRHP homolog, which is the homolog pictured here.

importer. A two-electron reduction yields SO_3^{2-} through the activated adenosine-5′-phosphate sulfate (APS), either directly by APS reductase (E. C. 1.8.99.2) [9] or via a 3′-phosphoadenosine-5′-phosphosulfate (PAPS) intermediate by PAPS reductase (E. C. 1.8.4.8) [10], described in more detail later. Then, SO_3^{2-} is reduced by six electrons to make S^{2-} by sulfite reductase (E. C. 1.8.7.1) [11, 12]. Sulfur oxidation can also be used to drive some forms of anaerobic respiration in a diverse group of Archaeal and Bacterial species, re-oxidizing S^{2-} to SO_3^{2-} then back to SO_4^{2-} [13–15].

The dissimilatory and assimilatory sulfur reduction pathways are independent in terms of biochemical function, but they share common sulfite reductase homologs that catalyze the central six-electron reduction of SO_3^{2-} to S^{2-}. There are a few exceptions [16–19], but the vast majority of known sulfite reductases are iron metalloenzymes that use an iron-containing porphyrinoid called siroheme that is covalently coupled to a [4Fe-4S] cluster through a shared cysteine ligand [20–24] (Figure 4). These siroheme-dependent sulfite reductases are also homologous to a large class of siroheme-dependent nitrite reductases that catalyze dissimilatory or assimilatory nitrogen reduction, either for respiration or biomass formation [12]. Unlike the sulfite reductases, where the majority of known enzymes are siroheme-dependent, there are many known examples of multi-heme and copper-dependent nitrite reductases that catalyze the same chemistry [25, 26].

1.3. Biological Niches of Sulfur-Reducing Organisms

Sulfur-reducing organisms, including sulfate-reducing bacteria (SRB), are a diverse group of prokaryotic organisms, bacteria and archaea, that can use sulfate as a final electron acceptor for energy conservation, at the same time that another carbon source is oxidized [27]. Dissimilatory sulfate reduction is an anaerobic process, therefore SRB are abundant in anoxic environments. Some SRB have also been identified in the aerobic zones surrounding anoxic environments, although sulfate dissimilation has not been detected in the aerobic zones [28, 29]. Analysis of 16S rRNA identified four groups of SRB [30]: Gram-negative mesophilic SRB, that includes the genera *Desulfovibrio* and *Desulfomicrobium*; Gram-positive spore-forming SRB including the genus *Desulfotomaculum*; thermophilic bacterial SRB that includes well characterized *Thermodesulfobacterium commune* and *Thermodesulfovibrio yellowstonii*; and thermophilic archaeal SRB, this group includes two species of *Archaeoglobus, A. fulgidus*, and *A. profundus.*

SRB couple sulfate reduction to oxidation of a wide variety of compounds that range from CO_2 to more complex compounds such as sugars, alcohols, and aromatic compounds, but SRB prefer organic compounds of low-molecular weight as their carbon source [31, 32]. Thus, SRB are central in the overall sulfur cycle, but also have an important role in the carbon cycle, because they participate in the process of carbon mineralization in anaerobic environments.

Additionally, SRB can use different sources of sulfur, which includes sulfate, sulfite, and thiosulfate. Some studies have shown that thiosulfate is the preferred sulfur source in SRB found in sewer systems [32]. There is also evidence that organic sources of sulfur, such as cysteine, are also used by SRB, but to a minor extent [32]. This metabolic flexibility allows SRB to colonize a wide variety of environments including lakes, marine, and fresh-water sediments; deep-sea hydrothermal vents; and mine drainages, petroleum-polluted environments and rice fields [31]. Given the wide distribution of SRB, it is not surprising that they thrive in the laboratory under diverse conditions. For example, the optimum growth temperatures for pure cultures of SRB range from 24 °C for some *Desulfobacteria* to 70 °C for *Thermodesulfobacterium commune* [32]. Additionally, isolated cultures of SRB have shown growth only between pH 5.5 and 9, even for those species of SRBs isolated from mine drainages, where pH can range between 3 and 4. It has been proposed that microniches of higher pH exist in those acidic environments that allow for SRB survival [32].

Finally, it is important to consider that many of the environments in which SRB thrive are polluted with heavy metals. SRB seem to be less affected than other microorganisms by heavy metals, however, because the product of dissimilatory sulfate-reduction, hydrogen sulfide, reacts with heavy metals producing highly stable metal-sulfides [27].

2. SULFUR IN BIOLOGY

2.1. Activation of Sulfur for Biochemical Reactions

Sulfur-based anaerobic respiration links to an ancient time before the advent of an oxygen-rich environment. SRB produce energy by coupling oxidation of organic compounds or hydrogen to the reduction of sulfate (SO_4^{2-}) to sulfide (S^{2-}). This process cannot be performed in a single step due to the chemical stability of sulfate. Therefore, sulfate needs to be activated prior to being used by the cell as a terminal electron acceptor through the dissimilatory sulfate reduction pathway (Figures 2 and 3). A parallel process called sulfate assimilation reduces sulfate to produce cysteine. Sulfate assimilation is widely distributed amongst prokaryotic and eukaryotic organisms (Figure 2).

The first step in sulfate dissimilation is the transport of sulfate into the cytoplasm, where the intracellular ATP pool can be used for sulfate anion activation. In SRB organisms there are two types of sulfate transporter systems: high-affinity transporters, which are expressed under limiting sulfate conditions; and low-affinity transporters that are expressed under high concentrations of environmental sulfate [33, 34]. A comparative genomic analysis of 44 SRB identified that sulfate transporters in these organisms belong to four different families [34]. They include: (1) a member of the APC superfamily of anion transporters, sulfate permease (SulP, TCDB 2.1.53) [35]; (2) a divalent anion:Na^+ symporter [36] (DASS, TCDB 2.A.47); (3) a sulfate-binding protein encoded in the *cysP* gene of the *cys* operon [37] (CysP, TCDB 2.A.20.4.1); and (4) a high affinity transporter, encoded as *cysZ*, also in the *cys* operon [38] (CysZ, TCDB 2.A.20.102).

Interestingly, the ABC transporter family member CysAWTP [39, 40] is not represented in SRB organisms despite being the most studied sulfate transporters in sulfate assimilatory organisms like *Escherichia coli*. To explain this phenomenon, it is important to consider that ABC family transporters couple hydrolysis of one molecule of ATP to the transport of one molecule of sulfate. Other bacteria like *Desulfovibrio vulgaris* produce only one ATP per imported sulfate molecule when grown with hydrogen as electron donor. Thus, the use of ATP to transport sulfate would produce zero net energy gain for the cells [34]. Moreover, the use of CysAWTP inhibitors in different species of *Desulfovibrio* did not affect sulfate transport [34].

Once a sulfate molecule is incorporated into the cell, its activation is catalyzed by the enzyme sulfate adenylyltransferase (SAT, or ATP sulfurylase, E. C. 2.7.7.4). SAT is a Mg-dependent nucleotidyl transferase that belongs to the α/β phosphodiesterase superfamily. The SAT superfamily includes a structurally diverse group of enzymes that are involved in several sulfate-related metabolic pathways, and therefore they are widespread in nature. In the sulfate dissimilation pathway, sulfate activation uses one molecule of ATP to produce APS and PP_i. To drive the reaction towards the formation of APS, the PP_i molecule is quickly hydrolyzed by a pyrophosphatase in a reaction that consumes a second molecule of ATP. Additionally, SAT is strongly inhibited by APS [41].

The first reductive step of APS is catalyzed by the enzyme APS reductase (AprBA) that produces sulfite and AMP. This is a two-electron reduction step. AprBA is a heterodimeric protein that contains a [4Fe-4S] cluster and FAD as cofactors [42]. Sulfite is then reduced to sulfide in a six-electron reaction by dissimilatory sulfite reductase (DSR) [42], discussed in detail below.

How the electron transport chain is linked to the sulfite reduction is not well understood because SRB lack the main respiratory complex that is well-characterized in other bacteria, such as complex I and the bc_1 complex [43]. Moreover, the terminal reductase, DSR, is not associated to the membrane, so it does not participate in the generation of an electrochemical potential across the membrane. In addition, the primary electron donors for AprBA and DSR in SRB are not known [43]. In *Desulfovibrio* the coupling of electron transport and sulfate reduction has been studied to some extent. In general, sulfate reduction is coupled to fermentation of some organic source, such as lactate. Two models have been proposed to couple both processes. First, the hydrogen cycling model proposes that upon fermentation of lactate, the protons released are transformed to H_2 by a hydrogenase. This H_2 will diffuse to the periplasm, where electrons and protons will be reformed by periplasmatic hydrogenases, to produce a chemical potential across the membrane [27, 44]. The second model proposes that AprBA and DSR receive electrons directly from the menaquinone pool in the membrane, which receives reducing equivalents from the fermentation of a carbon source [27, 45]. There is no clear evidence as to which model prevails and comparative genetic analyses proposed that SRBs use diverse processes, which provides them with greater metabolic versatility than being locked into a single pathway [45].

2.2. Sulfide in Biomolecules

Sulfur is found in a range of biomolecules in all organisms. There are two sulfur-containing amino acids, cysteine and methionine (Figure 1). Cysteine's free sulfhydryl group confers on this amino acid several specialized roles in protein structure. It can form cysteine-cysteine disulfide linkages that play an important role in protein structure, for example in holding together multi-subunit complexes like immunoglobulins or keeping peptide hormones like insulin folded. Methionine's sulfur is not free to react with other chemicals so methionine's structural role is more passive than cysteine's.

Cysteine's reactive sulfhydryl also drives diverse reactivity within protein structure. Cysteine can serve as the first shell ligand to iron-sulfur clusters of varying stoichiometry ([2Fe-2S], [3Fe-4S], and [4Fe-4S]) [46]. Cysteine can take a special role in catalyzing chemistries as it does in a family of DNA repair proteins (O^6-alkylguanine-DNA alkyltransferase) that irreversibly transfer methyl or alkyl groups from an alkylated guanine to a reactive cysteine [47]. Cysteine's reactivity to heavy metals allows it to serve a special role in metalloenzymes but also is one cause of heavy metal poisoning where cysteines in

essential proteins like the Na/K ATPase, tubulin, or metallothionein become irreversibly bound by a heavy metal ion like mercury, cadmium, lead, or arsenic [48].

The reactive sulfur–methyl bond in SAM allows it to serve as a methyl donor for a variety of biological chemical reactions important from bacterial chemotaxis to eukaryotic histone methylation. A highly conserved class of SAM-dependent methyltransferases called *S*-adenosyl-L-uroporphyrinogen III methyltransferases (SUMTs) play an essential role in siroheme, siroheme-derived heme, and cobalamin biosynthesis [49].

2.3. Sulfur-Based Redox Regulation

Sulfur's redox flexibility makes sulfur-containing cofactors special because it confers unique reactivity on those molecules. Glutathione (GSH) and its close partners bacillithiol or mycothiol (BSH/MSH, Figure 1), respectively, serve as cellular redox buffers to maintain homeostasis under reactive oxygen species (ROS) stress. Glutathione is primarily found in γ-proteobacteria and eukaryotes, suggesting an evolutionary link upon the symbiotic merging of a prokaryotic cell and a proto-eukaryote that eventually lead to the mitochondrion [50]. BSH or MSH, used by Firmicutes or Actinomycetes, serve a similar function as redox buffers, but also allow those potentially pathogenic bacteria to evade the ROS reaction from the mammalian immune system [51].

2.4. Sulfur-Containing Lipids

Sulfur is also found in the headgroup of a broad class of sulfolipids. Sulfatides have a sulfate ester moiety attached at the carbohydrate group of a glycosphingolipid [52]. The sulfur can be directly attached to the carbon lipid chain, forming a sulfolipid, or through conjugation through an oxygen, forming a sulfonated lipid. Sulfolipids have a reduced sulfur whereas sulfonated lipids have oxidized sulfur. Sulfur-containing lipids play a role in host-pathogen interactions [53–55]; photosynthesis [56, 57]; and formation of the myelin sheath [58].

3. SIROHEME-DEPENDENT SULFITE REDUCTASE HEMOPROTEIN

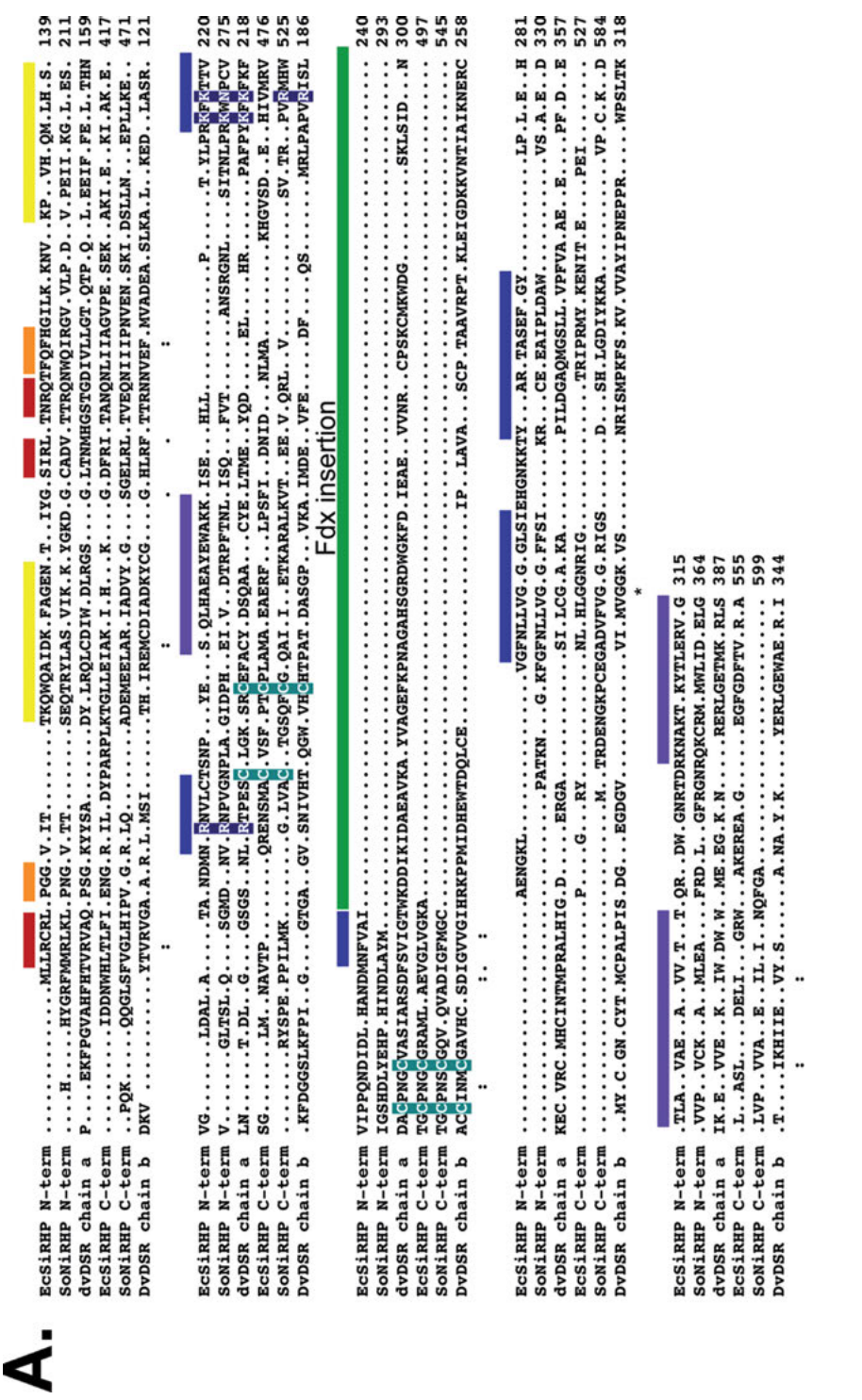
3.1. Two Types of Sulfite Reductase Are Evolutionarily Linked by Use of a Siroheme Cofactor

The siroheme-dependent sulfite reductases are marked by their uniquely coupled [4Fe-4S] cluster and the siroheme porphyrinoid that bind within a central

cleft of a shared protein fold called the sulfite and nitrite reductase repeat [22] (SNiRR, Figure 5). Siroheme is an ancient iron-containing tetrapyrrole that is synthesized from the common tetrapyrrole precursor uroporphyrinogen III through a distinct branch off the pathways that yield the more geochemically modern porphyrinogen III derivatives heme and chlorophyll [59–62], described in detail below. In fact, siroheme is more closely chemically related to the cobalt-containing corrin ring that makes up more than half the mass of the complex cofactor vitamin B_{12} [63] because it maintains all of its original propionyl and acetyl sidechains at the corners of the pyrroline/pyrrole rings (Figure 4). In contrast, heme and chlorophyll undergo extensive decarboxylation and further modification along their maturation pathways [64]. [4Fe-4S] clusters are used as structural and redox-active cofactors across a broad range of proteins and chemistries.

Bridging between distinct clusters or to other free metals is not infrequent, as seen in the Fe-Mo cofactor of nitrogenase [65] or Mo-containing carbon monoxide dehydrogenase (see Chapters 8 and 11) [66]. Some carbon monoxide dehydrogenases couple a [3Fe-4S] cluster to a nickel [67]. Coupling of tetrapyrroles is also fairly common, as in multiheme nitrite reductases [25] or the photosystems [68]. To our knowledge, however, use of a [4Fe-4S] cluster cysteine ligand simultaneously as the distal ligand to the iron of an iron-containing tetrapyrrole is unique.

The electrons that reduce the SO_3^{2-} substrate ultimately derive from cellular pools of the ubiquitous two-electron carrier NADPH. In SiRHP, the siroheme heme iron has a redox potential of –340 mV whereas the [4Fe-4S] cluster has a redox potential of –405 mV [20, 69] and the enzyme can exist in an oxidized, one-electron, or two-electron reduced state [20, 69]. How electrons get to the various sulfite reductase homologs differs, one aspect that differentiates amongst sulfite reductase family members.



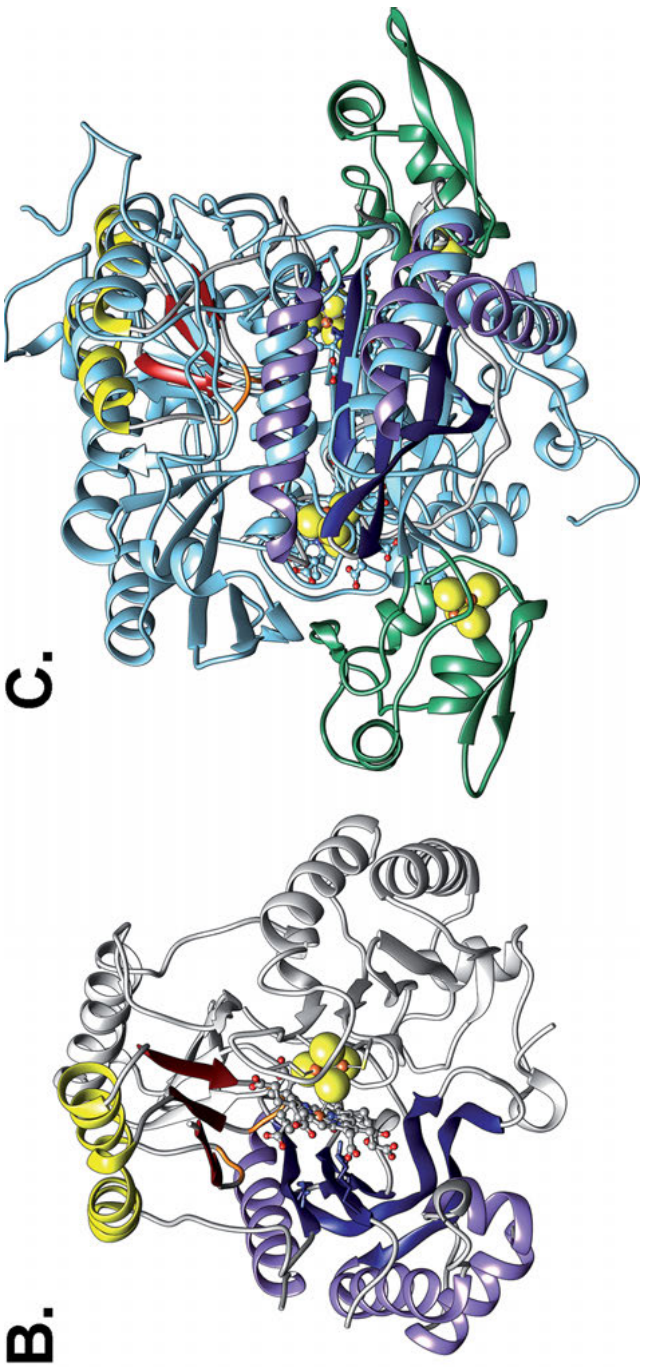


Figure 5. DSR and SiRHP are structural homologs. **A.** Sequence homology for the SNiRR core fold shows conservation of the cluster-binding cysteines (teal) and distal cationic amino acids (dark blue). The structural elements of the SNiRR are colored as follows: domain 1 β-sheet core (red), parachute domain harnesses (orange), α-helical parachute (yellow), domain 2 β-sheet core (blue), and domain 2 α-helices (purple). The ferredoxin insertion in DSR is green. The sequences shown are for the N-terminal SNiRR of *E. coli* SiRHP, the N-terminal SNiRR of *S. olerecea* NiRHP, the α chain of *D. vulgaris* DSR, the C-terminal SNiRR of *E. coli* SiRHP, the C-terminal SNiRR of *S. olerecea* NiRHP, and the β-chain of *D. vulgaris* DSR. Sequence alignments were performed in Clustal Omega [164]. **B.** SiRHP is a two-fold symmetric monomer where the active site cofactors are bound in a deep cleft on one of its faces. The structural elements are colored as in A. Coordinates are from PDB ID 6C3X [23]. **C.** The core of the *D. vulgaris* DSR is homologous to the SNiRR identified in SiRHP and the ferredoxin insertion is positioned near the active site siroheme. Only one α/β pair from the heterotetramer is shown. Coordinates from PDB ID 2V4J [73].

3.1.1. Heterotetrameric Dissimilatory Sulfite Reductase

Heterotetrameric dissimilatory sulfite reductase (DSR) with an $\alpha_2\beta_2(\gamma_2)$ oligomeric state is found in a range of SRB, allowing those organisms to reduce sulfur for energy conservation. DSR is widely believed to have resulted from a gene duplication event early in evolutionary history, before the splitting of Archaea and Bacteria domains, because the α and β chains are paralogs [70,71]. Both the α and β chains contain a ferredoxin domain insertion (Figure 5), suggesting that gene fusion between the protypical DSR and ferredoxin occurred before the duplication event that created the heterodimer core. The γ chain, on the other hand, binds the heterotetramer more weakly and is a homolog of TusE, a bacterial protein that takes sulfur from sulfur mobilization Isc proteins and transfers it to uridine in thiouridine synthesis [72].

Each α or β subunit binds 2 [4Fe-4S] clusters, one that makes up the active site and another as part of the ferredoxin [73–76]. There is some discrepancy in the tetrapyrrole and Fe-S cluster composition of the subunits between four DSR classes, defined by the UV/vis spectroscopic properties that arise from the cofactors. Desulforubidin, found in *Desulfomicrobium* and *Desulforsarcina*, has an absorbance maximum at 545 nm. These homologs house four [4Fe-4S]-coupled sirohemes [77]. DSR from *Archaeoglobus* has the same $\alpha_2\beta_2$ composition with four sirohemes per heterotetramer and an absorbance maximum at 545 nm [78]. Interestingly, the active site in the β subunit is blocked by a bulky tryptophan that projects into the active site, suggesting that siroheme-[4Fe-4S] serves a structural role [74]. Desulfoviridin, found in *Desulfovibrio*, has an absorbance maximum at 628 nm. These homologs house two [4Fe-4S]-coupled sirohemes in the α subunit with two [4Fe-4S]-coupled sirohydrochlorins in the β subunit [12, 77]. One form of DSR from *Desulfovibrio gigas* has a [3Fe-4S] cluster coupled with siroheme [75]. Desulfofuscidin, found in *Thermodesulfobacterium* and *Desulfovibrio thermophilus*, have an absorbance maximum at 576 nm. These homologs bind four sirohemes with the expected [4Fe-4S] clusters [79, 80]. P-582, found in *Desulfomaculum*, has an absorbance maximum at 582 nm and is one of the less-well studied DSRs so its siroheme content has not been confirmed [81, 82]. It is important to note, however, that the metal centers in the DSR sulfite reductase homologs are particularly oxygen-sensitive so it is possible that some discrepancies in siroheme/sirohydrochlorin or [4Fe-4S]/[3Fe-4S] cluster description could be artefactual. As purified, each of the above DSRs contains a paramagnetic ($S = 5/2$) siroheme and a diamagnetic ($S = 0$) [4Fe-4S] cluster [83].

As described above, the direct physiological electron donor for the DSRs is not firmly established, consistent with the uncertainty of the physical integration of DSRs into the oxidative phosphorylation chain. The fused ferredoxin domain is presumed by many to play a role in funneling electrons from a transient, multiheme cytochrome c3 partner [83]. Two other proteins have also been identified as involved in regulating dissimilatory sulfur reduction, DsrC and DsrD. DsrC binds the DsrAB heterotetramer more weakly than the α/β subunits bind each other and has been proposed to serve as a redox hub, linking DsrAB to energy conservation [84]. Specifically, the catalytically-important C-terminal

cysteines in DsrC have been proposed to play a role coupling DsrAB to the membrane protein complex DsrMKJOP, allowing two of the six electrons to derive from the membrane quinone pool (see Chapter 2) [73]. Further, addition of DsrC to the *Archaeoglobus fulgidus* DsrAB complex allows that enzyme to fully reduce SO_3^{2-} without releasing partially-reduced intermediates and appears to serve as the electron donor for the final four electron equivalents [85]. DsrD, on the other hand, is a small (78 amino acids) protein that adopts a winged helix motif, suggesting a role in regulating expression of DSR proteins [86]. In keeping with the diversity of DSR proteins that are found in the SRB, DsrD has also been found as a protein fusion with DsrB in a desulfoviridin from *Bilophila wadsworthia* [87]. The diversity of DSR complexes found in the SRBs means that phylogenetic analysis of the *dsrA* and *dsrB* genes is a powerful genomic tool for discriminating amongst SRBs across different habitats, even more discriminatory than the more typical used 16S RNA gene [6, 88, 89].

3.1.2. Monomeric Assimilatory Sulfite Reductase

In ancient time, a prehistoric sulfite reductase underwent a gene duplication event that replicated the core SNiRR, followed by a gene fusion event that created a monomeric sulfite reductase with a single siroheme-[4Fe-4S] active site and no ferredoxin insertion. This monomeric enzyme further evolved into what is now the prototypical assimilatory sulfite reductase hemoprotein (SiRHP) [22], which is homologous to siroheme-dependent assimilatory nitrite reductase hemoprotein (NiRHP) [90]. Interestingly, the linker that serves to fuse the two SNiRRs also fills the cavity that would have been occupied by a second siroheme-[4Fe-4S] active site, with side chains serving to mimic the siroheme's interactions with the core of the enzyme [22]. Additionally, this stretch of the protein structure is predicted to share properties typically associated with intrinsically-disordered peptides, a novel use of this type of protein structure where a plastic linker evolved to stabilize a vestigial active site cavity [91].

DSRs together comprise a varied family of enzymes, which enables them to function anaerobically under the extreme and diverse environments inhabited by the SRB. In contrast, assimilatory enzymes are more similar to one another than are their anaerobic dissimilatory homologs (Figure 5A) because they function under aerobic conditions. Nevertheless, S/NiRHPs are found in a range of organisms, from prokaryotic bacteria to eukaryotic fungi and plants, and allow those organisms to reduce sulfite or nitrite for incorporation into various biomolecules.

Quite a bit more is also known about the reducing partners that allow S/NiRHPs to function because of the stability conferred by their aerobic stability and stronger evolutionary conservation. The prototypical SiRHP found in proteobacteria like *E. coli* and *Salmonella typhimurium* binds tightly to a diflavin reductase within a heteromeric holoenzyme [11, 92, 93], described in detail below. Other sulfite reductases like those from *Mycobacterium tuberculosis*, once misidentified as a nitrite reductase [94], *Zea mays* [95], or true nitrite reductases

like that from *Spinacia oleracea* [90, 96, 97] or *Z. mays* [98], are monomeric enzymes that use a transiently interacting, reduced, ferredoxin to supply the necessary six electrons, also described in detail below.

3.1.3. Low-Molecular Weight Assimilatory Sulfite Reductase

A low-molecular weight sulfite reductase was purified from *D. vulgaris* that contained one siroheme and one [4Fe-4S] cluster per monomer [99, 100]. Although half the size of the SiRHPs described above (26,800 Da compared to 64,000 Da), the oligomeric state, UV/vis spectroscopic signals, and activity are similar so this enzyme has also been assigned an assimilatory function [99, 100]. Mössbauer [99] and Resonance Raman [101] spectroscopy confirm the cofactors are also exchange coupled in this homolog. Curiously, EPR spectroscopy reveals that unlike the other sulfite reductases, this homolog purifies with a low-spin ($S = 1/2$) siroheme, indicative of a second, strong axial ligand [79], presumably protein-derived and bound to the distal face of the siroheme where PO_4^{2-} binds in the resting state of SiRHP [22] and substrate binds after reduction [102, 103]. Resolution of that six-coordinate siroheme upon two-electron reduction to a five-coordinate, low-spin ($S = 1/2$) siroheme [101], as in other sulfite reductases, would free the active site for substrate binding to the Fe(II) siroheme.

3.2. The Core of Sulfite Reductase Is a Coupled [4Fe-4S]-Siroheme Assembly

Interpretation of the unique spectral properties of sulfite reductase was the source of much controversy early in studies of sulfite reductase, in part because for the most part the assimilatory homologs are more similar to one another than the dissimilatory homologs. For example, detection of a $S = 9/2$ spin in *D. vulgaris* and *Desulfosarcina variabilis* DSRs were used to argue against the presence of a spin-coupled siroheme-[4Fe-4S] site [104, 105]. Further, the discrepancies amongst the DSRs in terms of the number of sirohemes and the importance of a non-integral γ-subunit was ill-defined early on. The answers to the questions came from development of better protein preparations that ultimately allowed for the structure determination of several assimilatory and dissimilatory homologs.

3.2.1. What Is Siroheme?

Siroheme was originally discovered in the *E. coli* SiR and identified as an iron-containing isobacteriochlorin marked by its eight carboxylic acid groups and unique spectral properties [60]. Structurally, isobacteriochlorins differ from porphyrins because addition of methyl groups on C2 of ring A and C7 of ring

B break the conjugation of those pyrroline groups (Figure 4A). Consequently, rings A and B are in the half-chair conformation whereas the unsaturated pyrrole groups of rings C and D are planar (Figure 4B). After acetone extraction the pigment absorbs with wavelength maxima at 376 and 594 nm [11]. Early evidence that the pigment was different than other, more common, protoporphyin IX-derived iron tetrapyrroles included the low wavelength of the Soret band (376 nm compared to ~ 400 nm), with a lower Soret:α-band ratio (A_{376}:A_{594} = 3.5 for siroheme in acetone), and its extremely polar properties [60]. Enzyme-extracted siroheme has pH-dependent spectral properties owing to its large number of carboxylates. Its EPR signals are consistent with that of a high-spin Fe(III) heme ($S = 5/2$) that converts to low-spin Fe(II) heme upon binding cyanide [106], as it does in the enzyme [107]. Free siroheme, reduced chemically, is also capable of reducing SO_3^{2-} and NH_2OH in a pH-dependent fashion [108].

3.2.2. [4Fe-4S] Clusters, in General, and Special Features of SiRHP [4Fe-4S] Clusters

A broad range of redox chemistries have been evolutionarily tasked to Fe-S cluster-containing metalloenzymes that span a range of redox potentials [109]. Low potential, +2/+1, ferredoxin-like clusters (–100 to –650 mV) are important for electron transfer in metabolic processes like photosynthesis, nitrogen/sulfur fixation, carbon metabolism, and fermentation. High potential, +2/+1, Rieske centers (–150 to 300 mV) are important for electron transfer during oxidative phosphorylation in bacteria and mitochondria or as soluble electron donors for a range of prokaryotic and eukaryotic oxygenases. High potential, +3/+2, high potential iron-sulfur proteins (HiPIPs, +50 to +450 mV) are important in anaerobic photosynthetic bacteria (see Chapter 7).

3.2.3. Structural Coordination Between the Cofactors

In sulfite reductase, the siroheme-coupled [4Fe-4S] is coordinated by four cysteine residues [22,73,76] except in *D. gigas* DSRII where the fourth cysteine does not coordinate an iron because it binds a [3Fe-4S] cluster distal to the siroheme [75]. In the β-subunit of *D. vulgaris*, the fourth cysteine is replaced by a threonine [73] (Figure 5A). A threonine and an asparagine that are conserved in the assimilatory isoforms (Thr477 and Asn481 in *E. coli*) each make a hydrogen bond to a cubane sulfide (Figure 6A). Removing the interaction between Thr477 and the cubane sulfide by altering it to an alanine, alters the geometry of the proximal ligand that is bound to the siroheme, supporting the idea that the cofactors are carefully tuned to work together for optimal substrate binding [23]. In *E.* coli, two hydrophobic, second shell ligands to the cluster, Met444 and Phe437 are solvent-exposed and gate access to the [4Fe-4S] cluster by chemical

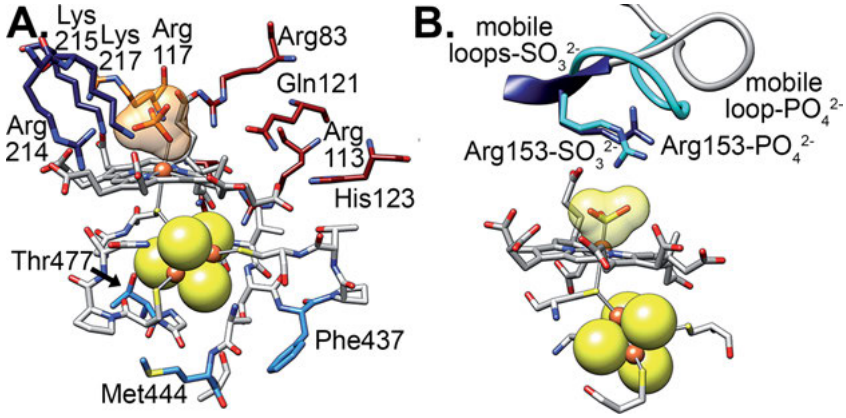


Figure 6. Anions bind to the siroheme iron. **A.** PO_4^{2-} binds to the siroheme iron in the oxidized enzyme through an oxygen, making extensive interactions with a cage of positively-charged side chains from domain 2 (blue). Positively-charged side chains from the parachute domain (red and orange) also interact with the siroheme carboxylates. Second shell amino acids in the loops that connect the Fe_4S_4 cysteine ligands (white and light blue) help tune the activity of the enzyme. Coordinates from PDB ID 6C3X [23]. **B.** Upon reduction, the PO_4^{2-} releases and SO_3^{2-} binds the siroheme iron through its sulfur atom. Arg153 (cyan) flips over to allow a loop to coalesce around the active site, trapping the anion. SO_3^{2-}-bound structure from PDB ID 2GEP [102].

reductant or the protein reductase partner (Figure 6A) [23]. A lack of conservation of these amino acids (Figure 5A) may contribute to the differences amongst the sulfite reductases that have been studied in terms of catalytic mechanism and reductase partners.

3.3. Details About the Coupled Iron Center

Sulfite reductase is unique in the way it has adopted the coupled siroheme-[4Fe-4S] cluster that acts like a metallic battery to pass six electrons to the evolving substrate. The nature of the coupling has been studied extensively with spectroscopy and structure determination, revealing commonalities amongst the assimilatory and dissimilatory homologs as well as the differences that make the homologs unique.

3.3.1. Spectroscopy Shows the Nature of the Coupling

Coupling of the siroheme and [4Fe-4S] cluster was first identified through extensive UV/vis [20], EPR [20, 77, 107], Mössbauer [77, 99, 110–112] spectroscopy, Resonance Raman [101, 113–115], and paramagnetic NMR [116, 117], and ^{57}Fe ENDOR [118, 119] spectroscopy, which all pointed to the fact that the

state of one metal cofactor impacted the state of the other. This review of the spectroscopic properties of sulfite reductase will focus on the *E. coli* SiRHP for simplicity.

In SiRHP, binding of a distal ligand to Fe(II) siroheme to lock it into a Fe^{2+} oxidation state followed by addition of a second reducing equivalent to form a $[4Fe\text{-}4S]^{1+}$ cluster results in an alteration of the UV/vis spectrum from the siroheme [20]. A similar observation was made by EPR spectroscopy. When the system is one-electron reduced, that electron goes to the siroheme because it is 65 mV more positive than the [4Fe-4S] cluster. The resulting species is EPR-silent, with a diamagnetic Fe(II) siroheme and a diamagnetic $[4Fe\text{-}4S]^{2+}$ cluster. Addition of a second electron further reduces the cluster to a paramagnetic $[4Fe\text{-}4S]^{1+}$. Measurement of the EPR spectrum of the two-electron reduced species shows the expected signal from the [4Fe-4S] cluster but also a novel signal that can be attributed to an induced paramagnetic state of the Fe(II) siroheme from coupling of the two metal centers [20]. Further, addition of a ligand, which binds on the distal face of the siroheme, affects the nature of the proximal, reduced [4Fe-4S] cluster EPR signal, depending on the nature of the ligand [107].

By low temperature Mössbauer spectroscopy, the phenomenon manifests itself in several ways. First, in the oxidized enzyme, the diamagnetic [4Fe-4S] cluster displays magnetic features in a complex spectrum [110]. Second, in the one-electron reduced system, the complex spectrum collapses to two overlapping quadrupole doublets, indicating that addition of a single electron affected all five of the active site irons [110]. Third, addition of a second electron to generate the paramagnetic $[4Fe\text{-}4S]^{1+}$ does not affect the redox state of the already Fe(II) siroheme but displays a Mössbauer spectrum with features expected from a half-integer spin heme iron [111].

Resonance Raman spectroscopy also supports electromagnetic coupling of the cofactor spin states. As with EPR and Mössbauer, Resonance Raman spectra of two-electron reduced SiRHP without added ligand show features associated with a paramagnetic Fe(II) siroheme [113]. By NMR spectroscopy, ligand binding to the siroheme resulted in upfield shifting of the resonances assigned to the cluster-binding cysteine ligands [116]. ^{1}H, ^{14}N, and ^{57}Fe couplings show that water does not serve as the proximal siroheme ligand; histidine does not serve as the distal siroheme ligand; and that there is anisotropy in the siroheme and [4Fe-4S] iron atoms that may explain the spin coupling between the two cofactors [118, 119].

3.3.2. X-Ray Crystallographic Analysis of the Cluster and Its Ligand Binding Modes Reveals Important Differences Between SiRHP, NiRHP, and DSR Homologs

3.3.2.1. SiRHP

The answer to how the two cofactors are coupled came with the first X-ray crystallographic structure of the *E. coli* assimilatory homolog that showed that one of the four cysteine ligands to the [4Fe-4S] cluster also serves as the distal ligand to the siroheme [22]. The siroheme-[4Fe-4S] cluster binds in a cleft formed between domain 1, which resembles a parachute whose harness contributes ami-

no acids that bind the siroheme and forms part of the anion binding site, and domain 2, a mixed α/β sheet that supplies the cysteine ligands to the [4Fe-4S] cluster (Figure 5B). Domain 3, which is pseudosymmetric to domain 2, completes the siroheme and anion binding site. Together half of the parachute domain and either domain 2 or 3 comprise a SNiRR. In the resting, oxidized form, a phosphate molecule is found on the distal siroheme face. Together, the presence of this phosphate and its subsequent release upon reduction demonstrate redox gating of the active site [103] and explain the observed lag phase in enzyme activity that is prevented by re-reducing the enzyme [120, 121].

Siroheme is coordinated by an extensive series of ionic interactions between each of its eight carboxylates and a large number of positively-charged lysine and arginine residues, specifically *E. coli* Arg214, Lys215, and Lys217 from the parachute domain as well as Arg83, Arg113, Arg117, Gln121, His123, and Gln396 from domain 2 (Figure 6A) [22]. Two of these, Lys215, and Lys217, also coordinate the bound anion, whether it be the gating phosphate or the evolving substrate [102]. Arg153, which also comes from the parachute domain, is particularly important because its conformation controls the structure of the loop containing amino acids 145 to 149 that is open and not well-ordered when phosphate is bound. Upon reduction and substrate binding, Arg153 flips over, causing the loop to coalesce around the bound substrate (Figure 6B).

Alteration of Lys215, Lys217, or Arg153 to a serine results in an enzyme that is unable to complement a SiRHP-deficient strain of *E. coli*, suggesting an essential role in enzyme function for each [24]. The serine-altered variants were active for electron transfer, however, suggesting that the defect in function arose from an inability to provide sufficient protons to the evolving substrate, not a deficiency in anion binding or electron transfer. Indeed, analysis of the products showed varying levels of sulfur reduction in each of the serine-altered variants that suggests Lys215 contributes the fourth proton; Arg153 contributes the fifth proton; and Lys217 contributes the final proton during the course of the six-electron reduction. This result also indicates that SiRHP is capable of performing single electron transfer events, in contrast to the proposal that the electrons move two at a time.

Arg83 is also essential for SiRHP function but its role appears to be different than the other positively charged amino acid residues that bind the anion and siroheme carboxylates [24]. Arg83 stacks against a propionate from ring B and forms salt bridge interactions with the bound phosphate but not sulfite (Figure 6A) [102]. Arg83 was one of the amino acids originally proposed to help provide protons to drive the reaction. Alteration of Arg83 had a quite different effect of SiRHP structure and function: it caused the bulk of the protein to express and purify as an apo-enzyme that oligomerizes to dimer and tetramer [24, 93]. The small fraction of metallated enzyme was more active than wild-type enzyme. Together, these observations suggest a role for Arg83 in binding siroheme, however, the contribution of a single π-π stacking interaction would seem fairly minor in the totality of siroheme binding. One possible interpretation is that Arg83 plays a role in assembling the cluster/siroheme cluster, the implications of which are discussed below.

The [4Fe-4S] cluster sits asymmetrically below the siroheme, directly under ring A (Figures 4 and 6). The uniquely bowed conformation of the siroheme that derives from its isobacteriochlorin scaffold means that the meso C9 bends down towards the cluster, coming within 3.6 Å of one of the cubane's sulfides. It has been proposed that perhaps electrons are able to move between cofactors through the covalent bonds mediated by the bridging Cys483 and also through close contact of the siroheme and [4Fe-4S] cluster [22]. This hypothesis would also explain why sulfite reductase never evolved to use a more modern tetrapyrrole cofactor like heme: the fully conjugated, unsaturated ring of protoporphyin IX-derived hemes would be planar.

3.3.2.2. NiRHP

The NiRHP active site is virtually identical to the SiRHP active site [90] and given the similarity between the SiRHPs and the NiRHPs, it can be difficult to discriminate between them based on sequence or structure. In fact, the SiRHP (SirA) from *M. tuberculosis* was originally mis-identified as a NiRHP (NirA) [94]. Biochemically, both homologs turn over nitrogen-containing substrates more readily than sulfur-containing substrates but SiRHP binds sulfite (SO_3^{2-}) more tightly than nitrite (NO_2^-) or hydroxylamine (NH_2OH), demonstrating substrate specificity [122, 123]. Structurally, this differentiation is explained because the loop that encloses SO_3^{2-} after substrate binding does not fold down when NO_2^- binds [102]. Consequently, SO_3^{2-} binds more tightly to SiRHP, which explains both its specificity for SO_3^{2-} and the more rapid turnover of the nitrogen-containing substrates. In NiRHP, the loop closes with NO_2^- binding [124]. The importance of the active site loop in determining anion specificity is further supported by the observation that altering Asn149 to a tryptophan in *E. coli* SiRHP blocks loop closure and increases the rate of SiRHP to SO_3^{2-} reduction [24]. Another important difference between SiRHP and NiRHP is that Lys217 is an asparagine in NiRHP so it does not interact with the bound substrate anion [124]. Alteration of Asn226 to a lysine increases the reactivity of tobacco leaf NiRHP to SO_3^{2-} reduction by 100 % [125].

3.3.2.3. DSR

The DSR siroheme-[4Fe-4S] cluster is similarly located within a single α/β heterodimer but with an important difference: the siroheme/anion binding residues derive from one subunit whereas the cluster-binding ligands derive from the other (Figure 5A) [73, 74, 76]. Another important difference between DSR and SiRHP is that there are two siroheme-[4Fe-4S] cluster assemblies per heterodimer (four total in the heterotetramer) as well as additional [4Fe-4S] clusters that derive from the ferredoxin domain insertion (four additional [4Fe-4S] clusters) (Figure 7A). This creates a unique, intricate network of cofactor-cofactor interactions in DSR because there is close approach between the ferredoxin [4Fe-4S] cluster and the siroheme-[4Fe-4S] cluster from an α/β heterodimer, which are about 10 Å apart (Figure 7B). Each pair of α/β siroheme units are also separated by about 10 Å across the heterodimer interface, bridged by a siroheme-coordinating arginine between the ring C carboxylates. In some

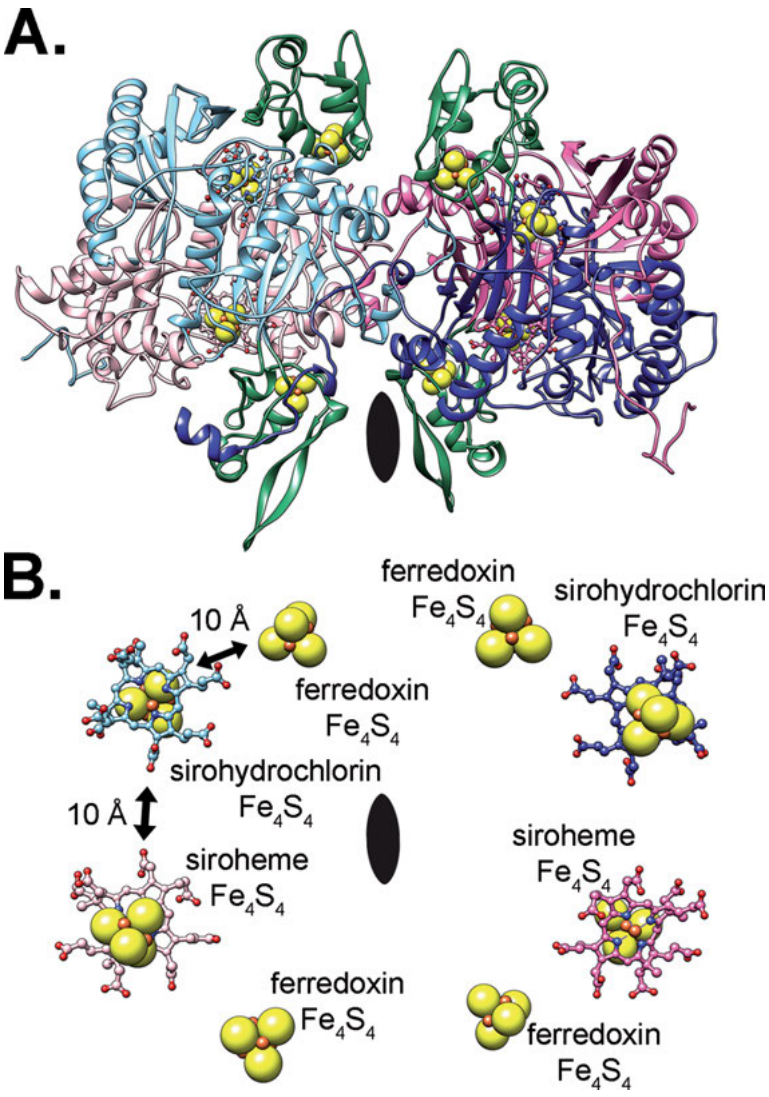


Figure 7. DSR is a heterotetramer. **A.** Two α/β heterodimers are related by a two-fold symmetry axis. In the *D. vulgaris* isoform, each subunit binds a siroheme-[4Fe-4S] cluster or a sirohydrochlorin-[4Fe-4S] cluster, so only one active site per heterodimer is active. **B.** The siroheme-sirohydrochlorin cofactors are about 10 Å apart. The isobacteriochlorin rings are about 10 Å from the ferredoxin [4Fe-4S] cluster. Coordinates from PDB ID 2V4J [73].

DSRs, this would link a functional and a structural siroheme unit [74] whereas in others it would link a siroheme to a sirohydrochlorin unit [73] or a [3Fe-4S] cluster-coupled siroheme [75]. In SiRHP, the vestigial active site is filled by the 18 amino acid-long linker that joins the two SNiRRs and whose amino acid

side chains fulfill many of the same interactions that are formed between the siroheme carboxylates and the positively-charged amino acids from domains 1 and 2 [22].

As in SiRHP, the active site of DSR has two lysines and two arginines that bind the anion and the siroheme carboxylates. Unlike SiRHP, however, the active site amino acids do not change conformation during the course of catalysis [126]. This more rigid active site and the more pronounced channel leading to the siroheme also explain why DSR is able to react with, and release, trithionite ($S_3O_6^{2-}$) and thiosulfate ($S_2O_3^{2-}$) because there is space for a second sulfite to bind adjacent to the siroheme-bound substrate and the two sulfurs could react if the siroheme-bound substrate becomes partially reduced [126].

3.3.3. Implications of the Cofactor Coupling on the Enzyme's "Push-Pull" Catalytic Mechanism

The overall reaction for sulfite reduction requires both electron for the reduction of the sulfur and proton transfer to facilitate the three dehydration events as the substrate changes, a push-pull type reaction. The highly charged ligand binding, lines with three or four lysines and arginines (Figure 6), serves to increase the local proton concentration, pulling charge away from the very negative siroheme and substrate. Three of the four charged amino acids in *E. coli* are implicated as specifically important for supplying the final three protons, whether directly or through the extensive array of ordered active site water molecules [24]. The first three protons likely derive from water molecules and the protonated substrate [22]. On the distal face of the siroheme, second shell ligands to the [4Fe-4S] cluster also help tune the reactivity of the siroheme by shielding the cluster from bulk solvent and gating access of the reducing equivalents [23]. In this way, the protein scaffold pulls electrons that are pushed from the reductase, whether a protein enzyme or a chemical, funneling them to the siroheme and then to the [4Fe-4S] cluster, which allows release of the resting-state phosphate to open the active site pocket for substrate binding. The covalent link between the cofactors is essential for supplying the required six electrons to complete turnover.

4. SULFITE REDUCTASE HOLOENZYME

4.1. SiR Is an Oxidoreductase

As with all redox enzymes, the hemoprotein oxidase does not act alone but, rather, works with a reductase partner. Although the cofactors in SiRHP and DSR are magnetically and covalently coupled, they each act as single electron acceptor, storing up to two electrons. The Fe^{3+}/Fe^{2+} electronic couple has a redox potential of –340 mV [69] whereas the $[4Fe\text{-}4S]^{2+}/[4Fe\text{-}4S]^{1+}$ couple has a redox potential of –405 mV [20], so the first electron is stored on the siroheme

and the second on the [4Fe-4S] cluster. In isolation, SiRHP can be chemically reduced but care must be taken to pick the correct chemical reductant because the commonly used sulfur-containing reductants like dithionite ($S_2O_4^{2-}$) might contain SO_3^{2-} [127], so complicate interpretation of any measurement. Cr(II)-EDTA-reduced methyl viologen is a better choice.

4.2. Different SiR Homologs Use Different Reductases

In keeping with the subtle differences between sulfite reductase homologs, there is also a fair amount of diversity in the makeup of the holoenzyme oxidoreductase complexes. As described above, there is some uncertainty about the ultimate electron source for the DSRs. The monomeric SiRHPs either use a stably-bound diflavin reductase or a dissociable ferredoxin. Ultimately, both the diflavin reductase and the ferredoxin are channels for electrons from cellular NADPH pools, either directly or through ferredoxin reduction by an NADPH ferredoxin reductase.

4.2.1. DSRs with Fused Ferredoxin Domains

DSRs evolved via a gene fusion that inserts a ferredoxin-like domain into domain 2 of a single SNiRR (Figures 5 and 7). Together with domains 1 and 3, the ferredoxin insertion forms an ~15 Å tunnel to the deeply-buried siroheme active site [74]. As positioned, the ferredoxin [4Fe-4S] cluster is about 10 Å from the siroheme [4Fe-4S] cluster. The presumption is that the ferredoxin participates in electron transfer, but how the ferredoxin domain is reduced is not known. Nevertheless, there are two direct paths between the clusters. One occurs through a solvent-bridged, side-chain/main-chain interaction between the two [4Fe-4S] cluster's cysteine ligands. Another occurs between a methionine that stacks against a cysteine ligand to the siroheme [4Fe-4S] cluster but whose main-chain amide-hydrogen interacts with a sulfur within the ferredoxin [4Fe-4S].

In addition to reduction from the second [4Fe-4S] cluster, DsrC has also been proposed to play a role in the catalytic cycle because two C-terminal cysteines can serve as electron carriers to supply the final electrons needed to complete turnover (see Chapter 2) [84]. In fact, inclusion of DsrC promotes full six electron turnover [85]. Functionally, of course, DSRs did not necessarily evolve under pressure to perform the full turnover because turnover of SO_3^{2-} is coupled to energy conservation in the cell, not synthesizing sulfide for biomass formation. Accordingly, DsrC is expressed in a different operon than DsrAB in some DSRs [128]; it is not a core component of the $\alpha_2\beta_2$ heterotetramer [75], and it is free in cells where it can interact with other proteins [76, 129].

4.2.2. SiRHPs with Ferredoxin

Monomeric SiRHP can work with a transiently interacting ferredoxin reductase whereas NiRHPs almost exclusively function as ferredoxin-coupled oxidases [83, 130]. A series of three structures of a *Z. mays* chloroplast SiRHP bound to ferredoxin is the only known structure thus far of a siroheme enzyme with its reductase (Figure 8A-C). Ferredoxin binds at the cavity formed by the N-terminus of SiRHP and the solvent-exposed edge of the siroheme. In each of the three ferredoxin-bound molecules, the ferredoxin [2Fe-2S] cluster is about

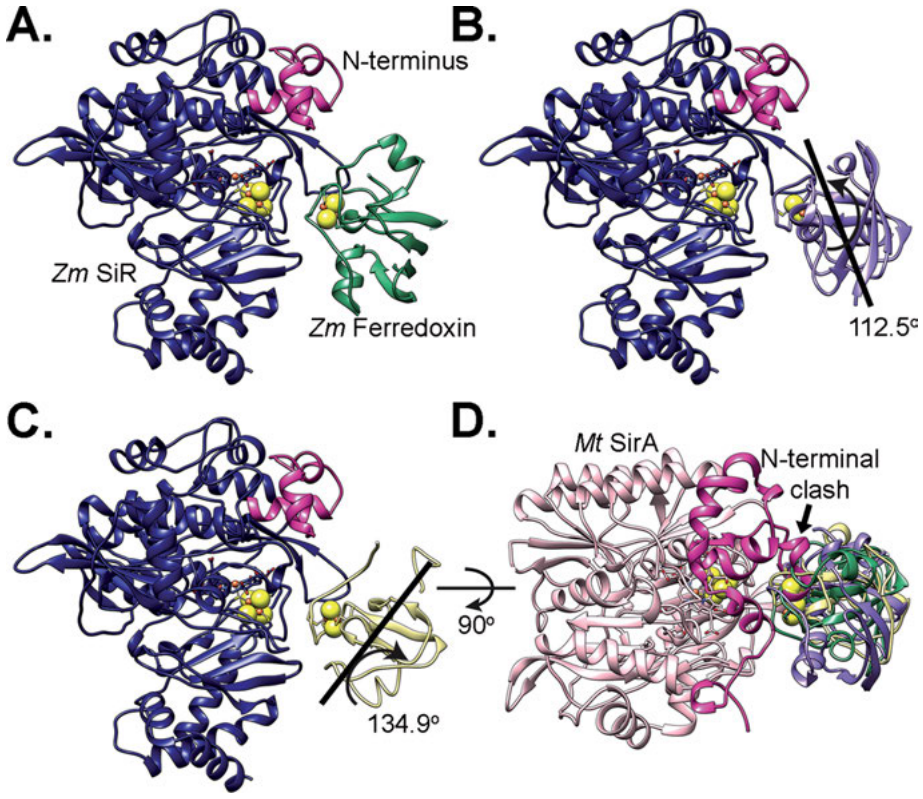


Figure 8. *Z. mays* SiRHP uses a ferredoxin reductase partner, which it binds in a heterogeneous way. **A.** One conformation of the SiRHP-ferredoxin dimer shows close approach of the Fe-S clusters. The N-terminus, which mediates the interaction, is pink and the ferredoxin is green. Coordinates from PDB ID 5H8Y [95]. **B.** Another conformation involves a rotation of the ferredoxin (purple) that maintains close contact of the Fe-S clusters. Coordinates from PDB ID 5H92 [95]. **C.** A third conformation, which was less well resolved than those in A or B, also maintains close contact between the Fe-S clusters. The ferredoxin is in yellow. Coordinates from PDB ID 5H92 [95]. **D.** The *M. tuberculosis* SirA, which also uses a ferredoxin reductase, shows steric clashing when superimposed on the *Z. mays* dimer, suggesting homolog-specific contacts direct this interaction. *M. tuberculosis* coordinates from PDB ID 1ZJ8 [94].

10 Å away from the siroheme [4Fe-4S] cluster [95], similar to the relative distance of the ferredoxin insertion [4Fe-4S] cluster to the siroheme [4Fe-4S] cluster in DSR. The relative orientation of the ferredoxin differs slightly between the three *Z. mays* structures and fairly dramatically between the insertion domain in DSR.

M. tuberculosis is another sulfur-reducing organism that uses an assimilatory sulfite reductase hemoprotein (SirA) coupled with a ferredoxin [94], however it is unknown which of the five *M. tuberculosis* isoforms it uses. Interestingly, superimposition of *M. tuberculosis* SirA and *Z. mays* SirA:ferredoxin shows that there is incongruity in the loops that mediate the interaction of the oxidase with its reductase (Figure 8D), suggesting that there are homolog-specific details about how the subunits interact. In contrast, ferredoxin fits well when *S. oleracea* NiR [90] is docked onto *Z. mays* SirA:ferredoxin.

4.2.3. SiRHPs Work with SiRFPs

Other SiRHPs from proteobacteria function as dodecameric holoenzyme complexes, where the oxidase and reductase make a tight complex with eight α subunits and four β-subunits ($\alpha_8\beta_4$) [11, 93]. The α subunit is a diflavin reductase called the flavoprotein (SiRFP) [92, 131] and is a homolog of the well-characterized cytochrome P_{450} reductase (CPR). This is a three-domain protein with an NADPH-reductase like FAD-binding domain interrupted by a connection domain that is joined to a ferredoxin-like FMN-binding domain by a flexible linker [132]. Compared to CPR, SiRFP is exceptionally dynamic and the FMN-binding domain undergoes much more dramatic conformational changes to pass electrons to SiRHP than in the analogous cytochrome system. The β-subunit is SiRHP [92].

As expressed, SiRFP is an octamer that is mediated by the N-terminal 52 amino acids [133]. This variable region of the diflavin reductases is what tethers monomeric CPR to the membrane [134] so the octameric nature of SiRFP makes it unique in this class of enzymes that includes the reductase domain of nitric oxide synthase, the reductase subunit of methionine synthase, and the flavin-binding domain of cytochrome P_{450} BM3 [135]. As expressed alone, of course, SiRHP is a monomer [136]. A minimal dimer formed by expressing SiRFP as a monomeric, N-terminal truncation and SiRHP maintains reduced activity [132, 137] so others have suggested that perhaps the uneven subunit stoichiometry is incorrect [137]. If equimolar octameric SiRFP and monomeric SiRHP are mixed, however, free SiRHP remains [93] suggesting that, indeed, the subunits assemble in a 2:1 ratio. The increase in local concentration afforded by the extra reductase subunits could, in part, explain why SiR holoenzyme but not DSR is able to fully reduce SO_3^{2-} without releasing partially-reduced intermediates.

There is no known structure of the SiR holoenzyme but recent progress has begun to tease out how the subunits interact. Early work suggested that the subunits interact via a charge-charge interaction network at the siroheme side of SiRHP facilitated by the FMN binding domain of SiRFP. The first claim was

based on the structure of *E. coli* SiRHP, which, importantly, is proteolytically-cleaved at amino acid 60 to allow crystallization, and then does not have the density for the next 20 amino acids and so has missing information [22]. The second claim is based on experiments in which truncated SiRFP was attached to a cyanogen bromide matrix and then SiRHP was crosslinked to it [138], not a straight-forward binding experiment.

Subsequent hydrogen-deuterium exchange/mass spectrometry experiments provided a map for exploring how the subunits interact [93]. First, the N-terminal 80 amino acids of SiRHP mediate its interaction with SiRFP. When SiRHP is expressed as an N-terminal truncation, the subunits can no longer bind. Second, there are two parts of SiRFP whose solvent accessibility changes when it binds SiRHP. One is within the FMN-binding domain and the other near the NADPH binding region of the FAD-binding domain. Only the site that is within the FAD-binding domain is able to mediate stable interactions with SiRHP. The model that emerges suggests that two interactions govern holoenzyme assembly: a stable, structural interface between the FAD, binding domain of SiRFP and the N-terminus of SiRHP and a transient, functional interaction that drives electron transfer. Further, the interaction is governed by hydrophobic interactions involving, minimally, SiRFP Phe497 and Val500 [91] and SiRHP's N-terminus, which has a conspicuous number of leucines and isoleucines that are not conserved in homologs that use a ferredoxin reductase and whose N-termini are structured in their X-ray crystal structures (Figures 5 and 8).

The coincidence of the extreme flexibility of SiRHP's N-terminus, which is protected upon holoenzyme formation, and its role in mediating binding to SiRFP suggests there could be functional consequences on SiR's mechanism. The active site loop that coalesces to trap the substrate after Arg153 flips over (Figure 6B) [102] sits right under the predicted position of the N-terminus of SiRHP, based on a small-angle X-ray scattering envelope [93]. It is conceivable that SiRHP itself must undergo a conformational change relative to SiRFP in order to open the active site and allow phosphate release/substrate binding. Combined with the conformational change that is required in SiRFP to transfer electrons from the NADPH to the FAD and to the FMN cofactors, from which they move to SiRHP [122], SiR has the potential to be a highly dynamic enzyme whose structure is linked to its redox state.

5. SIROHEME BIOGENESIS

Siroheme is, no doubt, a peculiar evolutionary relict. Why sulfite reductase did not broadly evolve away from it is not an easy question to answer. It is clear that sulfur metabolism was a critical chemistry before oxygen was the dominant gas, but once O_2-driven oxidative phosphorylation emerged as the most efficient form of ATP production, nitrogen reduction became the limiting chemistry in terms of biomass formation. Nitrite reductases diversified to optimize this chem-

istry but sulfite reductases did not, but with a few known exceptions [16–19]. Nevertheless, siroheme is the link between the iron tetrapyrroles that dominated before an oxygen-rich environment and hemes that are more commonly used now. For a long time, there was a mystery about how heme was synthesized in a broad class of heme-dependent, anaerobic bacteria because they did not have the usual genes that encode the heme biosynthetic enzymes that are evolutionarily related to the mitochondrial heme biosynthetic enzymes found in eukaryotes. In the past decade, it was discovered that siroheme, not precorrin-2 or sirohydrochlorin, is an intermediate in the synthesis of fully conjugated, decarboxylated protoheme or heme d_1 [139, 140]. That is, heme can be made without going through the protoporphyrin IX intermediate but rather through an iron isobacteriochlorin intermediate.

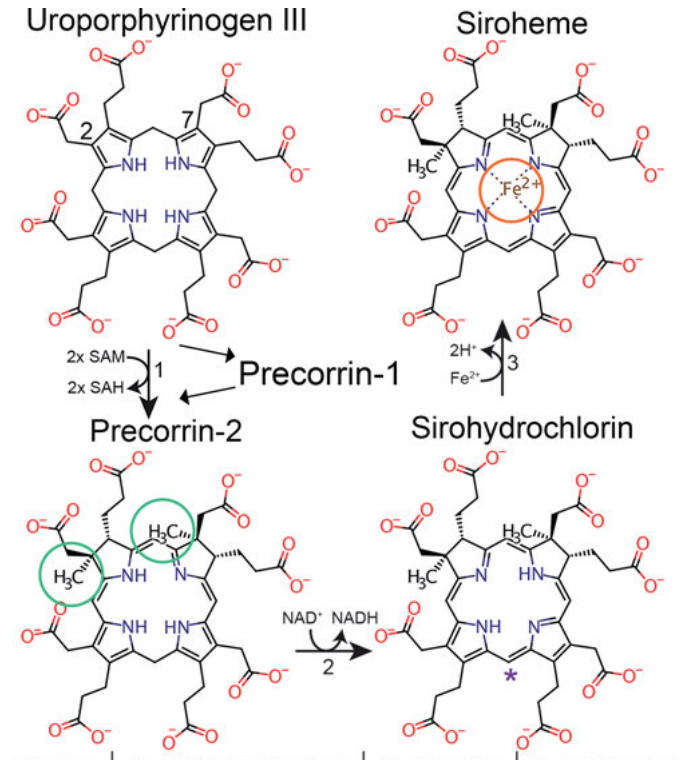
5.1. Siroheme Is an Isobacteriochlorin

The isobacteriochlorin nature of siroheme means that it is more readily oxidized than protoporphyin IX-derived hemes. It is possible that this could play an important function in the sulfite reductase catalytic mechanism because siroheme can exist in a super-oxidized state [141], where a third electron has been removed after transfer of the two electrons supplied from the reductase. If this state of the siroheme exists in the context of the enzyme, it would mean that three electrons, rather than just two, could move to the substrate without an additional interaction with the reductase subunit.

5.2. Siroheme Shares a Common Precursor with Other Tetrapyrroles

Siroheme is the simplest tetrapyrrole that derives from the common precursor uroporphyrinogen III (Figure 9) [64]. Two methylations at C2 and C7 saturate those carbons and unconjugate the resulting pyrroline A and B rings. A hydride anion (H^-) is abstracted from the bridging carbon C15 in a dehydrogenation reaction followed by deprotonation of the pyrrole/pyrroline nitrogens and iron insertion.

As a precursor to heme, then, the metallated siroheme is decarboxylated at the acetyl groups on C12 and C18 of rings C and D, resulting in a methyl group at those positions, yielding didecarboxysiroheme, catalyzed by siroheme decarboxylase (AhbAB, E. C. 4.1.1.111). Further, two decarboxylations at rings A and B remove the acetyl groups and leave single methyl groups on the now unsaturated C2 and C7, catalyzed by Fe-coproporphyrin III synthase (AhbC, E. C. 4.1.3.-) to make Fe-coproporphyrin III. Finally, the proponyl groups on ring A's C3 and ring B's C8 are oxidatively decarboxylated to make heme, by SAM-dependent heme synthase (AhbD, E. C. 1.3.98.f) [64, 139, 140].



Reaction	*E. coli/S. typhimurium*	*S. cerevisiae*	*B. megaterium*
1	CysG[A]	Met1p	SirA
2	CysG[B]	Met8p	SirC
3	CysG[B]	Met8p	SirB

Figure 9. Siroheme is synthesized from the common tetrapyrrole precursor uroporphyrinogen III in three steps, going through precorrin-2 and sirohydrochlorin intermediates. Different organisms use different sets of enzyme homologs to perform this chemistry.

5.3. Different Organisms Make Siroheme Differently

Not surprisingly, different organisms make siroheme differently but they do it with a defined set of enzyme modules (Figure 9). Three chemical reactions are needed to transform uroporphyrinogen III to siroheme: methylation, dehydrogenation, and ferrochelation (Figure 9). The enzyme module that performs the methylation is strictly conserved and is a homolog of the *S*-adenosyl-L-methionine-dependent uroporphyrinogen III methyl transferases (SUMTs, E. C. 2.1.1.107) that are involved in cobalamin (vitamin B_{12}) synthesis [142]. This single enzyme catalyzes both C2 and C7 methylations [143], positioning the twisted substrate in the central cavity between its two domains to place the unmethylated ring B, C7 proponyl near the *S*-adenosyl-L-homocysteine (SAH) while making a hydrogen bond with a main chain carbonyl [144]. Eight other charge-charge interactions hold the ring

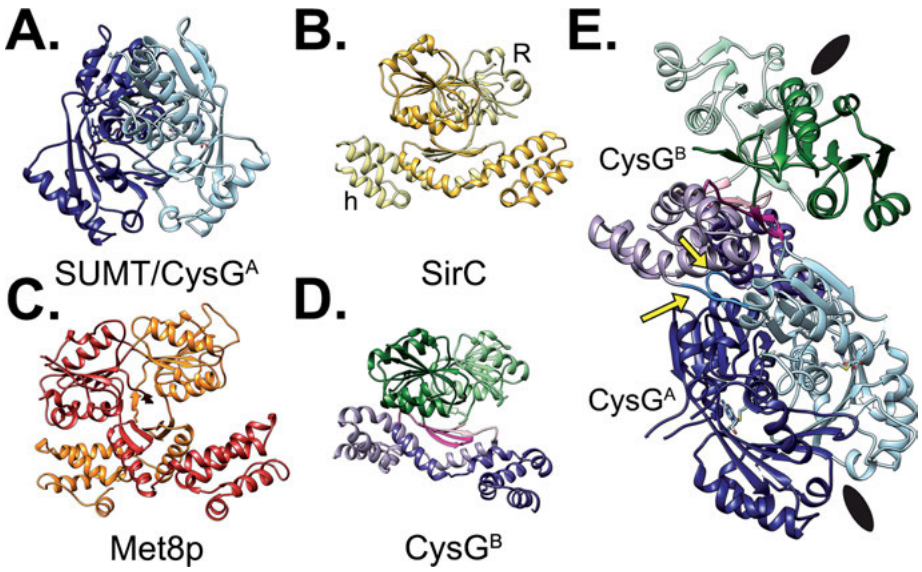


Figure 10. The enzymes for siroheme synthesis come from a set of enzyme homologs. **A.** Uroporphyrinogen III methyltransferase is a homodimer of SUMTs. Coordinates from PDB ID 1PJQ [146]. **B.** SirC is a monofunctional precorrin-2 dehydrogenase, also a homodimer. Coordinates from PDB ID 3DFZ [149]. **C.** Met8p is a bifunctional precorrin-2 dehydrogenase/sirohydrochlorin ferrochelatase that is homologous to SirC. Coordinates from PDB ID 1KYQ [148]. **D.** $CysG^B$ is a bifunctional precorrin-2 dehydrogenase/sirohydrochlorin ferrochelatase that is homologous to SirC and Met8p. Coordinates from PDB ID 1PJQ [146]. **E.** CysG is the product of a gene fusion even between the N-terminal $CysG^B$ and $CysG^A$, a C-terminal SUMT, that is able to transform uroporphyrinogen III into siroheme as a single siroheme synthase. Coordinates from PDB ID 1PJQ [146].

in the cavity and the pyrrole rings from B and D stack against an essential and conserved methionine [144, 145]. In this position, C2 from ring A is in steric conflict if a SAM were to be bound, so perhaps represents a transition-state like conformation, and either an active site arginine or glutamine could serve as the base to facilitate pyrrole nitrogen deprotonation [144]. For the second methylation to occur, the precorrin-1 and SAH products would need to diffuse out of the binding pocket and a second SAM bind followed by precorrin-1, rotated by 90°. SAH binds tightly to the enzyme as most SUMTs co-purify with bound product [145–147], so how complete product release occurs is unknown.

Dehydrogenation is catalyzed again by a single class of enzymes defined by their member's unique fold called uroporphyrinogen III dehydrogenase (E. C. 2.1.1.107) [146, 148, 149]. Uroporphyrinogen III dehydrogenase is a three-domain protein with an NAD^+-binding Rossmann fold followed by an all β dimerization domain, and then an all-helical domain that forms a homodimer with a large cavity between the Rossmann fold and the all-helical domain where sirohydrochlorin is presumed to bind [149]. A serine in the back pocket of the cavity regulates binding, either by phosphorylation [146] or tight binding of a compet-

ing divalent anion [148]. Ferrochelation is either catalyzed by a bi-functional enzyme that shares the uroporphyrinogen III dehydrogenase fold (E. C. 1.3.1.76) or by a ferrochelatase with structural, but not sequence, homology to a larger superfamily of cobalt and iron chelatases that includes the cobalamin and heme chelatases (E. C. 4.99.1.4) [146, 148, 150–152]. No metal binding by the mono-functional dehydrogenase or the bi-functional dehydrogenase/chelatase has been measured but the enzymes are highly homologous [149], so it is not clear how these enzymes differ mechanistically.

Firmicutes like *Bacillus megaterium* use three enzymes: SirA, a SUMT; SirC, a precorrin-2 dehydrogenase that binds metals without catalyzing metal insertion; and SirB, a chelatase that is homologous to heme and vitamin B_{12} chelatases (Figure 9) [153]. Fungi like *Saccharomyces cerevisiae* use a separate SUMT (Met1p) and a bi-functional precorrin-2 dehydrogenase/sirohydrochlorin chelatase (Met8p) (Figure 9) [154]. Proteobacteria like *E. coli* or *S. typhimurium* use a single gene product to synthesize siroheme, siroheme synthase (CysG) (Figure 9). CysG is the product of a gene fusion event between a C-terminal SUMT ($CysG^A$) and an N-terminal bifunctional dehydrogenase/chelatase ($CysG^B$) [151] that is a homodimer, so each dimeric enzyme module has a different two-fold symmetry axis as crystallized [146] (Figure 10). $CysG^B$ also acts as a sirohydrochlorin cobaltochelatase but, when challenged with high cobalt/low iron, is selective for iron, the mechanism for which is unknown [152]. Likewise, we do not understand the properties that discriminate Met8p/$CysG^B$ from SirC in allowing metal insertion [149].

6. FUTURE DIRECTIONS

6.1. Siroheme Chaperoning

We know much about the atomic properties that direct sulfite reductase function because of careful spectroscopic analysis followed by atomic-resolution structure determination. We do not yet know how these amazing structures come to be. There are some hints that siroheme is as carefully chaperoned as protoporphyrin IX-derived hemes, at least in *M. tuberculosis* [155], but no specific siroheme chaperone has been firmly identified. Free siroheme is chemically reactive but its chemistries are not carefully controlled [108] so there was likely evolutionary pressure away from allowing off-pathway reactions.

There are hints that the protein enzymes themselves might be involved in siroheme chaperoning or cofactor assembly into the protein scaffold because some amino acid alterations, designed so that they should not affect protein folding or siroheme binding because they are part of the disordered N-terminus of the *E. coli* SiRHP homolog, result in soluble, but unfolded and oligomeric, enzyme [93] (and unpublished observations). Formation of molten globule apo-metalloenzymes is common [156–158], but oligomerization of apo SiRHP via its N-terminus, which also facilitates SiRFP binding, might be a unique way to prevent formation of inactive SiRHP with its oxidase partner [93]. Exactly how all

the parts of the puzzle fit together remains to be seen, but its deconstruction will be important because of the potential for siroheme metabolic enzymes as rational drug design targets against pathogenic bacteria like *M. tuberculosis* [159], for example, or in developing more sophisticated sulfite reductase synthetic enzymes that could aid in bioremediation [160].

6.2. [4Fe-4S] Biogenesis in a Coupled System

Much progress has been made recently in dissecting the mechanisms for [4Fe-4S] cluster biogenesis using simple scaffolds (see Chapter 7) [161–163]. The coupled siroheme-[4Fe-4S] cluster is somewhat more complex than single clusters found in small, monomeric proteins like ferredoxin, however. We do not know which of the two *E. coli* [4Fe-4S] cluster biosynthetic pathways is at work on SiRHP, nor do we know how those enzymes might collaborate with the siroheme biosynthetic enzymes to make the complex metallic center that is central to the enzyme. Some preliminary observations we have made by purifying SiRHP from CysG-depleted *E. coli* suggests that the [4Fe-4S] cluster is inserted first, but is not stable under aerobic conditions (unpublished observations). Learning how the cell assembles such complex metal centers from potentially toxic components will be important in understanding how organisms maintain robust pathways for biomass formation and in identifying potential drug targets in sulfur-dependent pathogens.

ACKNOWLEDGMENTS

We would like to thank Mr. Joseph M. Pennington for help with proofreading and acknowledge funding to MES from the National Science Foundation (MCB1149763) for allowing study of SiR. Contribution by IA is based upon work supported by the Great Lakes Bioenergy Research Center, U.S. Department of Energy, Office of Science, Office of Biological and Environmental Research under Award Number DE-SC0018409.

ABBREVIATIONS AND DEFINITIONS

ABC	ATP-binding cassette transporter
AMP	adenosine 5′-monophosphate
APC	amino acid-polyamine organocation
AprBA	adenosine-5′-phosphate sulfate reductase
APS	adenosine-5′-phosphate sulfate
ATP	adenosine 5′-triphosphate
BSH	bacillithiol
CPR	cytochrome P_{450} reductase

CysG	siroheme synthase
DSR	dissimilatory sulfite reductase
FAD	flavin adenine dinucleotide
FMN	flavin mononucleotide
GSH	glutathione
HiPIP	high-potential iron-sulfur protein
MSH	mycothiol
NADPH	nicotinamide adenine dinucleotide
NiRHP	nitrite reductase hemoprotein
PAPS	3′-phosphoadenosine-5′-phosphosulfate
ROS	reactive oxygen species
SAH	*S*-adenosyl-L-homocysteine
SAM	*S*-adenosyl-L-methionine
SAT	sulfate adenylyl
SUMT	*S*-adenosyl-L-uroporphyrinogen III methyltransferase
SiR	assimilatory sulfite reductase
SirA	*M. tuberculosis* and *Z. mays* sulfite reductase
SiRHP	SiR hemoprotein
SiRFP	SiR flavoprotein
SNiRR	sulfite and nitrite reductase repeat
SAT	sulfate adenylyltransferase
SRB	sulfate reducing bacteria
S^{2-}	sulfide
SO_3^{2-}	sulfite
SO_2	sulfur dioxide
$S_2O_3^{2-}$	thiosulfate
$S_2O_4^{2-}$	dithionite
$S_3O_6^{2-}$	trithionate
SO_4^{2-}	sulfate

REFERENCES

1. M. Whiteman, P. G. Winyard, *Expert Rev. Clin. Pharmacol.* **2011**, *4*, 1751–2441.
2. A. Aroca, C. Gotor, L. C. Romero, *Front. Plant Sci.* **2018**, *9*, 1369, doi: 10.3389/fpls.2018.01369.
3. V. P. Utgikar, S. M. Harmon, N. Chaudhary, H. H. Tabak, R. Govind, J. R. Haines, *Environ. Toxicol.* **2002**, *17*, 40–48.
4. G. T. Townsend, K. Ramanand, J. M Suflita, *Appl. Environ. Microbiol.* **1997**, *63*, 2785–2791.
5. K. Craig, "A Review of the Chemistry, Pesticide Use, and Environmental Fate of Sulfur Dioxide", *California Rev. Environ. Contam. Toxicol.* **2019**, *246*, 33–64.
6. K. L. Cook, T. R. Whitehead, C. Spence, M. A. Cotta, *Anaerobe* **2008**, *14*, 172–180.
7. Y. S. Do, T. M. Schmidt, J. A. Zahn, E. S. Boyd, A. de la Mora, A. A. DiSpirito, *Appl. Environ. Microbiol.* **2003**, *69*, 1710–20.
8. S. S. Schiffman, E. A. Miller, M. S. Suggs, B. G. Graham, *Brain Res. Bull.* **1995**, *37*, 369–375.

9. G. B. Michaels, J. T. Davidson, H. D. Peck, *Biochem. Biophys. Res. Commun.* **1970**, *39*, 321–328.
10. C. Li, H. D. Peck, A. E. Przybyla, *Gene* **1987**, *53*, 227–234.
11. L. M. Siegel, M. J. Murphy, H. Kamin, *J. Biol. Chem.* **1973**, *248*, 251–264.
12. J. P. Lee, J. LeGall, H. D. Peck, *J. Bacteriol.* **1973**, *115*, 529–542.
13. C. Dahl, S. Engels, A. S. Pott-Sperling, A. Schulte, J. Sander, Y. Lubbe, O. Deuster, D. C. Brune, *J. Bacteriol.* **2005**, *187*, 1392–1404.
14. C. G. Friedrich, D. Rother, F. Bardischewsky, A. Quentmeier, J. Fischer, *Appl. Environ. Microbiol.* **2001**, *67*, 2873–2882.
15. C. G. Friedrich, F. Bardischewsky, D. Rother, A. Quentmeier, J. Fischer, *Curr. Opin. Microbiol.* **2005**, 8, 253–259.
16. B. Hermann, M. Kern, L. La Pietra, J. Simon, O. Einsle, *Nature* **2015**, *520*, 706–709.
17. E. F. Johnson, B. Mukhopadhyay, *Appl. Environ. Microbiol.* **2008**, 74, 3591–3595.
18. S. Shirodkar, S. Reed, M. Romine, D. Saffarini, *Environ. Microbiol.* **2011**, *13*, 108–115.
19. A. A. Trofimov, K. M. Polyakov, K. M Boyko, T. V. Tikhonova, T. N. Safonova, A. V. Tikhonov, A. N. Popov, V. O. Popov, *Acta Crystallogr. D Biol. Crystallogr.* **2010**, *66*, 1043–1047.
20. P. A. Janick, L. M. Siegel, *Biochemistry*, **1982**, *21*, 3538–3547.
21. D. E. McRee, D. C. Richardson, J. S. Richardson, L. M. Siegel, *J. Biol. Chem.* **1986**, *261*, 10277–10281.
22. B. R. Crane, L. M. Siegel, E. D. Getzoff, *Science*, **1995**, *270*, 59–67.
23. M. R. Cepeda, L. McGarry, J. M. Pennington, J. Krzystek, M. E. Stroupe, *Biochim. Biophys. Acta* **2018**, *1866*, 933–940.
24. K. W. Smith, M. E. Stroupe, *Biochemistry* **2012**, *51*, 9857–9868.
25. K. D. Bewley, K. E. Ellis, M. A. Firer-Sherwood, S. J. Elliott, *Biochim. Biophys. Acta* **2013**, *1827*, 938–948.
26. S. Horrell, D. Kekilli, R. W. Strange, M. A. Hough, *Metallomics* **2017**, *9*, 1470–1482.
27. X. Li, S. Lan, Z. Zhu, C. Zhang, G. Zeng, Y. Liu, W. Cao, B. Song, H. Yang, S. Wang, S. Wu, *Ecotoxicol. Environ. Saf.* **2018**, *158*, 162–170.
28. F. Ramel, A. Amrani, L. Pielle, O. Lamrabet, G. Voordouw, N. Seddiki, D. Brethes, M. Company, A. Dolla, G. Brasseur, *Microbiology* **2013**, *159*, 2663–2673.
29. H. Sass, H. Cypionka, H. Babenzien, *FEMS Microbiol. Ecol.* **1997**, *22*, 245–255.
30. H. F. Castro, N. H. Williams, A. Ogram, *FEMS Microbiol. Ecol.* **2000**, *31*, 1–9.
31. A. Hussain, A. Hasan, A. Javid, J. Qazi, *3 Biotech* **2016**, *6*, 119. doi: 10.1007/s13205-016-0437-3.
32. O. Hao, J. Chen, L. Huang, R. Buglass, *Crit. Rev. Environ. Sci. Technol.* **1996**, *26*, 155–187.
33. M. K. Stoeva, J. D. Coates, *Microbiology* **2019**, *165*, 254–269.
34. A. Marietou, H. Røy, B. B. Jørgensen, K. U. Kjeldsen, *Front. Microbiol.* **2018**, **9**, 309. https://www.frontiersin.org/article/10.3389/fmicb.2018.00309.
35. P. Loughlin, M. C. Shelden, M. L. Tierney, S. M. Howitt, *Cell Biochem. Biophys.* **2002**, *36*, 183–190.
36. R. Nie, S. Stark, J. Symersky, R. S. Kaplan, M. Lu, *Nat. Commun.* **2017**, *8*, 15009. https://doi:10.1038/ncomms15009.
37. A. Sirko, M. Zatyka, E. Sadowy, D. Hulanicka, *J. Bacteriol.* **1995**, *177*, 4134–4136.
38. L. Zhang, W. Jiang, J. Nan, J. Almqvist, Y. Huang, *Biochim. Biophys. Acta* **2014**, *1838*, 1809–1816.
39. M. Hryniewicz, A. Sirko, A. Pałucha, A. Böck, D. Hulanicka, *J. Bacteriol.* **1990**, *172*, 3358–3366.
40. A. Sirko, M. Hryniewicz, D. Hulanicka, A. Böck, *J. Bacteriol.* **1990**, *172*, 3351–3357.
41. L. Prioretti, B. Gontero, R. Hell, M. Giordano, *Front. Plant Sci.* **2014**, *5*, 597.

42. F. Grein, A. R. Ramos, S. S. Venceslau, I. A. Pereira, *Biochim. Biophys. Acta* **2013**, *1827*, 145–160.
43. P. M. Matias, I. A. Pereira, C. M. Soares, M. A. Carrondo, *Prog. Biophys. Mol. Biol.* **2005**, *89*, 292–329.
44. J. Odom, H. D. Peck, *FEMS Microbiol. Lett.* **1981**, *12*, 47–50.
45. I. A. Pereira, A. R. Ramos, F. Grein, M. C. Marques, S. M. da Silva, S. S. Venceslau, *Front. Microbiol.* **2011**, *2*, 69. doi: 10.3389/fmicb.2011.00069.
46. D. C. Johnson, D. R. Dean, A. D. Smith, M. K. Johnson, *Annu. Rev. Biochem.* **2005**, *74*, 247–281.
47. A. E. Pegg, M. Boosalis, L. Samson, R. C. Moschel, T. L. Byers, K. Swenn, M. E. Dolan, *Biochemistry* **1993**, *32*, 11998–12006.
48. D. Quig, *Altern. Med. Rev.* **1998**, *3*, 262–270.
49. H. Fang, J. Kang, D. Zhang, *Microb. Cell Fact.* **2017**, *16*, 15; doi: 10.1186/s12934-017-0631-y.
50. S. D. Copley, J. K. Dhillon, *Genome Biol.* **2002**, *3*, 0025.1–16.
51. M. L. Reniere, *J. Bacteriol.* **2018**, *200*, e00128–18; doi: 10.1128/JB.00128-18.
52. P. J. Stoffyn, *J. Am. Oil Chem. Soc.* **1966**, *43*, 69–74.
53. S. J. Kang, P. Cresswell, *Nat. Immunol.* **2004**, *5*, 175–181.
54. M. Gilleron, S. Stenger, A. Mazorra, F. Wittke, S. Mariotti, G. Böhmer, J. Prandi, L. Mori, G. Puzo, G. De Libero, *J. Exp. Med.* **2004**, *199*, 649–659.
55. C. A. James, K. K. Q. Yu, M. Gilleron, J. Prandi, V. R. Yedulla, Z. Z. Moleda, E. Diamanti, M. Khan, V. K. Aggarwal, J. F. Reijneveld, P. Reinink, S. Lenz, R. O. Emerson, T. J. Scriba, M. N. T. Souter, D. I. Godfrey, D. G. Pellicci, D. B. Moody, A. J. Minnaard, C. Seshadri, I. VanRhijn, *Cell Chem. Biol.* **2018**, *25*, 392–402.e14.
56. C. Benning, *Annu. Rev. Plant Physiol. Plant Mol. Biol.* **1998**, *49*, 53–75.
57. S. Güler, A. Seeliger, H. Härtel, G. Renger, C. Benning, *J. Biol. Chem.* **1996**, *271*, 7501–7507.
58. J. Marcus, S. Honigbaum, S. Shroff, K. Honke, J. Rosenbluth, J. L. Dupree, *Glia* **2006**, *53*, 372–381.
59. M. J. Murphy, L. M. Siegel, H. Kamin, D. V. DerVartanian, J. P. Lee, J. LeGall, H. D. Peck, *Biochem. Biophys. Res. Commun.* **1973**, *54*, 82–88.
60. M. J. Murphy, L. M. Siegel, H. Kamin, D. Rosenthal, *J. Biol. Chem.* **1973**, *248*, 2801–2814.
61. M. J. Murphy, L. M Siegel, *J. Biol. Chem.* **1973**, *248*, 6911–6919.
62. L. M. Siegel, P. S. Davis, M. J. Murphy, *Biochem. J.* **1977**, *167*, 669–674.
63. D. G. Hodgkin, J. Pickworth, J. H. Robertson, K. N. Trueblood, R. J. Prosen J. G. White, *Nature*, **1955**, *176*, 325–328.
64. H. A. Dailey, T. A. Dailey, D. Gerdes, D. Jahn, M. Jahn, M. R. O'Brian, M. J. Warren, *Microbiol. Mol. Biol. Rev.* **2017**, *81*, 1–62.
65. J. Kim, D. Woo, D. C. Rees, *Biochemistry* **1993**, *32*, 7104–7115.
66. P. Hänzelmann, H. Dobbek, L. Gremer, R. Huber, O. Meyer, *J. Mol. Biol.* **2000**, *301*, 1221–1235.
67. C. L. Drennan, J. Heo, M. D. Sintchak, E. Schreiter, P. W. Ludden, *Proc. Natl. Acad. Sci.* **2001**, *98*, 11973–11978.
68. S. Caffarri, T. Tibiletti, R. C. Jennings, S. Santabarbara, *Curr. Protein Pept. Sci.* **2014**, *15*, 296–331.
69. L. M. Siegel, D. C. Rueger, M. J. Barber, R. J. Krueger, N. R. Orme-Johnson, W. H. Orme-Johnson, *J. Biol. Chem.* **1982**, *257*, 6343–6350.
70. M. Wagner, A. J. Roger, J. L. Flax, G. A. Brusseau, D. A. Stahl, *J. Bacteriol.* **1998**, *180*, 2975–2982.
71. D. A. Stahl, S. Fishbain, M. Klein, B. J. Baker, M. Wagner, *Antonie Van Leeuwenhoek*, **2002**, *81*, 189–195.

72. J. R. Cort, U. Selan, A. Schulte, F. Grimm, M. A. Kennedy, C. Dahl, *J. Mol. Biol.* **2008**, *382*, 692–707.
73. T. F. Oliveira, C. Vonrhein, P. M. Matias, S. S. Venceslau, I. A. Pereira, M. Archer, *J. Biol. Chem.* **2008**, *283*, 34141–34149.
74. A. Schiffer, K. Parey, E. Warkentin, K. Diederichs, H. Huber, K. O. Stetter, P. M. Kroneck, U. Ermler, *J. Mol. Biol.* **2008**, *379*, 1063–1074.
75. Y. C. Hsieh, M. Y. Liu, V. C. Wang, Y. L. Chiang, E. H. Liu, W. G. Wu, S. I. Chan, C. J. Chen, *Mol. Microbiol.* **2010**, *78*, 1101–1116.
76. T. F. Oliveira, E. Franklin, J. P. Afonso, A. R. Khan, N. J. Oldham, I. A. Pereira, M. Archer, *Front. Microbiol.* **2011**, *2*, 71. doi: 10.3389/fmicb.2011.00071.
77. I. Moura, J. LeGall, A. R. Lino, H. D. Peck, G. Fauque, A. V. Xavier, D. V. DerVartanian, J. J. G. Moura, B. H. Huynh, *J. Am. Chem. Soc.* **1988**, *110*, 1075–1082.
78. C. Dahl, N. M. Kredich, R. Deutzmann, H. G. Trüper, *J. Gen. Microbiol.* **1993**, *139*, 1817–1828.
79. G. Fauque, A. R. Lino, M. Czechowski, L. Kang, D. V. DerVartanian, J. J. Moura, J. LeGall, I. Moura, *Biochim. Biophys. Acta* **1990**, *1040*, 112–118.
80. E. C. Hatchikian, J. G. Zeikus, *J. Bacteriol.* **1983**, *223*, 79–80.
81. P. A. Trudinger, *J. Bacteriol.* **1970**, *104*, 158–170.
82. J. M. Akagi, V. Adams, *J. Bacteriol.* **1973**, *116*, 392–396.
83. B. R. Crane, E. D. Getzoff, *Curr. Opin. Struct. Biol.* **1996**, *6*, 744–756.
84. S. S. Venceslau, Y. Stockdreher, C. Dahl, I. A. Pereira, *Biochim. Biophys. Acta* **2014**, *1837*, 1148–1164.
85. A. A. Santos, S. S. Venceslau, F. Grein, W. D. Leavitt, C. Dahl, D. T. Johnston, I. A. Pereira, *Science* **2015**, *350*, 1541–1545.
86. N. Mizuno, G. Voordouw, K. Miki, A. Sarai, Y. Higuchi, *Structure* **2003**, *11*, 1133–1140.
87. H. Laue, M. Friedrich, J. Ruff, A. M, Cook, *J. Bacteriol.* **2001**, *183*, 1727–1733.
88. W. M. Hipp, A. S. Pott, N. Thum-Schmitz, I. Faath, C. Dahl, H. G. Trüper, *Microbiology* **1997**, *143*, 2891–2902.
89. E. Ben-Dov, A. Brenner, A. Kushmaro, *Microb. Ecol.* **2007**, *54*, 439–451.
90. U. Swamy, M. Wang, J. N. Tripathy, S. K. Kim, M. Hirasawa, D. B. Knaff, J. P. Allen, *Biochemistry* **2005**, *44*, 16054–16063.
91. I. Askenasy, D. T. Murray, R. M. Andrews, V. N. Uversky, H. He, M. E. Stroupe, *Biochemistry* **2018**, *57*, 3764–3772.
92. L. M. Siegel, P. S. Davis, *J. Biol. Chem.* **1974**, *249*, 1587–1598.
93. I. Askenasy, J. M. Pennington, Y. Tao, A. G. Marshall, N. L. Young, W. Shang, M. E. Stroupe, *J. Biol. Chem.* **2015**, *290*, 19319–19333.
94. R. Schnell, T. Sandalova, U. Hellman, Y. Lindqvist, G. Schneider, *J. Biol. Chem.* **2005**, *280*, 27319–27328.
95. J. Y. Kim, M. Nakayama, H. Toyota, G. Kurisu, T. Hase, *J. Biochem.* **2016**, *160*, 101–109.
96. M. Hirasawa, J. N. Tripathy, R. Somasundaram, M. K. Johnson, M. Bhalla, J. P. Allen, D. B. Knaff, *Mol. Plant* **2009**, *2*, 407–415.
97. L. S. Privalle, C. T. Privalle, N. J. Leonardy, H. Kamin, *J. Biol. Chem.* **1985**, *260*, 14344–14350.
98. K. W. Joy, R. H. Hageman, *Biochem. J.* **1966**, *100*, 263–273.
99. B. H. Huynh, L. Kang, D. V. DerVartanian, H. D. Peck, J. LeGall, *J. Biol. Chem.* **1984**, *259*, 15373–15376.
100. H. L. Drake, J. M. Akagi, *J. Bacteriol.* **1977**, *132*, 132–138.
101. T. Underwood-Lemons, I. Moura, K. T. Yue, *Biochim. Biophys. Acta* **1993**, *1157*, 275–284.
102. B. R. Crane, L. M. Siegel, E. D. Getzoff, *Biochemistry*, **1997**, *36*, 12120–12137.

103. B. R. Crane, L. M. Siegel, E. D. Getzoff, *Biochemistry* **1997**, *36*, 12101–12119.
104. A. F. Arendsen, M. F. Verhagen, R. B. Wolbert, A. J. Pierik, A. J. Stams, M. S. Jetten, W. R. Hagen, *Biochemistry* **1993**, *32*, 10323–10330.
105. A. J. Pierik, W. R. Hagen, *Eur. J. Biochem.* **1991**, *195*, 505–516.
106. L. Kang, J. LeGall, A. T. Kowal, M. K. Johnson, *J. Inorg. Biochem*. **1987**, *30*, 273–290.
107. P. A. Janick, L. M. Siegel, *Biochemistry* **1983**, *22*, 504–515.
108. Y. Seki, N. Sogawa, M. Ishimoto, *J. Biochem.* **1981**, *90*, 1487–1492.
109. J. Liu, S. Chakraborty, P. Hosseinzadeh, Y. Yu, S. Tian, I. Petrik, A. Bhagi, Y. Lu, *Chem. Rev.* **2014**, *114*, 4366–4469.
110. J. A. Christner, E. Munck, P. A. Janick, L. M. Siegel, *J. Biol. Chem.* **1981**, *256*, 2098–2101.
111. J. A. Christner, E. Munck, P. A. Janick, L. M. Siegel, *J. Biol. Chem.* **1983**, *258*, 11147–11156.
112. J. A. Christner, P. A. Janick, L. M. Siegel, E. Munck, *J. Biol. Chem.* **1983**, *258*, 11157–11164.
113. S. H. Han, J. F. Madden, R. G. Thompson, S. H. Strauss, L. M. Siegel, T. G. Spiro, *Biochemistry* **1989**, *28*, 5461–5471.
114. S. H. Han, J. F. Madden, L. M. Siegel, T. G. Spiro, *Biochemistry* **1989**, *28*, 5477–5485.
115. J. F. Madden, S. J. Han, L. M. Siegel, T. G. Spiro, *Biochemistry* **1989**, *28*, 5471–5477.
116. J. Kaufman, L. M. Siegel, L. D. Spicer, *Biochemistry* **1993**, *32*, 8782–8791.
117. J. Kaufman, L. D. Spicer, L. M. Siegel, *Biochemistry* **1993**, *32*, 2853–2867.
118. J. F. Cline, P. A. Janick, L. M. Siegel, B. M. Hoffman, *Biochemistry* **1985**, *24*, 7942–7947.
119. J. F. Cline, P. A. Janick, L. M. Siegel, B. M. Hoffman, *Biochemistry* **1986**, *25*, 4647–4654.
120. P. A. Janick, D. C. Rueger, R. J. Krueger, M. J. Barber, L. M. Siegel, *Biochemistry* **1983**, *22*, 396–408.
121. L. J. Young, L. M. Siegel, *Biochemistry* **1988**, *27*, 4991–4999.
122. L. M. Siegel, P. S. Davis, H. Kamin, *J. Biol. Chem.* **1974**, *249*, 1572–1586.
123. R. J. Krueger, L. M. Siegel, *Biochemistry* **1982**, *21*, 2892–2904.
124. S. Nakano, M. Takahashi, A. Sakamoto, H. Morikawa, K. Katayanagi, *Proteins* **2012**, *80*, 2035–2045.
125. S. Nakano, M. Takahashi, A. Sakamoto, H. Morikawa, K. Katayanagi, *Chem. Biodivers.* **2012**, *9*, 1989–1999.
126. K. Parey, E. Warkentin, P. M. Kroneck, U. Ermler, *Biochemistry* **2010**, *49*, 8912–8921.
127. S. G. Mayhew, *Eur. J. Biochem.* **1978**, *85*, 535–547.
128. R. R. Karkhoff-Schweizer, M. Bruschi, G. Voordouw, *Eur. J. Biochem.* **1993**, *211*, 501–507.
129. S. S. Venceslau, J. R. Cort, E. S. Baker, R. K. Chu, E. W. Robinson, C. Dahl, L. M. Saraiva, I. A. Pereira, *Biochem. Biophys. Res. Commun.* **2013**, *441*, 732–736.
130. M. Nakayama, T. Akashi, T. Hase, *J. Inorg. Biochem.* **2000**, *82*, 27–32.
131. E. J. Faeder, P. S. Davis, L. M. Siegel, *J. Biol. Chem.* **1974**, *249*, 1599–1609.
132. A. M. Tavolieri, D. T. Murray, I. Askenasy, J. M. Pennington, L. McGarry, C. B. Stanley, M. E. Stroupe, *J. Struct. Biol.* **2019**, *205*, 170–179.
133. M. Zeghouf, M. Fontecave, D. Macherel, J. Coves, *Biochemistry* **1998**, *37*, 6114–6123.
134. S. Bar-Nun, G. Kreibich, M. Adesnik, L. Alterman, M. Negishi, D. D. Sabatini, *Proc. Natl. Acad. Sci.* **1980**, *77*, 965–969.
135. T. Iyanagi, C. Xia, J. J. Kim, *Arch. Biochem. Biophys.* **2012**, *528*, 72–89.
136. J. Y. Wu, L. M. Siegel, N. M. Kredich, *J. Bacteriol.* **1991**, *173*, 325–333.
137. M. Zeghouf, M. Fontecave, J. Coves, *J. Biol. Chem.* **2000**, *275*, 37651–3766.

138. L. Champier, N. Sibille, B. Bersch, B. Brutscher, M. Blackledge, J. Coves, *Biochemistry* **2002**, *41*, 3770–3780.
139. S. Bali, A. D. Lawrence, S. A. Lobo, L. M. Saraiva, B. T. Golding, D. J. Palmer, M. J. Howard, S. J. Ferguson, M. J. Warren, *Proc. Natl. Acad. Sci.* **2011**, *108*, 18260–18265.
140. S. Bali, D. J. Palmer, S. Schroeder, S. J. Ferguson, M. J. Warren, *Cell Mol. Life Sci.* **2014**, *71*, 2837–2863.
141. L. J. Young, L. M. Siegel, *Biochemistry* **1988**, *27*, 5984–5990.
142. M. J. Warren, N. J. Stolowich, P. J. Santander, C. A. Roessner, B. A. Sowa, A. I. Scott, *FEBS Lett.* **1990**, *261*, 76–80.
143. M. J. Warren, C. A. Roessner, P. J. Santander, A. I. Scott, *Biochem. J.* **1990**, *265*, 725–729.
144. S. Storbeck, S. Saha, J. Krausze, B. U. Klink, D. W. Heinz, G. Layer, *J. Biol. Chem.* **2011**, *286*, 26754–26767.
145. J. Vévodová, R. M. Graham, E. Raux, H. L. Schubert, D. I. Roper, A. A. Brindley, I. A. Scott, C. A. Roessner, N. P. Stamford, M. E. Stroupe, E. D. Getzoff, M. J. Warren, K. S. Wilson, *J. Mol. Biol.* **2004**, *344*, 419–433.
146. M. E. Stroupe, H. K. Leech, D. S. Daniels, M. J. Warren, E. D. Getzoff, *Nat. Struct. Biol.* **2003**, *10*, 1064–1073.
147. P. H. Rehse, T. Kitao, T. H. Tahirov, *Acta Crystallogr. D Biol. Crystallogr.* **2005**, *61*, 913–919.
148. H. L. Schubert, E. Raux, A. A. Brindley, H. K. Leech, K. S. Wilson, C. P. Hill, M. J. Warren, *EMBO J.* **2002**, *21*, 2068–2075.
149. H. L. Schubert, R. S. Rose, H. K. Leech, A. A. Brindley, C. P. Hill, S. E. Rigby, M. J. Warren, *Biochem. J.* **2008**, *415*, 257–263.
150. E. Raux, H. K. Leech, R. Beck, H. L. Schubert, P. J. Santander, C. A. Roessner, A. I. Scott, J. H. Martens, D. Jahn, C. Thermes, A. Rambach, M. J. Warren, *Biochem. J.* **2003**, *370*, 505–516.
151. M. J. Warren, E. L. Bolt, C. A. Roessner, A. I. Scott, J. B. Spencer, S. C. Woodcock, *Biochem. J.* **1994**, *302*, 837–844.
152. S. C. Woodcock, E. Raux, F. Levillayer, C. Thermes, A. Rambach, M. J. Warren, *Biochem. J.* **1998**, *330*, 121–129.
153. H. K. Leech, E. Raux-Deery, P. Heathcote, M. J. Warren, *Biochem. Soc. Trans.* **2002**, *30*, 610–613.
154. E. Raux, T. McVeigh, S. E. Peters, T. Leustek, M. J. Warren, *Biochem. J.* **1999**, *338*, 701–708.
155. J. L. Small, S. W. Park, B. D. Kana, T. R. Ioerger, J. C. Sacchettini, S. Ehrt, *MBio.* **2013**, *4*, e00475–13.
156. D. Andersson, P. Hammarstrom, U. Carlsson, *Biochemistry* **2001**, *40*, 2653–2661.
157. V. E. Bychkova, A. E. Dujsekina, S. I. Klenin, E. I. Tiktopulo, V. N. Uversky, O. B. Ptitsyn, *Biochemistry* **1996**, *35*, 6058–6063.
158. S. S. Leal, C. M. Gomes, *Proteins* **2007**, *68*, 606–616.
159. H. Paritala, K. S. Carroll, *Infect. Disord. Drug Targets* **2013**, *13*, 85–115.
160. E. N. Mirts, I. D. Petrik, P. Hosseinzadeh, M. J. Nilges, Y. Lu, *Science* **2018**, *361*, 1098–1101.
161. Y. Bai, T. Chen, T. Happe, Y. Lu, A. Sawyer, *Metallomics* **2018**, *10*, 1038–1052.
162. B. Blanc, C. Gerez, S. Ollagnier de Choudens, *Biochim. Biophys. Acta* **2015**, *1853*, 1436–1447.
163. R. Puglisi, A. Pastore, *FEBS Lett.* **2018**, *592*, 4011–4019.
164. F. Sievers, D. G. Higgins, *Protein Sci.* **2018**, *27*, 135–145.

11

Nickel, Iron, Sulfur Sites

Yulia Ilina, Berta M. Martins, Jae-Hun Jeoung, and Holger Dobbek

Institut für Biologie, Strukturbiologie/Biochemie, Humboldt-Universität zu Berlin,
D-10099 Berlin, Germany
<holger.dobbek@hu-berlin.de>

Metal Ions in Life Sciences, Volume 20 Guest Editors: Martha E. Sosa Torres and Peter M. H. Kroneck
Series Editors: Astrid Sigel, Eva Freisinger, and Roland K. O. Sigel

https://doi.org/10.1515/9783110589757-011

Abstract: Enzymes relying on the interplay of nickel, iron, and sulfur in their active sites are used by prokaryotes to catalyze reactions driving the global carbon and hydrogen cycles. The three enzymes, [NiFe] hydrogenases, Ni,Fe-containing carbon monoxide dehydrogenases and acetyl-CoA synthases share an ancient origin possibly derived from abiotic processes. Although their active sites have different compositions and assemble Ni, Fe, and S in different ways and for different purposes, they share a central role of Ni in substrate binding and activation, with sulfur linking the Ni ion to one or more Fe ions, which, although indispensable for function, supports the catalytic process in less understood ways. The review gives a short overview on the properties of the three individual enzymes highlighting their parallels and differences.

Keywords: acetogenesis · acetyl-CoA synthase · carbon monoxide dehydrogenase · electron transfer · hydrogenase · iron-sulfur clusters · small molecule activation

1. INTRODUCTION

Three enzymes share the property that their active sites are formed by Ni, Fe, and S: [NiFe] hydrogenases, Ni,Fe-containing carbon monoxide dehydrogenases (CODH), and acetyl-CoA synthases (ACS). These three enzymes will be presented and compared. All three enzymes have recently been reviewed individually. Overviews on the structure and function of CODH and ACS can for example be found in [1–3], while recent overviews on [NiFe] hydrogenases may be found in [4–7]. In the following the enzymes will only be discussed as far as needed to understand the function of Ni, Fe, and S, to be able to appreciate the similarities and differences in the roles they are playing in the respective mechanisms.

1.1. Biological Nickel

Elemental nickel is typically found in enrichments together with iron, where the iron exceeds the nickel, such as in iron meteorites and the earth's core. Within the lithosphere, nickel is found in nickel ores and can be isolated in high purity by the classical Mond process [8]. The predominant oxidation state of nickel is Ni(II), but oxidation states ranging from –1 to +4 have been reported [8]. Ni(II) has a $3d^8$ electron configuration and is isoelectronic to but substantially less nucleophilic than Co(I). Ni(II) is also the predominant form found in enzymes and favors coordination by four, five or six ligands. The prevalent ligand coordination is with four ligands either arranged tetragonally, producing a paramagnetic species or square-planar, resulting in diamagnetic Ni(II) complexes.

Compared to other biological metals, the role of Ni as constituent of active sites was discovered only lately, which is to some extent due to the lack of characteristic and intense optical (or electronic, or UV-vis) transitions by Ni, as well as by the oxygen-sensitivity of some of the enzymes. Several enzymes have been detected to contain nickel, such as glyoxylase I, the acidoreductone dioxygenase, and the quercetinase QueD, are also active with other metals, whereas the apoenzyme of ureases, class IV superoxide dismutases, methyl-

coenzyme M reductase, and lactate racemase strictly depend on the incorporation of Ni or a Ni-containing cofactor to gain activity [9].

To achieve their function, the organisms need to take up nickel from the environment, which is supported by metallophores and dedicated transporters through the inner and outer membrane, as well as Ni storage proteins, Ni chaperones, and gene regulators [10, 11].

The three Ni enzymes, [NiFe] hydrogenase, CODH, and ACS, strictly depend on Ni and its integration into a cofactor containing in addition iron and sulfur. Although the number of enzymes depending on Ni is still small compared to those containing iron, zinc, copper or molybdenum, Ni-dependent enzymes drive the global hydrogen, carbon and nitrogen cycles by catalyzing critical small molecule activation steps [9].

[NiFe] hydrogenase and the Ni-containing CODH and ACS are considered evolutionarily old and phylogenetic analysis suggests their presence in the last universal common ancestor LUCA [12, 13]. Their active sites may even pre-date the enzymes by being present in form of precipitates at membranes around the hydrothermal vents of the Hadean ocean (~4 billion years ago). Ni-, Fe-, and S-containing minerals such as mackinawite, nickel-bearing greigite, and violarite have similar coordination as the active sites discussed in this chapter and may exhibit similar properties as the enzymes at defect sites in the minerals [12, 14].

1.2. Nickel Together with Iron and Sulfur

Nickel and iron share many characteristics. As middle 3d transition metals, both are found in at least three different oxidation states in enzymes (Fe: II, III, IV; Ni: I, II, III), both readily bind to thiolates and sulfides but are rather versatile and can also form complexes with biological anionic ligands such as carboxylates or the large organic tetrapyrroles (Fe: heme cofactors; Ni: cofactor F_{430}), as well as with neutral ligands like the side chain of histidine. Both transition metals may form octahedral or tetrahedral complexes with their ligands producing diamagnetic or paramagnetic species.

Despite these parallels, the biological distribution and physiological roles of iron and nickel are surprisingly different. While iron is the fourth most abundant element in the lithosphere, nickel is much less abundant but better bioavailable as it is stable in solution under oxic, aqueous conditions, conditions under which Fe(II) is oxidized and forms Fe(III) oxide complexes of negligible solubility, quenching its bioavailability.

The only currently known enzymes in which Ni forms a complex with Fe are the three enzymes covered in this chapter. All contain Ni only in the active site, while Fe and S are also found outside the active site forming gateways for the electrons turned over in the overall process.

The three enzymes all have Ni and Fe linked by sulfur-containing ligands. While in our mechanistic proposals, we typically focus on the role of the two transition metals, the role of sulfur can hardly be overestimated and may be appreciated once we compare it with the lighter and more abundant chalcogen oxygen [15]. Due to its high electronegativity, oxygen tends to pull electrons to it

forming (protonated) oxides, alcoholates or carboxylates with strongly polarized bonds, thus stabilizing higher oxidation states of the coordinated metal. In contrast, sulfur is more versatile and may form bonds of varying levels of covalency, depending and reacting to the oxidation state of the metal. By shifting towards a more ionic bond character once the metal becomes reduced, the sulfur-containing ligands take some of the charge, acting as an electronic buffer [16]. Also in contrast to oxygen, with sulfur the 3d-orbitals start to play a role, opening new possibilities of metal-ligand interactions.

2. THE [NiFe] SITE OF HYDROGENASES

2.1. Hydrogenases

Hydrogenases catalyze the reversible oxidation of H_2 to two protons and two electrons. Research on hydrogenases is motivated not only by their central role in various bacterial metabolisms, but also for developing a sustainable economy using dihydrogen as an energy vector [17]. To reach an H_2-based economy, several obstacles have still to be overcome and breakthroughs in our understanding on how H_2 may be efficiently produced, transported, stored, and oxidized are needed, motivating the study of nature's solutions to these quests [18].

2.1.1. The Three Types of Hydrogenases

Three principal types of non-homologous hydrogenases are found in nature with different active site architectures, folds, and physiological roles [19]. Hydrogenases, called [FeFe] hydrogenases, containing a dinuclear Fe-site with CO and CN^- ligands together with a [4Fe-4S] cluster are found in several bacteria and eukaryotes [20–22], while [Fe] hydrogenases containing a single Fe coordinated by a cysteine side chain, 2-pyridinol, and two CO ligands are found predominantly in methanogenic archaea, but recently also in bacteria [23]. The third group, [NiFe] hydrogenases contain a dinuclear Ni,Fe center in the active site and are found wide-spread in prokaryotes. Several recent reviews may be consulted for a general information on hydrogenases [19, 24, 25], or more specifically on the [NiFe] hydrogenases [4–7]. Here, we will only focus on the structure and function of [NiFe] hydrogenases.

2.1.2. The [NiFe] Hydrogenase Groups

Of the three types of hydrogenases, the [NiFe] hydrogenases are the most numerous and are connected to various metabolic functions [26, 27]. Wide-scaled bioinformatic studies indicate that the [NiFe] hydrogenases may be divided into four main groups, of which each has a conserved sequence pattern [27]. Group 1 covers the uptake [NiFe] hydrogenases, including membrane-bound respiratory hydrogenases oxidizing H_2, while group 2 contains cyanobacterial uptake hydro-

genases and regulatory hydrogenases. Group 3 comprises bidirectional heteromultimeric cytoplasmic [NiFe] hydrogenases, where "bidirectional" indicates that their physiological function may be associated with proton reduction or H_2 oxidation. Here, we also find hydrogenases from methanogenic archaea such as *Methanosarcina barkeri* allowing H_2 to be used as energetic intermediate produced from oxidizing reduced deazaflavin-type cofactor F_{420} and consumed by a membrane-bound [NiFe] hydrogenase [28–31]. Finally, group 4 contains the energy-conserving membrane-associated hydrogenases, which are homologous to the complex I of the respiratory chain, containing the CO-induced hydrogenases allowing to couple CO oxidation with proton-reduction while generating a proton-motive force [32, 33].

2.1.3. Overall Structure of [NiFe] Hydrogenases

[NiFe] hydrogenases have different subunit compositions, with the common motif of combining the [NiFe] active site within a larger subunit with one or more subunits forming an electron transfer chain supported by iron-sulfur clusters and eventually a flavin as the terminal electron donor or acceptor. The [NiFe] hydrogenases from sulfate-reducing bacteria, which have been intensively studied, provide us with the typical core motif of [NiFe] hydrogenases.

The crystal structure of the [NiFe] hydrogenase from the sulfate-reducing *Desulfovibrio gigas* revealed the principal architecture of the small and large subunits with the large subunit carrying the bimetallic active site and the small subunit binding three iron-sulfur clusters, two [4Fe-4S] clusters with a [3Fe-4S] cluster between them (Figure 1, left side). Based on distance to the [NiFe] site, the Fe-S clusters are denoted as proximal, medial, and distal [34]. In addition to the transport of two electrons, two protons need to be transferred, which may be helped by a chain of conserved histidine and glutamate residues [34].

Substantially larger assemblies containing longer electron transfer chains are known, for example in the F_{420}-reducing hydrogenases (group 3), forming a dodecameric complex of 1.2 MDa, whose structure was initially determined by cryo-electron microscopy [35, 36], before being resolved at 1.7 Å by X-ray crystallography [37], but also in the soluble hydrogenase (group 3) [38] (Figure 1, right side). Different kinds of oligomers are also found to electronically connect two or more active sites, such as in actinobacterial-type hydrogenases (group 1) where a V-shaped electron transfer pathway has been found [39], which may contribute to O_2 tolerance [40].

In addition to different lengths of the electron transfer chains, variants in the type of iron-sulfur clusters occur. In the [NiFeSe] hydrogenase of *Desulfomicrobium baculatum*, a [4Fe-4S] cluster is found where some [NiFe] hydrogenases harbor the [3Fe-4S] cluster [41], whose exchange against a [4Fe-4S] cluster has only slight effects on turnover [42]. Still, the type and redox potential of the electron transfer chain remote from the active site may tweak the catalytic bias of hydrogen oxidation versus proton reduction [43].

The importance of the electron transfer chain for the catalytic properties are paramount for the monophyletic and wide-spread oxygen-tolerant group 1

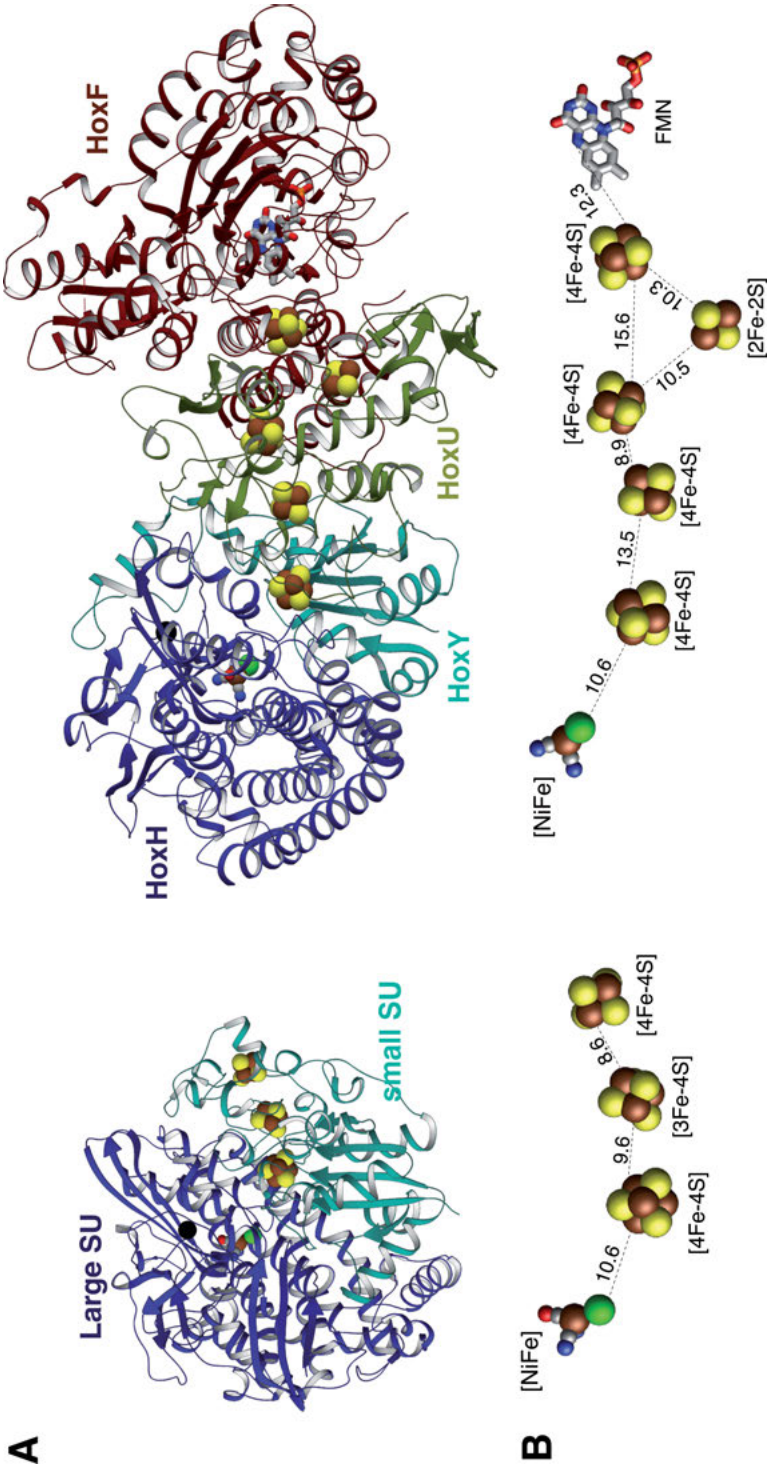


Figure 1. Overall structure of [NiFe] hydrogenases. (**A**) Cartoon representation of the prototype [NiFe] hydrogenase from *D. gigas* (PDB: 1YQ9 left side) [34] and the soluble [NiFe] hydrogenase from *H. thermoluteolus* (PDB: 5XF9 right side) [38]. The protein backbone is shown as ribbon with domains individually colored. Atoms are shown as spheres (dark brown for Fe, green for Ni, blue for N, red for O). (**B**). Electron transfer chain between the [NiFe] active site and its solvent-accessible endmost member ([4Fe-4S] cluster for *D. gigas* and FMN for *H. thermoluteolus*). Atoms are shown as spheres (sienna for Fe, green for Ni, blue for N, red for O). Distances are shown in Ångstrom.

[NiFe] hydrogenases containing a conserved cysteine motif [44]. The proximal [4Fe-4S] cluster found in the similar O_2-sensitive group 1 [NiFe] hydrogenases is exchanged in the O_2-tolerant group 1 [NiFe] hydrogenases against a unique [4Fe-3S] cluster, which is coordinated by six cysteines and is able to adopt three redox states [45–47]. The [4Fe-3S] cluster undergoes redox-dependent rearrangements and has an additional OH^- ligand in the oxidized state [48]. However, in addition to the electron transfer chain other parts of the structure likely contribute to O_2 tolerance, such as access to the active site [49, 50].

2.2. Active Site Structure of [NiFe] Hydrogenases

The surprising result of the first high-resolution crystal structure of a [NiFe] hydrogenase was the presence of a bimetallic site [34], followed by a second surprise, the presence of three diatomic ligands, one CO and two CN^- ligands bound to iron [51, 52] (Figure 2A).

Ni is coordinated by four cysteine residues with a geometry of a strongly distorted square-plane. Variants in Ni coordination exist in the different enzymes, such as in [NiFeSe] hydrogenases, where a selenocysteine ligand at Ni replaces one of the terminal cysteine ligands found in [NiFe] hydrogenases and an additional sulfur ligand was detected [41, 53].

Two of the Ni-coordinating cysteine residues act as bridging ligands to Fe, which is furthermore coordinated by two CN^- and one CO ligand [51, 52]. The

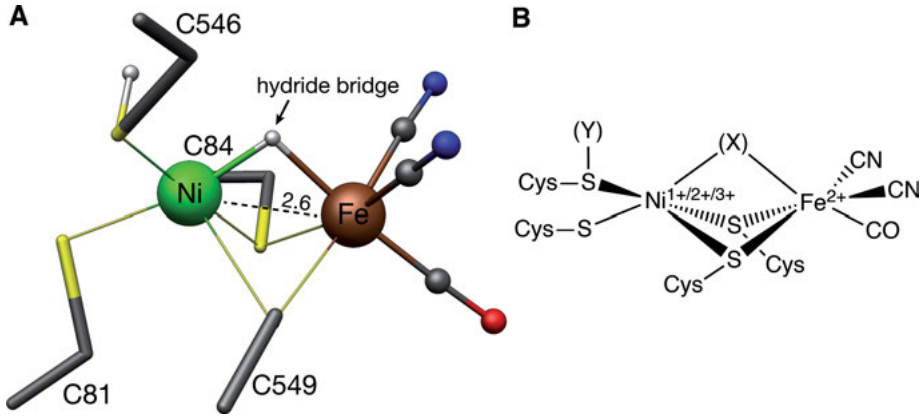


Figure 2. Active site of [NiFe] hydrogenase from *D. vulgaris* Miyazaki F (PDB: 4U9H) [59]. (**A**) Ball-and-stick model of the fully reduced Ni-R state. Color code is green for Ni, dark brown for iron, red for O, blue for N, gray for carbon, and silver for H. Coordinating cysteines are labelled. Distances are shown in Ångstrom. (**B**) Schematic drawing of the cysteine coordinated [NiFe] active site. Ni was observed in three distinct oxidation states, whereas the oxidation state of Fe remained unchanged. Mono- or diatomic ligands were observed in different states between Ni and Fe atoms (depicted as (X)), as well as at one of the Ni-ligating cysteine sulfur (depicted as (Y)).

sixth coordination site of Fe is in most oxidation states occupied by another bridging ligand, which is either an oxygen-containing ligand in the inactive oxidized states, or a hydride anion in the reduced states, completing a distorted octahedral ligand arrangement (Figure 2B). The configuration of the Fe ligands has been deduced from a hydrogen bond pattern and favors the CO in *trans* position to the bridging coordination site. Structural identification of these states is not easy, due to redox heterogeneity [34, 53].

At least two paramagnetic inactive, oxidized states exist, which are termed Ni-A and Ni-B (Figure 3). Ni-A exhibits a long lag phase during reductive activation, while Ni-B is quickly reactivated. Crystal structures of the two states found evidence for a diatomic ligand, potentially a μ-hydroperoxo or a μ-oxo bridge in Ni-A, and a μ-hydroxo-bridge in Ni-B [54–56].

Oxygen-containing ligands are not present in the reduced states, where the distance between Ni and Fe shrinks compared to the oxidized states [41, 57]. The catalytic, reduced states are termed Ni-R and Ni-C and contain a Ni,Fe-bridging hydride ligand, demonstrated by HYSCORE and ENDOR spectroscopy and subatomic-resolution crystallography [58, 59] (Figure 2A). The 0.89 Å resolution structure (PDB ID: 4U9H) allowed to resolve many protons and indicated an additional proton on the Ni-coordinating Cys546 in the [NiFe] hydrogenase from *D. vulgaris* [59].

Other active site variants exist such as an oxidative inactivation by thiol oxidation. These S oxygenations are likely reversible [60]. CO is a reversible inhibitor of H_2 oxidation in [NiFe] hydrogenases by binding as a distinctly non-linear end-on ligand to Ni^{2+} [61]. A further reversible variation of the Ni coordination is found in the complex I-related soluble hydrogenase, a representative of an O_2-tolerant group 3 [NiFe] hydrogenase. Its structure revealed a six-coordinated Ni in the oxidized state with a glutamate side chain as additional ligand [38]. The glutamate carboxylate is eliminated in the reduced state giving the same coordination found in group 1 hydrogenases. This additional redox switch may contribute to the O_2 tolerance of the enzyme [38].

2.3. Reversible Dihydrogen Oxidation at the [NiFe] Site

The mechanism of reversible proton reduction has been studied by a variety of different approaches including computational, structural, and spectroscopic methods. Specifically vibrational spectroscopy yielded more information than with most other metalloenzymes, using the CO and CN^- ligands as sensors for the active site, which could be applied to enzymes in solution, immobilized on electrodes or in their cellular context [62–65]. Furthermore electrochemistry, especially protein-film voltammetry has been used extensively for these highly active enzymes [66–69].

Independent of the state of the enzyme, Fe stays as low-spin Fe(II) ion, whereas Ni may adopt the oxidation states Ni(I), Ni(II), and Ni(III) (Figure 3). Thus, Ni is the redox active element of the active site. The two paramagnetic oxidized states, Ni-A and Ni-B and the reduced Ni-C state contain Ni(III), whereas the Ni-L state contains Ni(I). The latter may be part of the catalytic cycle [70],

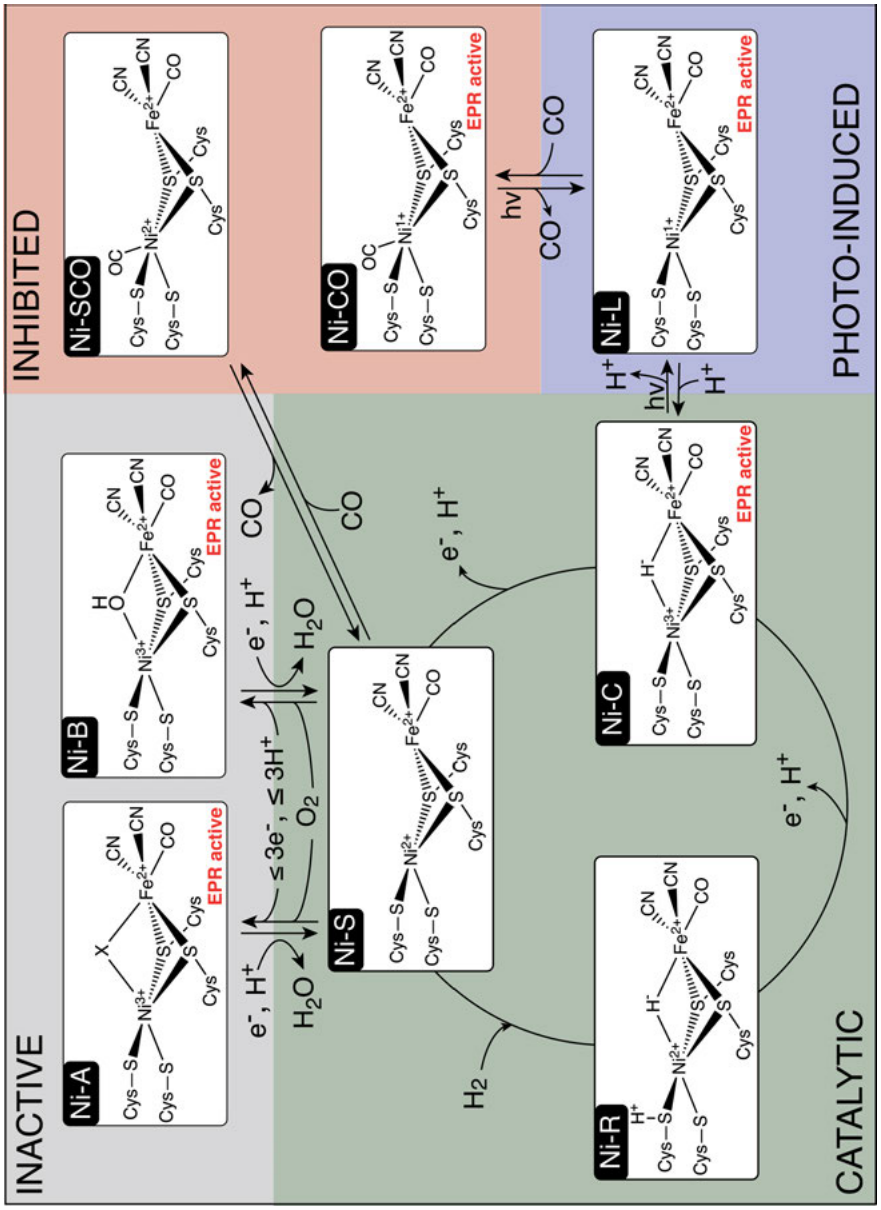


Figure 3. Overview of the simplified catalytic cycle, inhibition, oxidation, and photo-induction mechanisms for [NiFe] hydrogenases.

which may be minimally described by three states, EPR-silent Ni_a-S (also called Ni-SIa), Ni-R, and Ni-C [19]. Ni_a-S contains Ni(II) and is able to bind and activate dihydrogen forming the catalytic Michaelis complex. Despite its small size, dihydrogen likely reaches the active site in Ni_a-S through channels detected using Xe derivatization of crystals [71] and molecular dynamics calculations [72]. The Ni_a-S state is presumed to have no bridging ligand between Ni and Fe and is susceptible to inhibition by CO, forming the Ni-SCO state.

How dihydrogen is activated is one of the fundamental, but still hardly understood steps of the mechanism and it is likely that the ligand arrangement around Ni in the Ni_a-S state plays a major role [73, 74]. The following heterolysis of dihydrogen generates Ni-R, in which the H_2-derived hydride bridges Ni(II) and Fe(II), while the proton is either taken up by a Ni-coordinating cysteine thiolate or an arginine above the hydride ligand (Figure 3) [69]. An exchange of the arginine against a lysine lowers hydrogenase activity by two orders of magnitude [69] and is in line with the central role of pendant amines as proton relays in Ni-containing hydrogenase model complexes [75]. One electron oxidation of Ni-R by electron transfer to the electron transfer chain results in a deprotonation of the Ni-coordinating cysteine thiolate and generates Ni-C in which Ni(III) binds the bridging hydride [58]. If Ni-C is illuminated at low temperatures, Ni-L develops [76]. The following proton-coupled electron transfer regenerates Ni_a-S, ready for another round of H_2 oxidation. The reverse order of steps would allow to reduce protons to dihydrogen.

The role of sulfur in the mechanism is partly illuminated by comparing the [NiFe] with the homologous [NiFeSe] hydrogenases [77]. While there are some structural differences remote from the active site [78], the higher activity of [NiFeSe] hydrogenases is likely a direct consequence of having a selenolate instead of a thiolate ligand [77]. Se forms weaker bonds than S, making it more acidic, but also less prone to form Se–O bonds after reacting with dioxygen [79]. Furthermore, it affects the redox equilibria of the [NiFe] site [80].

3. THE [Ni4Fe-4S] SITE OF CARBON MONOXIDE DEHYDROGENASES

3.1. Carbon Monoxide Dehydrogenases

Carbon monoxide dehydrogenases (CODHs) catalyze CO oxidation according to Equation (1) and can formally also be called CO:acceptor oxidoreductases [1, 3]:

$$CO + H_2O \rightleftharpoons CO_2 + 2H^+ + 2e^- \quad (1)$$

With a midpoint redox potential of the CO/CO_2 couple below –500 mV [81, 82], CO oxidation supplies highly energetic electrons, while CO_2 reduction requires a strong reductant. CO oxidation with water also occurs in the water-gas-shift reaction to change the ratio of $CO:H_2$ in synthesis gas [83]. However, while the water-gas-shift reaction produces dihydrogen, CO oxidation by CODHs generates electrons and protons separately.

Two principal types of CODHs have been found in bacteria and archaea, containing different cofactors and different metals in the active site. Furthermore, they are distinguished by their capacity to reduce CO_2 and their stability in the presence of air [1]. Cu,Mo-containing CODHs have been isolated from aerobic carboxydotrophic bacteria [84], contain Cu and Mo in their active site [85] and are stable in the presence of air under non-turnover conditions. But Cu,Mo-CODHs have the disadvantage that they cannot reduce CO_2. In contrast, the Ni,Fe-containing CODHs are found only in anaerobically living bacteria and archaea and are oxygen-sensitive. Ni,Fe-CODHs contain Ni and Fe in their active site and efficiently reduce CO_2.

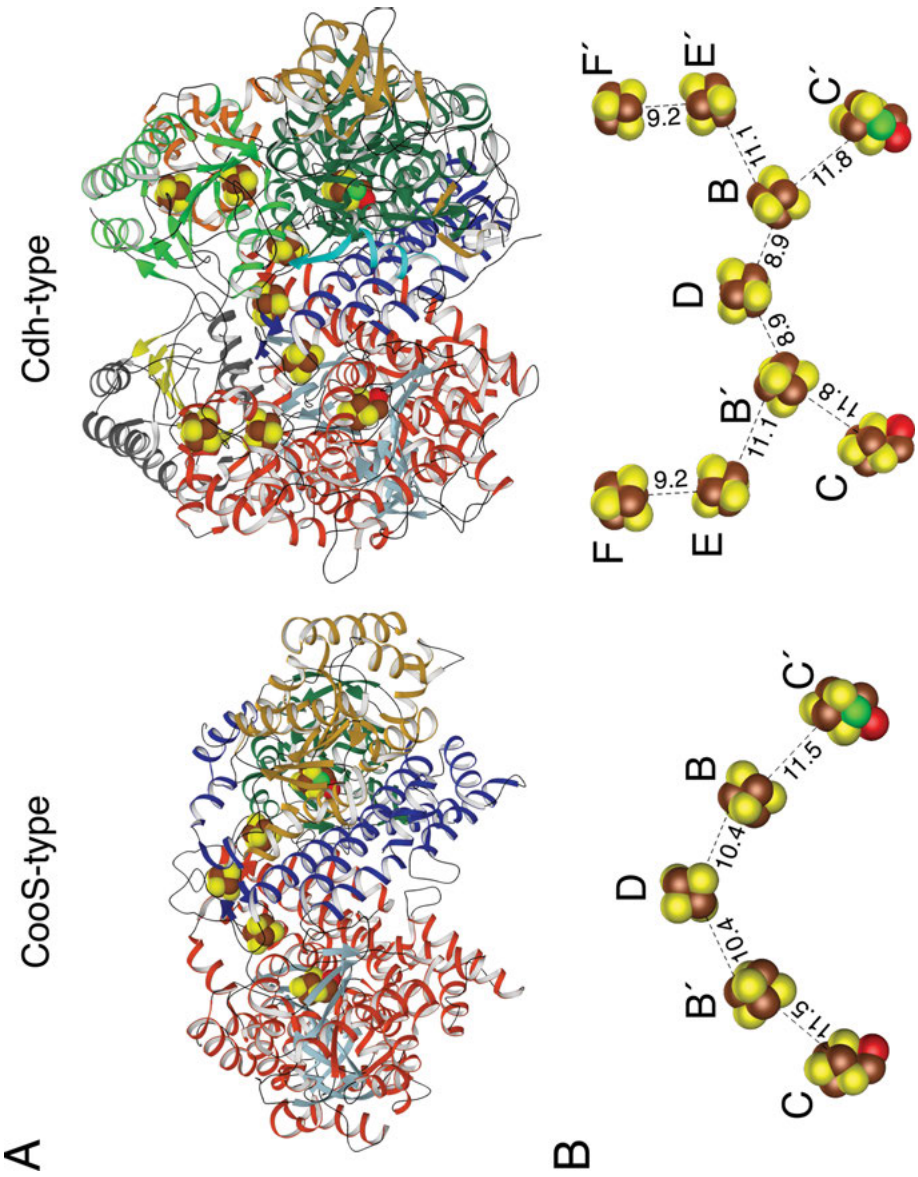
Based on subunit composition and amino acid sequence, Ni,Fe-containing CODHs can be grouped into four classes [86, 87]. Classes I and II are found in archaea with class I CODHs being typical for but not restricted to autotrophic methanogens where they generate acetyl-CoA, while class II CODHs are found in facultative chemo-autotrophic methanogens, typically catabolizing acetate. CODHs of class I and II are components of a large protein complex, termed acetyl-CoA decarbonylase/synthase (ACDS) complex, comprising six [88] or eight [89] copies of a heteropentamer, which is encoded by the *cdh* operon. CODHs of class I and II are larger when compared to the CODHs of other classes and harbor additional Fe-S clusters. Based on the operon name enzymes of class I and II are denoted as Cdh-type CODHs.

In contrast to these archaeal Cdh-type CODHs, classes III and IV are the bacterial-type CODHs, called CooS-type. Class III CODHs are found in acetogens and are associated with acetyl-CoA synthases (ACS, see Section 4.1), together catalyzing the transformation of CO_2, two electrons, a methyl cation and CoA to acetyl-CoA. The simplest type of CODH are the monofunctional CODHs of class IV, which catalyze CO oxidation allowing several bacteria to use CO as electron source [90, 91].

Several anaerobic microorganisms encode more than one CODH in their genome [92], which can be differently regulated according to metabolic needs [93–96]. Phylogenetic analysis of CODH sequences as well as the early chemistry of life indicate that a common ancestor of Ni,Fe-CODHs was already present in the last-universal ancestor more than 2.5 billion years ago [13, 14, 97, 98].

3.2. [Ni4Fe-4S] Cluster: Structure and Ligand-Binding Site

The overall structure of CODHs has been analyzed at different levels of resolution, starting with the 26-Å resolution structure obtained by scanning transmission electron microscopy of the negatively stained Cdh-type CODH from *Methanosarcina thermophila* [99] to the 1.0-Å resolution X-ray structures of CooS-type CODH-II from *Carboxydothermus hydrogenoformans* [100]. We have now highly-resolved structures of different CODHs from classes II [101], III [102, 103], and IV [104–107], allowing us to appreciate parallels and differences in their architectures (Figure 4).



We will focus on the direct environment of the Ni,Fe cluster in the active site. A common feature of enzymes with an Fe-S cluster in the active site, is its location at the merging point of Rossman-fold domains [108]. This holds also true for CODHs, in which the NiFe cluster is coordinated and embraced by two Rossman-fold domains, each domain contributing three amino acid residues to coordinate the cluster [104, 105]. The surrounding protein domains are not only holding the cluster, but are also forming the channels needed to transport substrates towards and products away from the active site cluster. CO may enter the active site along two channels [103, 109] while water may be guided along a pathway of conserved protonable residues, mostly histidines [110]. The same network may also transfer the two protons. CO_2 is likely to leave either via one of the static gas channels or by a dynamically formed gas channel, uncovered by molecular dynamics calculations [111]. Electrons take the obvious path along the chain of [4Fe-4S] clusters [104, 105].

Channeling of substrates and products is vital as turnover of CO is very rapid in CODHs reaching a turnover number of 39,000 s^{-1} at optimal temperatures [1, 90]. CODH-IV from *C. hydrogenoformans* reaches a k_{cat}/K_m of 3×10^8 to 10^9 M^{-1} s^{-1} and is therefore only limited by the diffusion of CO to the active site [106]. CO_2 reduction occurs about two to three orders of magnitude slower than CO oxidation and is highly dependent on the reducing agent and mediators [112, 113]. In addition to CO_2, CODHs are able to reduce carbonyl sulfide (COS) [112], cyanate (NCO^-) [100, 114, 115], nitrous oxide (N_2O) [116, 117], and hydroxylamine (NH_2OH) [118], and oxidize isocyanides (R-NC) [109, 119], likely relying on the same mechanistic strategies used to convert CO and CO_2.

The active site is formed by a [Ni4Fe-4S] cluster called cluster C. Cluster C consists of a distorted [Ni3Fe-4S] heterocubane connected to an Fe(II) ion in *exo* (Fe_1) (Figure 5). Ni and Fe are each coordinated by one cysteine residue, with the Fe_1 ion in *exo* having an additional histidine ligand. The [3Fe-4S] subsite has a similar Fe-S arrangement as found in [4Fe-4S] clusters. The Ni(II) and the

Figure 4. Overall structure of CODHs. (**A**) Cartoon representation of homodimeric Ni,Fe-CODHs: CooS-type (PDB: 3B53 [104]) and Cdh-type (PDB: 3CF4 [101]). The two core subunits of both CODHs (α-subunit of Cdh-type) are shown with different color scheme. One subunit is drawn in red for α-helices and light blue for β-strands and the other is highlighted in marine-blue, sea green and brown for the N-terminal, middle, and C-terminal domain, respectively. Additional domains found in the α-subunit of Cdh-type CODH are colored in cyan for the N-terminal portion and orange for the [4Fe-4S] cluster-binding domain. One of the ε-subunits from Cdh-type CODH is drawn with light green and the other is highlighted in gray for α-helices and yellow for β-strands. The metal clusters are depicted as spheres (green for Ni, brown for Fe, yellow for S, and red for O). (**B**) Metal cluster arrangements in CODHs. A [4Fe-4S] cluster D connects covalently the two core subunits in both CODHs and is in electron transfer distance to clusters B/B′ ([4Fe-4S]). The active site [Ni4Fe-4S-OH] clusters C/C′ are situated at the end of the electron transfer chain. In Cdh-type CODH, four additional [4Fe-4S] clusters (E/E′ and F/F) are present with clusters E/E′ bridging possible electron transfer between clusters B/B and clusters F/F. The closest distances between Fe atoms of individual clusters are shown in Ångstrom.

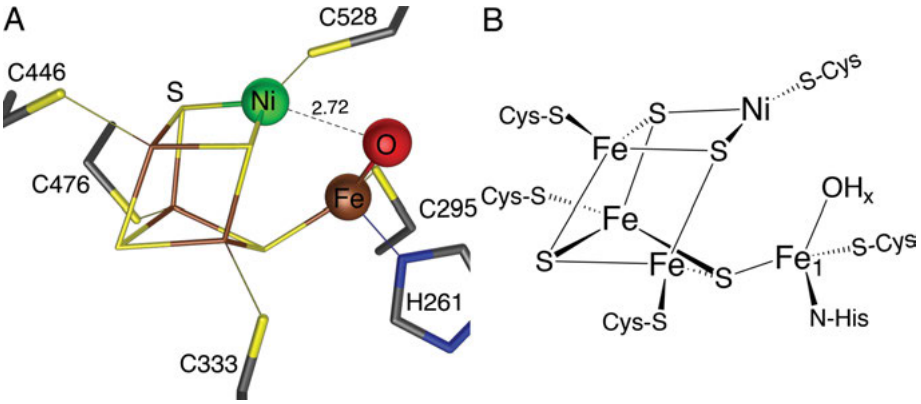


Figure 5. Cluster C in CODHs. (**A**) Ball-and-stick model of active site cluster C from CooS-type in the –320 mV state (PDB: 3B53 [120]). Color code is green for nickel, brown for iron, red for oxygen, yellow for sulfur, blue for nitrogen, and gray for carbon. Coordinating cysteines and histidine are labeled. Alternative positions of Fe_1 and Cys295 are omitted for clarity. Metal to S/O distances are shown in Ångstrom. Same color code as in Figure 4 and gray is used for C. (**B**) Schematic drawing of the [Ni4Fe-4S] cluster C.

Fe_1(II) ions have an open coordination site, which is either filled by a OH^- ligand [101, 120, 121] or a bridging CO_2 ligand [100, 120], both relevant for the mechanism. A μ-sulfido ligand, originally claimed to be an essential component of the active site cluster [104, 122], is not necessary for catalysis [120] and appears to indicate an inactive, oxidized state [114, 123].

Ni in cluster C is, under all conditions investigated so far, a low-spin Ni(II) ion [1]. The available crystal structures are consistent with this observation, as they reveal a planar T-shaped coordination in the OH^--bound state and a square-planar coordination when CO_2 or a sulfido ligand bridge Ni and Fe – both planar arrangements should be low-spin for an atom with $3d^8$ electron configuration [124]. Shortest Ni–Fe(II) distances are 2.8–2.9 Å in the active states and 2.6 Å in the CN^--inhibited state [125, 126].

The second coordination sphere around the [Ni4Fe-4S] site of cluster C is highly conserved and hosts a histidine and a lysine residue (His93 and Lys563 in CODH-II numbering) [104, 105]. These two residues are positioned to form hydrogen bonds with ligands bound at Ni and Fe_1 and likely participate in stabilizing and (de)protonating substrates and reaction intermediates during the catalytic cycle.

3.3. Reversible Carbon Dioxide Activation by Nickel and Iron

In the catalytic cycle (Figure 6), cluster C adopts at least three oxidation states, called C_{red1}, C_{red2}, and C_{int}. These states have been studied by spectroscopy, electrochemistry, and to some degree by protein crystallography. In addition to

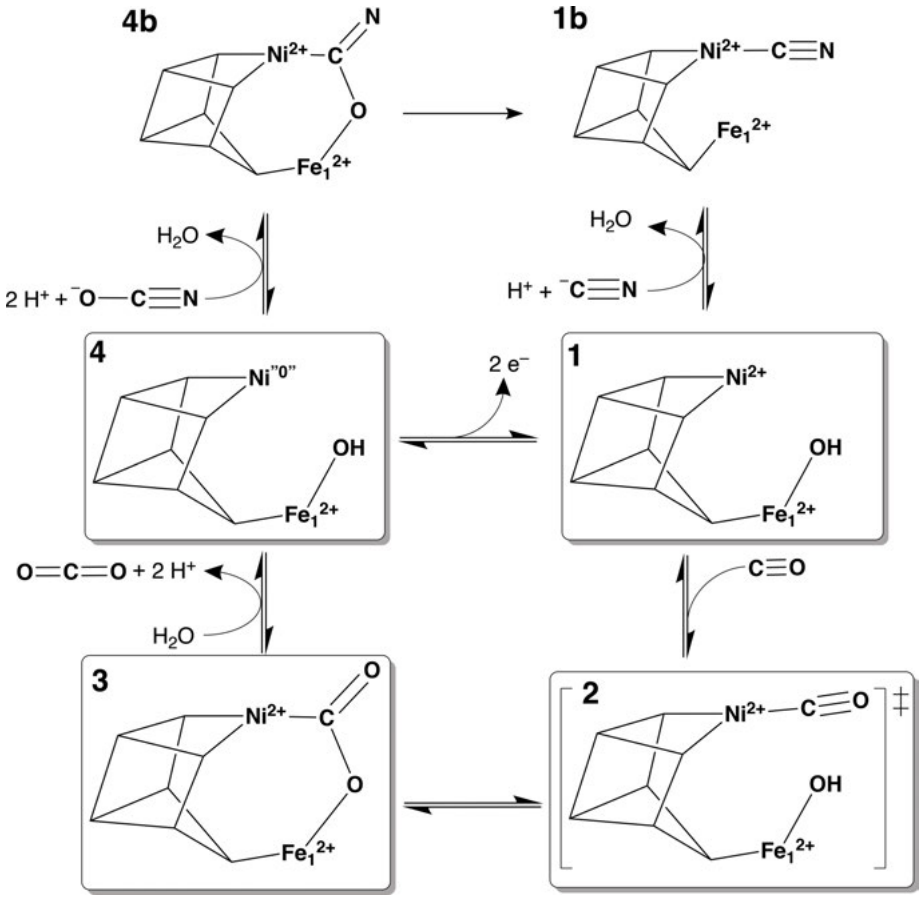


Figure 6. Mechanism of CO oxidation at cluster C. (**1**) The active C_{red1} state with a hydroxyl group bound to Fe_1 and stabilized by K583. (**2**) A transition state of CO bound to Ni^{2+} completing a square planar coordination. CO is in close proximity for a nucleophilic attack by the hydroxyl group. (**3**) CO_2 formation, where the oxidation product CO_2 is bridged between Ni and Fe_1. (**4**) The C_{red2} state, where CO_2/water are released/uptake, respectively. Cluster C contains two additional electrons compared to the C_{red1} state. The successive release of two electrons restores cluster C to the catalytically competent C_{red1} state. Cyanate-bound (4b), which is turned over to the (1b) cyanide-inhibited state. See text for details.

the catalytically relevant oxidation states, at least one more oxidized state, called C_{ox}, has been described. However, preparations of CODH proteins are often heterogenous with regard to the metal content and oxidation state, and the Ni,Fe center selectively bind ligands in the different oxidation states modulating their electronic structure. Consequently, their mechanism may contain more states and CODHs may be more diverse in their reactivity and states than earlier anticipated. Thus, the below listed states give a minimal description of what to expect from a Ni,Fe cluster reversibly oxidizing CO.

At redox potential above –200 mV the catalytic C_{red1} state is converted to the diamagnetic and catalytically inactive C_{ox} state. C_{ox} has the proposed oxidation state {[Ni^{2+} Fe^{3+}]:[3Fe-4S]$^{-}$}. Below –200 mV C_{ox} is reduced by a single electron to the C_{red1} state [87, 127].

C_{red1} is paramagnetic with S = 1/2 (g_{av} = 1.82, with g_1 = 2.01, g_2 = 1.81, g_3 = 1.65 for the bifunctional CODH from *Moorella thermoacetica* [128, 129], and g_{av} = 1.87, with g_1 = 2.01, g_2 = 1.88, g_3 = 1.71 for the monofunctional CODH from *Rhodospirillum rubrum* and CODH-I from *C. hydrogenoformans* [114, 130]) and its EPR and Mössbauer parameter could be reconciled with {[Ni^{2+} Fe^{2+}]:[3Fe-4S]$^{-}$} [87]. Crystal structures likely reflecting the C_{red1} state show an OH^- ligand bound to the Fe_1^{2+} ion and a T-shaped planar coordination for Ni^{2+} [101, 120, 121]. The OH^- bound to Fe_1^{2+} has a Ni–O distance of 2.7 Å (Figure 5A) suggesting a weak interaction.

C_{red1} is the state in which CO is supposed to bind to cluster C. Although reliable crystallographic data for the CO-bound state are absent, the open coordination site at Ni^{2+} and studies of a CN^--bound state, clearly indicate that CO binds to the Ni^{2+} ion (Figure 6). Cyanide (CN^-) is isoelectronic to CO and a slow-binding inhibitor competitive with CO [131]. Inhibition by CN^- has been extensively studied in CODHs. As a competitive inhibitor its inhibition constant reflects its dissociation constant and low μmolar K_is have been reported for the CODHs isolated from *R. rubrum* and *C. hydrogenoformans* [131, 132]. EPR spectroscopy shows that CN^- binds to the C_{red1} state [133, 134] and IR spectroscopy revealed a CN^- stretching band at 2110 cm^{-1} [126]. The structure of CODH-II crystals soaked with CN^- shows that CN^- binds to an open apical coordination site with only a slight tilt of the Ni-C-N axis [125, 126].

While many Ni carbonyl complexes are known starting from elemental Ni, Ni^{2+} seems an odd choice to bind CO, because as a "hard" metal ion it does not readily donate electrons needed to form a Ni-CO complex stabilized by back-bonding [83]. However, the electronic structure of Ni^{2+} is modulated by the surrounding ligands, in particular by strong donors, e.g., the thiolate CyS^- and the two μ_3-sulfido ligands of Ni^{2+} in cluster C, which may help stabilizing CO at Ni^{2+} (Figure 6) [135]. As electron back-donation is still weak, the CO molecule is most likely highly electrophilic and should readily react with the nucleophilic OH^- ligand at Fe_1, forming a metal-bound carboxylate (Figure 6) [83].

A bound carboxylate ligand could be stabilized and analyzed by X-ray crystallography [100, 120]. The CO_2-moiety bridges Ni and Fe_1 thereby completing the square-planar coordination of Ni (η^1-***C***O_2 mode) and one of the oxygens of CO_2 replaces the OH^- ligand at Fe_1 (η^1-***O***CO mode). The C–O bond lengths (1.30 and 1.32 Å) and the O-C-O bending angle of 117° indicate that CO_2 is reduced by two electrons and is best described as a metal-bound carboxylate [100]. The short Ni–C bond of 1.8 Å reveals substantial π-back-bonding and resembles a Ni carbene. Hydrogen bonds to a lysine and a histidine residue stabilize the bound CO_2. The Ni/Fe/sulfide frame of cluster C is unchanged by CO_2 binding.

In the next step, the bound CO_2 ligand needs to be released requiring the transfer of two electrons on cluster C forming the C_{red2} state. The midpoint potential of the C_{red1}/C_{red2} redox couple is about –530 mV [129, 136], coinciding

with the potential of the CO_2/CO couple of –558 mV [81]. Like the C_{red1} state, C_{red2} is paramagnetic with S = 1/2 and exhibits a similar EPR spectrum as C_{red1}, with $g_{av} = 1.86$ ($g_1 = 1.97$, $g_2 = 1.87$, $g_3 = 1.75$) [129, 130]. Where the two additional electrons distinguishing C_{red2} from C_{red1} reside is unclear, and a Ni-hydride as well as a Ni–Fe bond were postulated [137, 138]. All observable redox chemistry in cluster C seems to occur at the Fe ions, and spectroscopic evidence for a change in the oxidation state of Ni has not been documented yet.

The C_{red2} state returns to C_{red1} by donating two electrons in single-electron transfer steps to a [4Fe-4S] cluster termed cluster B from where the electron may be transferred to external redox acceptors or another [4Fe-4S] cluster sitting in the dimer interface, termed cluster D. Oxidation of C_{red2} by one electron yields an integer spin state called C_{int} [139].

4. THE [2Ni4Fe-4S] SITE OF ACETYL-COENZYME A SYNTHASES

4.1. Acetyl-Coenzyme A Synthases

Acetyl-CoA synthases (ACS) catalyze the reversible condensation of CO, a Co(III)-derived methyl cation and coenzyme A to acetyl-CoA, the universal form of activated acetic acid (Equation 2). ACS should not to be confused with acetyl-CoA **synthetases**, which are ligases forming acetyl-CoA from acetic acid and CoA at the expense of ATP, hence the name synthetase versus synthase, which requires no ATP input.

$$CH_3\text{-Co(III)CoFeSP} + CO + \text{CoASH} \rightleftharpoons CH_3C(O)SCoA + \text{Co(I)CoFeSP} + H^+ \quad (2)$$

ACS occupies a central place in two microbial pathways: the reductive acetyl-CoA pathway also called Wood-Ljungdahl pathway, and methanogenesis from acetate and from CO_2/H_2.

Phylogenetically diverse, anaerobic microorganisms, including sulfate-reducing bacteria and archaea, hydrogenogens, acetogens, and methanogens use the reductive acetyl-CoA pathway to generate biomass and energy from CO or CO_2 and H_2 [140, 141]. CO_2 enters the pathway along two branches: in one branch it is reduced by six electrons to a folate- or pterin-bound methyl cation, and in the other branch it is reduced by two electrons to CO. ACS catalyzes the reaction in the merging point of both branches by fusing CO and the methyl cation to an acetyl group. In methanogenesis from acetate by aceticlastic archaea, such as *Methanosarcina* species and the non-methanogenic *Archaeoglobus fulgidus*, ACS catalyzes the reverse reaction, the splitting of acetyl-CoA [89, 142, 143].

Mono-, bi-, and multifunctional ACS complexes are found in Nature with the bifunctional CODH/ACS being the dominant form in bacteria and the multifunctional ACDS complex typically found in archaea (see also CODH classes in Section 3.1) [86]. Bifunctional CODH/ACS complexes are typically involved

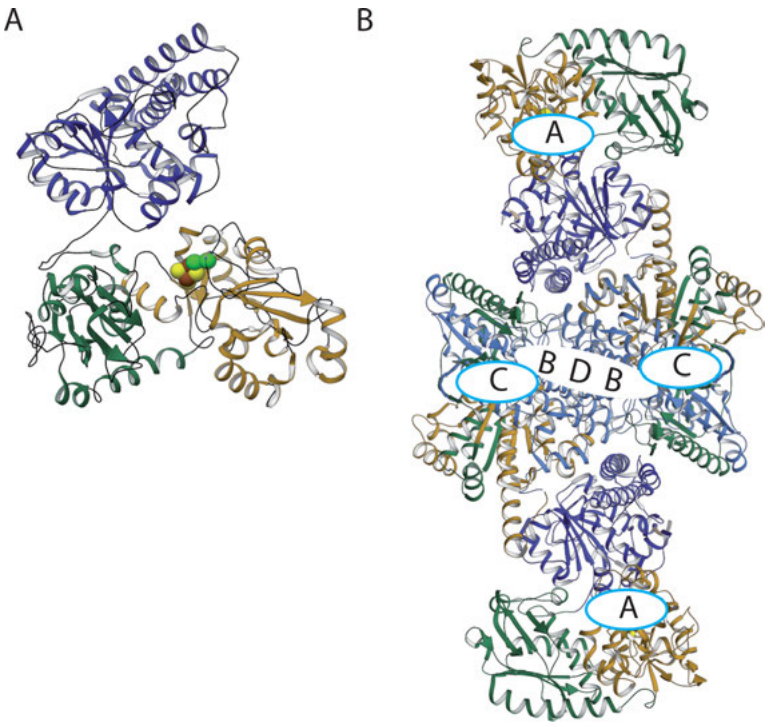


Figure 7. Overall structure of ACS. (**A**) Cartoon-representation of monomeric ACS (PDB: 1RU3 [145]). The three domains are colored in blue, sea green, and brown for the N-terminal, middle, and C-terminal domain, respectively. The active site cluster A (a binuclear nickel site bridged to a [4Fe-4S] cluster) is depicted as sphere (green for nickel, brown for Fe, and yellow for S). (**B**) Cartoon representation of bifunctional CODH-ACS (PDBs: 1MJG [102] or 1OAO [103]). The three domains of the ACS component (top and bottom) are colored as for monomeric ACS. The dimeric CODH central component is also colored marine blue, sea green, and brown for its N-terminal, middle, and C-terminal domain, respectively. Clusters A in ACS and clusters C in CODH are indicated by white boxes labeled "A" and "C". The three bridging [4Fe-4S] clusters of CODH are symbolized by a white box labeled "B D B". The distance between cluster A of ACS and all three clusters of CODH is 60–65 Ångstrom (not shown).

in acetogenesis from CO_2 and H_2, with a class-III CODH [86] reducing CO_2 to CO, which is channeled to the active site of ACS, where CO is used to form the carbonyl group of acetyl-CoA. ACDS complexes containing class I and class II CODHs [86] may be called multifunctional as they not only incorporate the two activities of the bifunctional CODH/ACS complex, but additionally convert a methylated pterin, either methanopterin or sarcinapterin, to give an intermediary methylated Co(III) complex that donates the methyl group to the active site of ACS [144].

A monofunctional ACS has so far only been described for *C. hydrogenoformans* acting as a monofunctional enzyme under physiological conditions [145].

When *C. hydrogenoformans* grows in the presence of elevated CO concentrations, it stops synthesizing the CODH component (CODH-III) from the bifunctional CODH/ACS but keeps producing the ACS component as a monofunctional enzyme, which uses the exogenously supplied CO as a substrate [145]. When *C. hydrogenoformans* grows at low CO concentrations, the same ACS protein is found in complex with CODH-III [145].

Crystal structures of two different ACS enzymes have been resolved to date: the bifunctional CODH/ACS complex from *M. thermoacetica* ($CODH/ACS_{Mt}$) [102, 103], of which also a shortened protein version has been published [146], and the monofunctional ACS from *C. hydrogenoformans* (ACS_{Ch}) [145] (Figure 7). The bifunctional CODH/ACS is an $\alpha_2\beta_2$ heterotetramer of about 310 kDa and dimensions of approx. $200 \times 75 \times 105$ Å, through which a 138 Å long gas channel connects cluster C and cluster A [102, 103, 147]. The active sites, marked by the positions of cluster C (CODH) and cluster A (ACS), are 67 Å apart without bridging Fe-S clusters. The CODH component of the bifunctional enzyme is identical to that of monofunctional CODHs, so that we will focus on the structure of the ACS subunit alone.

Each ACS subunit is composed of three domains: an N-terminal domain of about 315 amino acids with a Rossmann-fold and sharing similarity with CODH [102, 146], a middle domain from residues 316–500 and a C-terminal domain consisting of residues 501–732, both with $\alpha+\beta$ fold. The C-terminal domain is the binding site for the Ni,Fe-cluster (cluster A) [102, 103, 145]. The presence of the N-terminal domain is characteristic for bacterial-type ACS proteins, while the archaeal ACS subunit of the ACDS complex lacks the N-terminal domain and carries a 75 amino acid C-terminal extension, which may serve as a docking module for the other proteins in the ACDS complex [148, 149].

4.2. Structure of the [4Fe-4S]-(μ_2-SCys)-[Ni((μ_2-SCys)$_2$Gly)Ni] Site

It had long been anticipated that cluster A consists of a [4Fe-4S] cluster to which a Ni ion is linked. It came as a surprise that not one, but two different metal binding sites are next to the [4Fe-4S] cluster [102] (Figure 8). The two metal binding sites are named after their relative location to the [4Fe-4S] cluster as proximal and distal site. The distal site harbors a Ni^{2+} ion, which is held by a Cys-Gly-Cys motif, coordinating the Ni ion by two backbone amide nitrogens and the thiolates of the two Cys residues with a square-planar arrangement (Figure 8). The Cys-Gly-Cys motif is a versatile and almost autonomous metal binding site, productively used for synthetic complexes mimicking ACS chemistry [150, 151]. Remarkably, the two thiolates coordinating the distal Ni^{2+} ion (Ni-d; in Figure 8 Ni_d) are also coordinating the proximal Ni ion (Ni-p; in Figure 8 Ni_p). A third ligand of the proximal Ni ion is a cysteine thiolate, which the Ni ion shares with the [4Fe-4S] cluster next to it. Cluster A may therefore be described as a [4Fe-4S]-(μ_2-SCys)-[Ni((μ_2-SCys)$_2$Gly)Ni]-site [152].

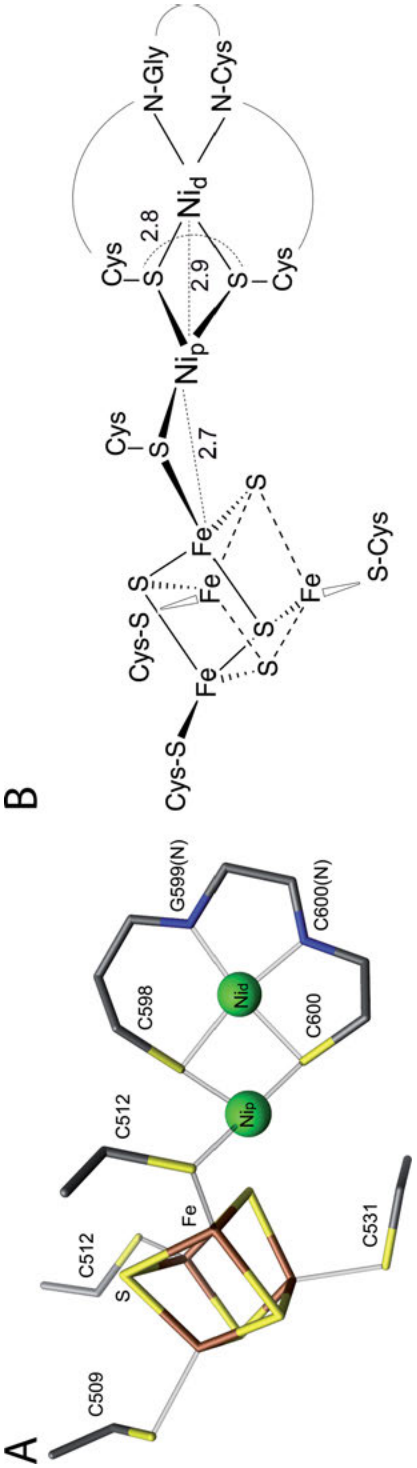


Figure 8. Active site structure of cluster A in ACS. (**A**) Ball-and-stick model of active site cluster A of monomeric ACS from *C. hydrogenoformans* (PDB: 1RU3 [145]). Color code is green for Ni, brown for Fe, yellow for S, gray for C, and blue for N. (**B**) Schematic drawing of the [4Fe-4S]-(μ_2-SCys)-[Ni((μ_2-SCys)$_2$Gly)Ni] site. Distances are given in Ångstrom.

The unusual T-shaped coordination by three shared thiolates is not only providing the proximal Ni with unusual properties, but is most likely also responsible for the promiscuous metal binding and associated heterogeneity that bothered ACS researchers early on [2]. The problem may be inherent to sulfur ligation in general [153] and N_2S_2Ni complexes in particular, which show a higher affinity for Cu than for Ni [150], a tendency also observed in the first CODH/ACS structure [102].

A gas tunnel connecting cluster C and cluster A ends directly above the proximal Ni ion, one of the reasons making it the most likely site for substrate coordination. In contrast to the other two Ni,Fe,S-containing active sites, cluster A is not buried within the protein matrix, but is rather solvent-exposed, likely because transmethylation directly from the corrinoid bound to the corrinoid-Fe-S protein (CoFeSP) requires sufficient space.

4.3. Ligand Binding and Acetyl-Coenzyme A Formation

The mechanism of acetyl-CoA formation by ACS attracts considerable attention from chemists, because Ni, Fe, and S manage a chemistry, which in industrial settings requires Rh (Monsanto process) or Ir (Cativa process) [83]. Evolutionary biologists on the other hand look at reaction mechanisms of CODH and ACS because of parallels to an autotrophic origin of life: both enzymes act in the earliest autotrophic metabolic pathway and they were most likely present in LUCA, they catalyze reactions such as the condensation of acetic acid that would be expected at the origin of life and use active site metal clusters resembling minerals at which these reactions could have started [14, 154].

Several different mechanisms have been proposed for ACS, which may be distinguished according to three choices: (a) monometallic versus bimetallic; (b) diamagnetic versus paramagnetic; and (c) methyl-first versus CO-first. All proposals deal directly with the function of the two Ni ions in substrate binding and turnover.

(a) Mono- versus Bimetallic Mechanism: From the beginning on, mono- and bimetallic mechanisms were both postulated and a bimetallic mechanism seemed indicated when the first crystal structure revealed two metal ions next to the [4Fe-4S] cluster [102]. However, apart from the presence of two metals, most arguments point to a monometallic mechanism, in which the binding and condensation of the substrates occur only at the proximal Ni ion. For one, monometallic mechanisms are known from synthetic catalysts catalyzing the condensation of CO and CH_3^+, while there is no precedence for a bimetallic mechanism [155]. Furthermore, density functional theory (DFT) calculations indicate that substrate binding to the distal Ni is clearly disfavored [156, 157]. Thus, although all evidence for substrate binding to cluster A is indirect, all points towards the proximal Ni site as the place of action: CO binding appears to invoke coordination changes at Ni-p [158, 159], methylation prevents extraction of Ni-p by 1,10-phenanthroline, ACS treated by 1,10-phenanthroline to

remove Ni-p cannot be methylated [160], and also model calculations favor methylation of Ni-p [157, 161].

(b) Dia- versus Paramagnetic Mechanism: Assuming binding at Ni-p, Ni-p may cycle through diamagnetic (Ni^0, Ni^{2+}) or paramagnetic (Ni^+, Ni^{3+}) states. Here, CO binding is unproblematic as it may well occur at Ni^0 as well as at Ni^+. However, the transfer of a methyl cation from CoFeSP (methyl cation donor protein) requires a nucleophilic Ni ion at least matching the nucleophilicity of Co(I). Furthermore, binding of the methyl cation is concomitant with a formal two-electron oxidation of the methyl accepting metal site (Figure 9). Thus, a low-valent Ni site such as a zero-valent Ni or Ni^+ is required as methyl group acceptor. As Ni^0 is diamagnetic ($3d^{10}$) and the resulting Ni^{2+} ($3d^8$) could be square-planar and hence also diamagnetic, a mechanism with only diamagnetic intermediates would ensue. In contrast, if Ni^+ ($3d^9$) is the methyl group acceptor, Ni^{3+} ($3d^7$) would be generated, both of which are paramagnetic. Zero-valent Ni would be unprecedented in Nature and would fit with a nucleophilicity superior to that of Co(I) [162]. DFT studies argue against the possibility of Ni^0 next to an oxidized [4Fe-4S] cluster [157, 163], however, that may also be related to the difficulty to correctly model the electron affinity of [4Fe-4S] clusters [162]. Nevertheless, a zero-valent Ni in ACS has not been observed by XAS so far.

The paramagnetic mechanism predicts cycling between Ni^+ and Ni^{3+}. In contrast to Ni^0, Ni^+ has been observed alone [164] and in complex with CO, where it gives rise to the Ni,Fe-C EPR signal, first described by Ragsdale and coworkers [165, 166]. In the paramagnetic mechanism, Ni^+ is the methyl group acceptor, which is oxidized to Ni^{3+}. Methylated ACS with Ni^{3+} has not been observed, however, that may well be due to the reaction conditions, requiring the presence of strong reducing agents, which would swiftly reduce methylated Ni^{3+} to Ni^{2+}, which is the species observed spectroscopically [2].

(c) Methyl- versus CO-first Mechanism: the third question is about the order of substrate binding. Specifically, the question *what binds first*? Does CO bind before the methyl group, or is it the methyl group that has to bind first (Figure 9)? This centers on the question if the CO-bound state associated with the Ni,Fe-C EPR signal is a true reaction intermediate or an inhibited species. Ni,Fe-C forms when ACS is treated with CO and a reducing agent, such as sodium dithionite or Ti(III)-citrate [165, 166]. The Ni,Fe-C state is stable and has been investigated for most ACS preparations. The EPR signal is most likely produced from the state containing $[4Fe\text{-}4S]^{2+}$-$Ni(I)_p$-$Ni(II)_d$, with CO bound to $Ni(I)_p$. The EPR spectra would be consistent with the spin density being delocalized over $Ni(I)_p$, $[4Fe\text{-}4S]^{2+}$, and CO, giving a signal with g-values at 2.08, 2.07, and 2.03, found in bacterial as well as archaeal ACS variants [166, 167]. The EPR signal appears to be sensitive to the conformation of the three domains and the associated solvent exposure of the active site [168]. Infrared spectroscopy indicates that the Ni,Fe-C species consists of a single Ni,Fe-CO species with an absorption at 1995 cm^{-1} and a bandwidth of 8 cm^{-1} [169, 170], which is consistent with a terminal CO ligand. Very similar IR spectra have been produced from synthetic Ni-CO complexes [171–173], as well as from Ni-substituted azurin [174] modeling the active site of ACS, corroborating the electronic structure of the Ni,Fe-C state.

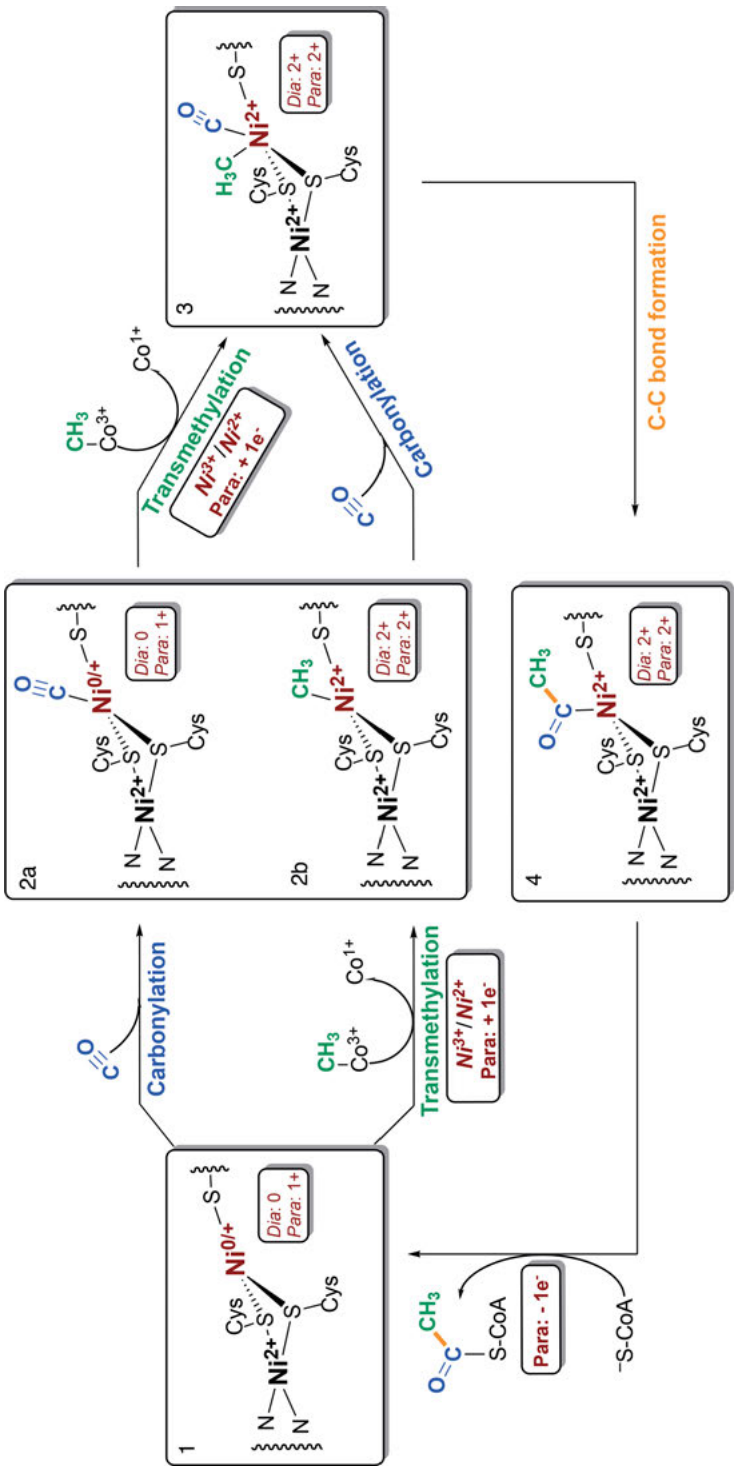


Figure 9. Postulated reaction mechanism of acetyl-CoA synthesis at cluster A. (**1**) The active state with a divalent distal Ni (Ni^{2+}) and a diamagnetic (Ni^{0}) or paramagnetic (Ni^{+}) proximal Ni. (**2**, **3**) Carbonylation and transmethylation occur sequentially. The transmethylation requires an electron if proximal Ni is paramagnetic. (**4**) Condensation of CO and CH_3 to acetyl and its transfer to CoA generates the reaction product acetyl-CoA. An electron is transferred when proximal Ni is paramagnetic. See text for details.

The sequence of substrate binding has been investigated using pulse-chase kinetics, which is consistent with a random binding of either CO or the methyl group in the first step and CoA in the last step [175]. This would argue for the Ni,Fe-C state as a potential reaction intermediate. The prime requirement for a potential reaction intermediate is its kinetic competence, requiring its rate of formation and transformation agreeing with the rate of the overall turnover [176]. Kinetic competence has been repeatedly shown for the Ni,Fe-C state, involving stopped-flow IR spectroscopy and rapid freeze quench EPR sampling [170, 177–179]. However, kinetic competence is a required but not sufficient property for a state to be a true intermediate.

It has been argued against the catalytic relevance of the CO-bound state that CO does not only act as a substrate, but also as an inhibitor of ACS, particularly when bound before the methyl group. This is primarily based on the CO dependency of the rate of acetyl-CoA synthesis, which has no simple hyperbolic increase of rate on CO concentration, but shows a pronounced maximum at low CO concentrations, after which the rate approaches a low non-zero value [180]. This behavior may originate from Ni,Fe-C inhibiting methyl group transfer [181, 182]. This can be either due to CO occupying the site to which the methyl group may be transferred, or by CO reducing the nucleophilicity of Ni, making it less reactive in transmethylation. But inhibition by CO appears more complicated. While inhibition is observed when CODH/ACS_{Mt} is supplied with exogenous CO, no inhibition is found in the presence of CO_2 and reducing agent, generating endogenous CO. Surprisingly, when the ACS_{Mt} subunit is produced in absence of CODH, or when the tunnel connecting CODH and ACS is closed by substituting small for bulky side chains, CO acts as a weak binding substrate, but does not inhibit ACS activity [183], indicating that the inhibitory effect may be due to accumulating CO in the substrate tunnel or by conformational changes. In contrast, the monofunctional ACS_{Ch} shows CO inhibition even though no CODH subunit and tunnel network is associated to it, while deleting the N-terminal domain gives an uninhibited ACS_{Ch} variant [168].

The mechanistic alternatives center around the proximal Ni (Ni-p) leaving the mechanistic role of the other two metal sites, Ni-distal (Ni-d) and the [4Fe-4S] cluster, open. While on a first glance redox activity appears a likely role, specifically because methylation requires a reservoir of two electrons, oxidation state changes for both Ni-d and the [4Fe-4S] cluster appear unlikely. Ni-d with its square-planar coordination is difficult to reduce and no changes of its oxidation state have been observed so far. The [4Fe-4S] cluster of ACS_{Mt} can be reduced at low redox potentials when Ni-p is removed [184, 185], but only with a reduction rate about 200-fold slower than the rate of methyl transfer [186]. Furthermore, both Ni-d and the [4Fe-4S] cluster do not have obvious open coordination sites, making them unlikely binding positions for substrates. Thus the sole role of Ni-d and the [4Fe-4S] cluster may be in (in)directly modulating the electronic properties of Ni-p and to anchor the coordinating cysteine thiolates. Looking at it from the perspective of Ni-p, the charge of the three thiolate ligands are neutralized by binding to other charged metals. Thus, one role for Ni-d and the [4Fe-4S] cluster would be charge neutralization [162]. However, that is likely

not the complete story and electronic interactions, as evident from the spin delocalization in the Ni,Fe-C state, and the two additional metal sites likely contribute in other, (currently) less obvious ways to the properties of Ni-p.

In addition to the metal sites, the conformation of the three ACS domains likely directs and controls substrate access to the active site. The crystal structures of ACS from *M. thermoacetica* and *C. hydrogenoformans* containing Ni-p show cluster A and specifically the square-planar Ni-p exposed to the solvent. As the three domains are in a more stretched conformation, this is called the "open" conformation. In this conformation the tunnel connecting cluster C and cluster A is closed by the side chain of Phe512 (ACS_{Mt} numbering) [103, 145]. In contrast, in the structures of ACS_{Mt} containing Zn^{2+} or Cu^{+}, both with tetrahedral ligand arrangement, the three ACS domains interact more tightly, most visibly by a closer interaction of the N-terminal and C-terminal domain. This domain organization has been termed the "closed" conformation [102, 103]. In this conformation the side chain of Phe512 moved aside opening the channel between cluster C and cluster A. Although this conformation has not been directly shown to be relevant for the active Ni-p containing ACS, it is likely of physiological relevance as the presence of the N-terminal domain and Phe512 determines the catalytic repertoire of ACS catalyzing reversible acetyl-CoA synthesis [149, 168]. In summary, while most metalloenzymes may be treated as more or less solid protein matrixes with embedded metal(s) site(s), in ACS conformational changes drive and control the overall reaction and thus the observed kinetics.

5. UNIFYING FEATURES OF BIOLOGICAL Ni, Fe, S SITES

The three enzymes have some obvious parallels which may help not only to appreciate their uniqueness, but also to learn from them for future research. These parallels include their physiological roles and evolution, their architecture, and the central role of Ni versus a supportive role of Fe in small molecule activation. All three enzymes are required in ancestral autotrophic pathways converting CO_2 and H_2 into organic matter. Acetogenesis and methanogenesis share not only their ancestral roots but rely like no other known biochemical pathways on the catalytic proficiency of Ni-containing enzymes. Why Ni enzymes are so abundant in these two pathways may have several reasons: an evolution from an abiotic background of Ni,Fe,S minerals, the bioavailability of the metals and parallels in the evolution of their maturation systems. As recently shown, the two ATPases required to insert Ni into cluster C (CooC) and cluster A (AcsF) share a common ancestor and likely very similar metal insertion strategies [187, 188]. The sequence identity between CooC and AcsF is actually surprising when one thinks on the different number of Ni ions and architectures of the respective target clusters. The AcsF/CooC homologous HypB protein is involved in GTP/GDP-dependent Ni insertion into [NiFe] hydrogenases [189].

The three Ni,Fe,S active sites have obviously different architectures and are even distinguished by different Ni to Fe ratios, 1:1 in [NiFe] hydrogenase, 1:4 in

CODH, and 2:4 in ACS. Nevertheless, they share some similarities. In contrast to other Ni-containing enzymes, the Ni ions are predominantly coordinated by S-containing ligands. This allows more covalency in the ligand-Ni interaction and probably supports the redox activity of Ni in [NiFe] hydrogenase and ACS enabling redox transitions in the low potential regime. Again in contrast to other Ni-containing enzymes, Ni is not alone but close to Fe. Ni-Fe distances in the three enzymes vary between 2.5 to 3.0 Å, which may not be surprising as the distance is restrained by linking the two metals by one or more S-containing ligands. Still, it may produce similar electronic interactions allowing similar Ni-X-Fe ligand complexes such as Ni-Fe hydrides or even metal–metal bonds as recently suggested [137]. These possibilities touch on one of the tantalizing questions in the mechanisms of CODH and ACS: the source and storage place for two electrons. In [NiFe] hydrogenase Ni manages to change by only one unit despite catalyzing a two-electron chemistry due to hydride formation. Hydride formation has been suggested to be energetically feasible for CODH based on DFT calculations [138]. But there is no experimental evidence for the presence of hydrides in ACS and CODH, which should also result in H_2 formation if not shielded from protons.

Despite the differences in reactivity and their substrates, there are also some parallels in the role of Ni versus Fe, which may reflect on the respective properties. In all three enzymes discussed here, it is the Ni ion that binds and activates the substrate, and it is redox-active. Ni seems predestined for this role as it is held in a strained coordination geometry. The Fe ions seem to have a more passive role: Fe helps directly in stabilizing bound substrates in CODH and [NiFe] hydrogenase and is likely modulating the electronic properties of all three active sites via metal-metal interactions. While the active site Fe ions of CODH and ACS have a tetrahedral coordination and therefore a high-spin state, the strong-field ligands in hydrogenases produce a low-spin state. As a last and maybe most speculative link, all three enzymes require CO, either as a stably bound ligand or as a substrate. Indeed, no other enzymes using CO as a substrate or ligand are known to date (apart from the other two types of hydrogenases and of Mo-containing CODH), so that the evolutionary link between the three enzymes is reinforced. Again, one is reminded on the primordial ancestry of the enzymes, in an environment where CO may have been abundant.

ACKNOWLEDGMENTS

Funding of our research on the discussed enzymes by the DFG (DO 785) and the German excellence initiative (EXC 314 − UniCat) is gratefully acknowledged.

ABBREVIATIONS AND DEFINITIONS

ACS	acetyl-CoA synthase
ACDS	acetyl-CoA decarbonylase/synthase
Ch	*Carboxydothermus hydrogenoformans*

CN^-	cyanide
CoA	coenzyme A
CODH	carbon monoxide dehydrogenase
CODH/ACS	complex between carbon monoxide dehydrogenase and acetyl-CoA synthase
CoFeSP	corrinoid iron sulfur protein
DFT	density functional theory
ENDOR	electron nuclear double resonance
EPR	electron paramagnetic resonance
FMN	flavine mononucleotide
HYSCORE	hyperfine sub-level correlation (spectroscopy)
IR	infrared
LUCA	last universal common ancestor
Mt	*Moorella thermoacetica*
Ni-d	distal Ni ion in acetyl-CoA synthase (in Figure 8 Ni_d)
Ni-p	proximal Ni ion in acetyl-CoA synthase (in Figure 8 Ni_p)
XAS	X-ray absorption spectroscopy

REFERENCES

1. A. M. Appel, J. E. Bercaw, A. B. Bocarsly, H. Dobbek, D. L. DuBois, M. Dupuis, J. G. Ferry, E. Fujita, R. Hille, P. J. A. Kenis, C. A. Kerfeld, R. H. Morris, C. H. F. Peden, A. R. Portis, S. W. Ragsdale, T. B. Rauchfuss, J. N. H. Reek, L. C. Seefeldt, R. K. Thauer, G. L. Waldrop, *Chem. Rev.* **2013**, *113*, 6621–6658.
2. M. Can, F. A. Armstrong, S. W. Ragsdale, *Chem. Rev.*. **2014**, *114*, 4149–4174.
3. J.-H. Jeoung, J. Fesseler, S. Goetzl, H. Dobbek, *Met. Ions Life Sci* **2014**, *14*, 37–69.
4. H. Ogata, W. Lubitz, Y. Higuchi, *Dalton Trans.* **2009**, 7577–7587.
5. H. Ogata, W. Lubitz, Y. Higuchi, *J. Biochem. Tokyo* **2016**, *160*, 251–258.
6. Y. Shomura, Y. Higuchi, *Rev. Inorg. Chem.* **2013**, *33*, doi: 10.1515/revic-2013–0005.
7. H. S. Shafaat, O. Rüdiger, H. Ogata, W. Lubitz, *Biochim. Biophys. Acta – Bioenergetics* **2013**, *1827*, 986–1002.
8. A. F. Holleman, E. Wiberg, N. Wiberg, G. Fischer, *Lehrbuch der Anorganischen Chemie*, Walter De Gruyter, Berlin, New York, **2007**, 1–2186
9. J. L. Boer, S. B. Mulrooney, R. P. Hausinger, *Arch. Biochem. Biophys.* **2013**, *544*, 142–152.
10. C. J. Zeer-Wanklyn, D. B. Zamble, *Curr. Op. Chem. Biol.* **2017**, *37*, 80–88.
11. H. Kaluarachchi, K. C. Chan Chung, D. B. Zamble, *Nat. Prod. Rep.* **2010**, *27*, 681.
12. W. Nitschke, S. E. McGlynn, E. J. Milner-White, M. J. Russell, *Biochim. Biophys. Acta – Bioenergetics* **2013**, *1827*, 871–881.
13. P. S. Adam, G. Borrel, S. Gribaldo, *Proc. Natl. Acad. Sci.* **2018**, *115*, E1166–E1173.
14. M. J. Russell, W. Martin, *Trends Biochem. Sci.* **2004**, *29*, 358–363.
15. H. Beinert, *Eur. J. Biochem.* **2000**, *267*, 5657–5664.
16. L. Noodleman, T. Lovell, T. Liu, F. Himo, *Curr. Op. Chem. Biol.* **2002**, *6*, 259–273.
17. J. A. Turner, *Science* **2004**, *305*, 972–974.
18. R. F. Service, *Science* **2004**, *305*, 958–961.
19. W. Lubitz, H. Ogata, O. Rüdiger, E. Reijerse, *Chem. Rev.* **2014**, *114*, 4081–4148.
20. J. W. Peters, W. N. Lanzilotta, B. J. Lemon, L. C. Seefeldt, *Science* **1998**, 1853–1858.

21. Y. Nicolet, C. Piras, P. Legrand, C. E. Hatchikian, J. C. Fontecilla-Camps, *Structure* **1999**, *7*, 13–23.
22. A. J. Pierik, M. Hulstein, W. R. Hagen, S. P. J. Albracht, *Eur. J. Biochem.* **1998**, *258*, 572–578.
23. T. Watanabe, T. Wagner, G. Huang, J. Kahnt, K. Ataka, U. Ermler, S. Shima, *Angew. Chem. Int. Ed. Engl.* **2019**, *58*, 3506–3510.
24. A. Volbeda, J. C. Fontecilla-Camps, in *Bioorganometallic Chemistry*, Volume 17 of *Topics in Organometallic Chemistry*, Ed. G. Simonneaux, Springer-Verlag, Berlin/ Heidelberg, Germany, 2006, pp. 57–82.
25. J. C. Fontecilla-Camps, A. Volbeda, C. Cavazza, Y. Nicolet, *Chem. Rev.* **2007**, *107*, 4273–4303.
26. P. M. Vignais, B. Billoud, J. Meyer, *FEMS Microbiol. Rev.* **2001**, *25*, 455–501.
27. P. M. Vignais, B. Billoud, *Chem. Rev.* **2007**, *107*, 4206–4272.
28. G. Kulkarni, T. D. Mand, W. W. Metcalf, *MBio* **2018**, *9*, 28.
29. G. Kulkarni, D. M. Kridelbaugh, A. M. Guss, W. W. Metcalf, *Proc. Natl. Acad. Sci. USA* **2009**, *106*, 15915–15920.
30. K. Fiebig, B. Friedrich, *Eur. J. Biochem.* **1989**, *184*, 79–88.
31. R. K. Thauer, A.-K. Kaster, M. Goenrich, M. Schick, T. Hiromoto, S. Shima, *Annu. Rev. Biochem.* **2010**, *79*, 507–536.
32. S. A. Ensign, P. W. Ludden, *J. Biol. Chem.* **1991**, *266*, 18395–18403.
33. B. Soboh, D. Linder, R. Hedderich, *Eur. J. Biochem.* **2002**, *269*, 5712–5721.
34. A. Volbeda, M. H. Charon, C. Piras, E. C. Hatchikian, M. Frey, J. C. Fontecilla-Camps, *Nature* **1995**, *373*, 580–587.
35. D. J. Mills, S. Vitt, M. Strauss, S. Shima, J. Vonck, *eLife* **2013**, *2*, 7237.
36. M. Allegretti, D. J. Mills, G. McMullan, W. Kühlbrandt, J. Vonck, *eLife* **2014**, *3*, e01963.
37. S. Vitt, K. Ma, E. Warkentin, J. Moll, A. J. Pierik, S. Shima, U. Ermler, *J. Mol. Biol.* **2014**, *426*, 2813–2826.
38. Y. Shomura, M. Taketa, H. Nakashima, H. Tai, H. Nakagawa, Y. Ikeda, M. Ishii, Y. Igarashi, H. Nishihara, K.-S. Yoon, S. Ogo, S. Hirota, Y. Higuchi, *Science* **2017**, *357*, 928–931.
39. C. Schäfer, M. Bommer, S. E. Hennig, J.-H. Jeoung, H. Dobbek, O. Lenz, *Structure* **2016**, *24*, 285–292.
40. P. Wulff, C. Thomas, F. Sargent, F. A. Armstrong, *J. Biol. Inorg. Chem.* **2016**, *21*, 121–134.
41. E. Garcin, X. Vernede, E. C. Hatchikian, A. Volbeda, M. Frey, J. C. Fontecilla-Camps, *Structure* **1999**, *7*, 557–566.
42. M. Rousset, Y. Montet, B. Guigliarelli, N. Forget, M. Asso, P. Bertrand, J. C. Fontecilla-Camps, E. C. Hatchikian, *Proc. Natl. Acad. Sci. USA* **1998**, *95*, 11625–11630.
43. A. Abou Hamdan, S. Dementin, P.-P. Liebgott, O. Gutierrez-Sanz, P. Richaud, A. L. De Lacey, M. Rousset, P. Bertrand, L. Cournac, C. Léger, *J. Am. Chem. Soc.* **2012**, *134*, 8368–8371.
44. M.-E. Pandelia, W. Lubitz, W. Nitschke, *Biochim. Biophys. Acta – Bioenergetics* **2012**, *1817*, 1565–1575.
45. J. Fritsch, P. Scheerer, S. Frielingsdorf, S. Kroschinsky, B. Friedrich, O. Lenz, C. M. T. Spahn, *Nature* **2011**, *479*, 249–252.
46. Y. Shomura, K.-S. Yoon, H. Nishihara, Y. Higuchi, *Nature* **2011**, *479*, 253–256.
47. A. Volbeda, P. Amara, C. Darnault, J.-M. Mouesca, A. Parkin, M. M. Roessler, F. A. Armstrong, J. C. Fontecilla-Camps, *Proc. Natl. Acad. Sci. USA* **2012**, *109*, 5305–5310.
48. S. Frielingsdorf, J. Fritsch, A. Schmidt, M. Hammer, J. Löwenstein, E. Siebert, V. Pelmenschikov, T. Jaenicke, J. Kalms, Y. Rippers, F. Lendzian, I. Zebger, C. Teutloff,

M. Kaupp, R. Bittl, P. Hildebrandt, B. Friedrich, O. Lenz, P. Scheerer, *Nat. Chem. Biol.* **2014**, *10*, 378–385.
49. P.-P. Liebgott, F. Leroux, B. Burlat, S. Dementin, C. Baffert, T. Lautier, V. Fourmond, P. Ceccaldi, C. Cavazza, I. Meynial-Salles, P. Soucaille, J. C. Fontecilla-Camps, B. Guigliarelli, P. Bertrand, M. Rousset, C. Léger, *Nat. Chem.* **2010**, *6*, 63–70.
50. P.-P. Liebgott, A. L. De Lacey, B. Burlat, L. Cournac, P. Richaud, M. Brugna, V. M. Fernandez, B. Guigliarelli, M. Rousset, C. Léger, S. Dementin, *J. Am. Chem. Soc.* **2011**, *133*, 986–997.
51. A. Volbeda, E. Garcin, C. Piras, A. L. De Lacey, V. M. Fernandez, E. C. Hatchikian, M. Frey, J. C. Fontecilla-Camps, *J. Am. Chem. Soc.* **1996**, *118*, 12989–12996.
52. R. P. Happe, W. Roseboom, A. J. Pierik, S. P. Albracht, K. A. Bagley, *Nature* **1997**, *385*, 126–126.
53. M. C. Marques, R. Coelho, A. L. De Lacey, I. A. C. Pereira, P. M. Matias, *J. Mol. Biol.* **2010**, *396*, 893–907.
54. H. Ogata, P. Kellers, W. Lubitz, *J. Mol. Biol.* **2010**, *402*, 428–444.
55. H. Ogata, S. Hirota, A. Nakahara, H. Komori, N. Shibata, T. Kato, K. Kano, Y. Higuchi, *Structure* **2005**, *13*, 1635–1642.
56. A. Volbeda, L. Martin, C. Cavazza, M. Matho, B. W. Faber, W. Roseboom, S. P. J. Albracht, E. Garcin, M. Rousset, J. C. Fontecilla-Camps, *J. Biol. Inorg. Chem.* **2005**, *10*, 239–249.
57. Y. Higuchi, H. Ogata, K. Miki, N. Yasuoka, T. Yagi, *Structure* **1999**, *7*, 549–556.
58. M. Brecht, M. van Gastel, T. Buhrke, B. Friedrich, W. Lubitz, *J. Am. Chem. Soc.* **2003**, *125*, 13075–13083.
59. H. Ogata, K. Nishikawa, W. Lubitz, *Nature* **2015**, *520*, 571–574.
60. M. Y. Darensbourg, W. Weigand, *Eur. J. Inorg. Chem.* **2011**, 994–1004.
61. H. Ogata, Y. Mizoguchi, N. Mizuno, K. Miki, S.-I. Adachi, N. Yasuoka, T. Yagi, O. Yamauchi, S. Hirota, Y. Higuchi, *J. Am. Chem. Soc.* **2002**, *124*, 11628–11635.
62. M. Horch, L. Lauterbach, M. Saggu, P. Hildebrandt, F. Lendzian, R. Bittl, O. Lenz, I. Zebger, *Angew. Chem. Int. Ed. Engl.* **2010**, *49*, 8026–8029.
63. S. Kamali, H. Wang, D. Mitra, H. Ogata, W. Lubitz, B. C. Manor, T. B. Rauchfuss, D. Byrne, V. Bonnefoy, F. E. Jenney, M. W. W. Adams, Y. Yoda, E. Alp, J. Zhao, S. P. Cramer, *Angew. Chem. Int. Ed. Engl.* **2013**, *52*, 724–728.
64. D. Millo, P. Hildebrandt, M.-E. Pandelia, W. Lubitz, I. Zebger, *Angew. Chem. Int. Ed.* **2011**, *50*, 2632–2634.
65. W. Lubitz, E. Reijerse, M. van Gastel, *Chem. Rev.* **2007**, *107*, 4331–4365.
66. K. A. Vincent, F. A. Armstrong, *Inorg. Chem.* **2005**, *44*, 798–809.
67. K. A. Vincent, J. A. Cracknell, O. Lenz, I. Zebger, B. Friedrich, F. A. Armstrong, *Proc. Natl. Acad. Sci. USA* **2005**, *102*, 16951–16954.
68. C. Léger, S. J. Elliott, K. R. Hoke, L. J. C. Jeuken, A. K. Jones, F. A. Armstrong, *Biochemistry* **2003**, *42*, 8653–8662.
69. R. M. Evans, E. J. Brooke, S. A. M. Wehlin, E. Nomerotskaia, F. Sargent, S. B. Carr, S. E. V. Phillips, F. A. Armstrong, *Nat. Chem. Biol.* **2015**, *12*, 46–50.
70. H. Tai, K. Nishikawa, M. Suzuki, Y. Higuchi, S. Hirota, *Angew. Chem. Int. Ed. Engl.* **2014**, *53*, 13817–13820.
71. Y. Montet, P. Amara, A. Volbeda, X. Vernede, E. C. Hatchikian, M. J. Field, M. Frey, J. C. Fontecilla-Camps, *Nat. Struct. Biol.* **1997**, *4*, 523–526.
72. V. H. Teixeira, A. M. Baptista, C. M. Soares, *Biophys. J.* **2006**, *91*, 2035–2045.
73. M. Horch, J. Schoknecht, M. A. Mroginski, O. Lenz, P. Hildebrandt, I. Zebger, *J. Am. Chem. Soc.* **2014**, *136*, 9870–9873.
74. M. Bruschi, L. Gioia, G. Zampella, M. Reiher, P. Fantucci, M. Stein, *J. Biol. Inorg. Chem.* **2004**, *9*, 873–884.

75. M. L. Helm, M. P. Stewart, R. M. Bullock, M. R. DuBois, D. L. DuBois, *Science* **2011**, *333*, 863–866.
76. M. Medina, E. C. Hatchikian, R. Cammack, *Biochim. Biophys. Acta – Bioenergetics* **1996**, *1275*, 227–236.
77. C. S. A. Baltazar, M. C. Marques, C. M. Soares, A. M. DeLacey, I. A. C. Pereira, P. M. Matias, *Eur. J. Inorg. Chem.* **2011**, 948–962.
78. C. S. A. Baltazar, V. H. Teixeira, C. M. Soares, *J. Biol. Inorg. Chem.* **2012**, *17*, 543–555.
79. A. Parkin, G. Goldet, C. Cavazza, J. C. Fontecilla-Camps, F. A. Armstrong, *J. Am. Chem. Soc.* **2008**, *130*, 13410–13416.
80. A. L. De Lacey, C. Gutiérrez-Sánchez, V. M. Fernandez, I. Pacheco, I. A. C. Pereira, *J. Biol. Inorg. Chem.* **2008**, *13*, 1315–1320.
81. D. A. Grahame, E. DeMoll, *Biochemistry* **1995**, *34*, 4617–4624.
82. R. K. Thauer, *Biochim. Biophys. Acta – Bioenergetics* **1990**, *1018*, 256–259.
83. R. H. Crabtree, *The Organometallic Chemistry of the Transition Metals*, John Wiley & Sons, Inc., Hoboken, NJ, USA, 2005, 1–546.
84. O. Meyer, H. G. Schlegel, *Annu. Rev. Microbiol.* **1982**, *37*, 277–310.
85. H. Dobbek, L. Gremer, R. Kiefersauer, R. Huber, O. Meyer, *Proc. Natl. Acad. Sci. USA* **2002**, *99*, 15971–15976.
86. P. A. Lindahl, B. Chang, *Orig. Life Evol. Biosph.* **2001**, *31*, 403–434.
87. P. Lindahl, *Biochemistry* **2002**, *41*, 2097–2105.
88. D. A. Grahame, *J. Biol. Chem.* **1991**, *266*, 22227–22233.
89. Y. R. Dai, D. W. Reed, J. H. Millstein, P. L. Hartzell, D. A. Grahame, E. DeMoll, *Arch. Microbiol.* **1998**, *169*, 525–529.
90. V. Svetlitchnyi, C. Peschel, G. Acker, O. Meyer, *J. Bacteriol.* **2001**, *183*, 5134–5144.
91. D. Bonam, P. W. Ludden, *J. Biol. Chem.* **1987**, *262*, 2980–2987.
92. M. Wu, Q. Ren, A. Scott Durkin, S. C. Daugherty, L. M. Brinkac, R. J. Dodson, R. Madupu, S. A. Sullivan, J. F. Kolonay, W. C. Nelson, L. J. Tallon, K. M. Jones, L. E. Ulrich, J. M. Gonzalez, I. B. Zhulin, F. T. Robb, J. A. Eisen, *PLoS Genet.* **2005**, *1*, 563–574.
93. J. E. Clark, S. W. Ragsdale, L. G. Ljungdahl, J. Wiegel, *J. Bacteriol.* **1982**, *151*, 507–509.
94. M. Rother, E. Oelgeschläger, W. W Metcalf, *Arch. Microbiol.* **2007**, *188*, 463–472.
95. E. Oelgeschläger, M. Rother, *Arch. Microbiol.***2008**, *190*, 257–269.
96. Q. Lin, X. Fang, A. Ho, J. Li, X. Yan, B. Tu, C. Li, J. Li, M. Yao, X. Li, *Appl. Microbiol. Biotech.* **2017**, *101*, 7303–7316.
97. F. L. Sousa, W. F. Martin, *Biochim. Biophys. Acta* **2014**, *1837*, 964–981.
98. F. L. Sousa, S. Nelson-Sathi, W. F. Martin, *Biochim. Biophys. Acta* **2016**, *1857*, 1027–1038.
99. E. Kocsis, M. Kessel, E. DeMoll, D. A. Grahame, *J. Struct. Biol.* **1999**, *128*, 165–174.
100. J. Fesseler, J.-H. Jeoung, H. Dobbek, *Angew. Chem. Int. Ed. Engl.* **2015**, *54*, 8560–8564.
101. W. Gong, B. Hao, Z. Wei, D. J. Ferguson, T. Tallant, J. A. Krzycki, M. K. Chan, *Proc. Natl. Acad. Sci. USA* **2008**, *105*, 9558–9563.
102. T. I. Doukov, T. M. Iverson, J. Seravalli, S. W. Ragsdale, C. L. Drennan, *Science* **2002**, *298*, 567–572.
103. C. Darnault, A. Volbeda, E. J. Kim, P. Legrand, X. Vernède, P. A. Lindahl, J. C. Fontecilla-Camps, *Nat. Struct. Biol.* **2003**, *10*, 271–279.
104. H. Dobbek, V. Svetlitchnyi, L. Gremer, R. Huber, O. Meyer, *Science* **2001**, *293*, 1281–1285.
105. C. L. Drennan, J. Heo, M. D. Sintchak, E. Schreiter, P. W. Ludden, *Proc. Natl. Acad. Sci. USA* **2001**, *98*, 11973–11978.

106. L. Domnik, M. Merrouch, S. Goetzl, J.-H. Jeoung, C. Léger, S. Dementin, V. Fourmond, H. Dobbek, *Angew. Chem. Int. Ed.* **2017**, *56*, 15466–15469.
107. E. C. Wittenborn, M. Merrouch, C. Ueda, L. Fradale, C. Léger, V. Fourmond, M.-E. Pandelia, S. Dementin, C. L. Drennan, *Elife* **2018**, *7*, 213.
108. D. Rees, *Annu. Rev. Biochem.* **2002**.
109. J.-H. Jeoung, H. Dobbek, *J. Biol. Inorg. Chem.* **2012**, *17*, 167–173.
110. E. J. Kim, J. Feng, M. R. Bramlett, P. A. Lindahl, *Biochemistry* **2004**, *43*, 5728–5734.
111. P.-H. Wang, M. Bruschi, L. De Gioia, J. Blumberger, *J. Am. Chem. Soc.* **2013**, *135*, 9493–9502.
112. S. A. Ensign, *Biochemistry* **1995**, *34*, 5372–5381.
113. A. Parkin, J. Seravalli, K. A. Vincent, S. W. Ragsdale, F. A. Armstrong, *J. Am. Chem. Soc.* **2007**, *129*, 10328–10329.
114. V. C. C. Wang, M. Can, E. Pierce, S. W. Ragsdale, F. A. Armstrong, *J. Am. Chem. Soc.* **2013**, *135*, 2198–2206.
115. A. Ciaccafava, D. Tombolelli, L. Domnik, J.-H. Jeoung, H. Dobbek, M. A. Mroginski, I. Zebger, P. Hildebrandt, *Angew. Chem. Int. Ed. Engl.* **2017**, *56*, 7398–7401.
116. W. P. Lu, S. W. Ragsdale, *J. Biol. Chem.* **1991**, *266*, 3554–3564.
117. V. C. C. Wang, S. T. A. Islam, M. Can, S. W. Ragsdale, F. A. Armstrong, *J. Phys. Chem. B* **2015**, *119*, 13690–13697.
118. T. Inoue, K. Takao, T. Yoshida, K. Wada, T. Daifuku, Y. Yoneda, K. Fukuyama, Y. Sako, *Biochem. Biophys. Res. Comm.* **2013**, *441*, 13–17.
119. M. Kumar, S. W. Ragsdale, *J. Am. Chem. Soc.* **1995**, *117*, 11604–11605.
120. J.-H. Jeoung, H. Dobbek, *Science* **2007**, *318*, 1461–1464.
121. Y. Kung, T. I. Doukov, J. Seravalli, S. W. Ragsdale, C. L. Drennan, *Biochemistry* **2009**, *48*, 7432–7440.
122. H. Dobbek, V. Svetlitchnyi, J. Liss, O. Meyer, *J. Am. Chem. Soc.* **2004**, *126*, 5382–5387.
123. J. Feng, P. A. Lindahl, *J. Am. Chem. Soc.* **2004**, *126*, 9094–9100.
124. Y. Jean, *Molecular Orbitals of Transition Metal Complexes*, Oxford University Press, Oxford, New York, 2005, 1–275.
125. J.-H. Jeoung, H. Dobbek, *J. Am. Chem. Soc.* **2009**, *131*, 9922–9923.
126. A. Ciaccafava, D. Tombolelli, L. Domnik, J. Fesseler, J.-H. Jeoung, H. Dobbek, M. A. Mroginski, I. Zebger, P. Hildebrandt, *Chem. Sci.* **2016**, *7*, 3162–3171.
127. J. Feng, P. A. Lindahl, *Biochemistry* **2004**, *43*, 1552–1559.
128. Z. Hu, N. Spangler, M. Anderson, J. Xia, P. W. Ludden, P. A. Lindahl, E. Münck, *J. Am. Chem. Soc.* **1996**, *118*, 830–844.
129. P. Lindahl, E. Münck, S. W. Ragsdale, *J. Biol. Chem.* **1990**, *265*, 3873–3879.
130. N. Spangler, P. Lindahl, V. Bandarian, P. W. Ludden, *J. Biol. Chem.* **1996**, *271*, 7973–7977.
131. S. A. Ensign, M. R. Hyman, P. W. Ludden, *Biochemistry* **1989**, *28*, 4973–4979.
132. S.-W. Ha, M. Korbas, M. Klepsch, W. Meyer-Klaucke, O. Meyer, V. Svetlitchnyi, *J. Biol. Chem.* **2007**, *282*, 10639–10646.
133. M. E. Anderson, P. A. Lindahl, V. J. DeRose, B. M. Hoffman, *J. Am. Chem. Soc.* **1993**, *115*, 12204–12205.
134. V. J. DeRose, J. Telser, M. E. Anderson, P. A. Lindahl, B. M. Hoffman, *J. Am. Chem. Soc.* **1998**, *120*, 8767–8776.
135. S. A. Macgregor, Z. Lu, O. Eisenstein, R. H. Crabtree, *Inorg. Chem.* **1994**, *33*, 3616–3618.
136. P. Lindahl, S. Ragsdale, *J. Biol. Chem.* **1990**, *265*, 3880–3888.
137. P. A. Lindahl, *J. Inorg. Biochem.* **2012**, *106*, 172–178.

138. P. Amara, J.-M. Mouesca, A. Volbeda, J. C. Fontecilla-Camps, *Inorg. Chem.* **2011**, *50*, 1868–1878.
139. M. E. Anderson, P. A. Lindahl, *Biochem.* **1996**, *35*, 8371–8380.
140. S. W. Ragsdale, E. Pierce, *Biochim. Biophys. Acta – Proteins & Proteomics* **2008**, *1784*, 1873–1898.
141. G. Fuchs, *Annu. Rev. Microbiol.* **2011**, *65*, 631–658.
142. R. K. Thauer, S. Shima, *Nature* **2006**, *440*, 878–879.
143. R. K. Thauer, D. M. Zinkhan, A. M. Spormann, *Annu. Rev. Microbiol.* **1989**, *43*, 43–67.
144. D. Grahame, E. DeMoll, *J. Biol. Chem.* **1996**, *271*, 8352–8358.
145. V. Svetlitchnyi, H. Dobbek, W. Meyer-Klaucke, T. Meins, B. Thiele, P. Römer, R. Huber, O. Meyer, *Proc. Natl. Acad. Sci. USA* **2004**, *101*, 446–451.
146. A. Volbeda, C. Darnault, X. Tan, P. A. Lindahl, J. C. Fontecilla-Camps, *Biochemistry* **2009**, *48*, 7916–7926.
147. T. I. Doukov, L. C. Blasiak, J. Seravalli, S. W. Ragsdale, C. L. Drennan, *Biochemistry* **2008**, *47*, 3474–3483.
148. S. Gencic, D. A. Grahame, *J. Biol. Chem.* **2003**, *278*, 6101–6110.
149. S. Gencic, K. Kelly, S. Ghebreamlak, E. C. Duin, D. A. Grahame, *Biochemistry* **2013**, *52*, 1705–1716.
150. M. L. Golden, C. M. Whaley, M. V. Rampersad, J. H. Reibenspies, R. D. Hancock, M. Y. Darensbourg, *Inorg. Chem.* **2005**, *44*, 875–883.
151. K. N. Green, S. P. Jeffery, J. H. Reibenspies, M. Y. Darensbourg, *J. Am. Chem. Soc.* **2006**, *128*, 6493–6498.
152. R. Panda, Y. Zhang, C. C. McLauchlan, P. Venkateswara Rao, F. A. Tiago de Oliveira, E. Münck, R. H. Holm, *J. Am. Chem. Soc.* **2004**, *126*, 6448–6459.
153. C. G. Riordan, *J. Biol. Inorg. Chem.* **2004**, *9*, 542–549.
154. M. C. Weiss, F. L. Sousa, N. Mrnjavac, S. Neukirchen, M. Roettger, S. Nelson-Sathi, W. F. Martin, *Nat. Microbiol.* **2016**, *1*, 16116.
155. E. L. Hegg, *Acc. Chem. Res.* **2004**, *37*, 775–783.
156. C. E. Webster, M. Y. Darensbourg, P. A. Lindahl, M. B. Hall, *J. Am. Chem. Soc.* **2004**, *126*, 3410–3411.
157. P. Amara, A. Volbeda, J. C. Fontecilla-Camps, M. J. Field, *J. Am. Chem. Soc.* **2005**, *127*, 2776–2784.
158. M. Can, L. J. Giles, S. W. Ragsdale, R. Sarangi, *Biochem.* **2017**, *56*, 1248–1260.
159. P. Schrapers, J. Ilina, C. M. Gregg, S. Mebs, J.-H. Jeoung, H. Dau, H. Dobbek, M. Haumann, *PLoS ONE* **2017**, *12*, e0171039.
160. D. Barondeau, P. Lindahl, *J. Am. Chem. Soc.* **1997**, *119*, 3959–3970.
161. R. Schenker, M. T. Mock, M. T. Kieber-Emmons, C. G. Riordan, T. C. Brunold, *Inorg. Chem.* **2005**, *44*, 3605–3617.
162. P. A. Lindahl, *J. Biol. Inorg. Chem.* **2004**, *9*, 516–524.
163. R. P. Schenker, T. C. Brunold, *J. Am. Chem. Soc.* **2003**, *125*, 13962–13963.
164. G. Bender, T. A. Stich, L. Yan, R. D. Britt, S. P. Cramer, S. W. Ragsdale, *Biochemistry* **2010**, *49*, 7516–7523.
165. S. Ragsdale, L. Ljungdahl, D. V. DerVartanian, *Biochem. Biophys. Res. Commun.* **1982**, *108*, 658–663.
166. S. W. Ragsdale, H. G. Wood, W. E. Antholine, *Proc. Natl. Acad. Sci. USA* **1985**, *82*, 6811–6814.
167. D. A. Grahame, S. Khangulov, E. DeMoll, *Biochemistry* **1996**, *35*, 593–600.
168. S. Gencic, E. C. Duin, D. A. Grahame, *J. Biol. Chem.* **2010**, *285*, 15450–15463.
169. M. Kumar, S. W. Ragsdale, *J. Am. Chem. Soc.* **1992**, *114*, 8713–8715.
170. S. J. George, J. Seravalli, S. W. Ragsdale, *J. Am. Chem. Soc.* **2005**, *127*, 13500–13501.

171. J. L. Craft, B. S. Mandimutsira, K. Fujita, C. G. Riordan, T. C. Brunold, *Inorg. Chem.* **2003**, *42*, 859–867.
172. T. C. Harrop, M. M. Olmstead, P. K. Mascharak, *Chem. Commun.* **2004**, 1744.
173. T. C. Harrop, M. M. Olmstead, P. K. Mascharak, *J. Am. Chem. Soc.* **2004**, *126*, 14714–14715.
174. A. C. Manesis, M. J. O'Connor, C. R. Schneider, H. S. Shafaat, *J. Am. Chem. Soc.* **2017**, *139*, 10328–10338.
175. J. Seravalli, S. W. Ragsdale, *J. Biol. Chem.* **2008**, *283*, 8384–8394.
176. D. L. Purich, *Enzyme Kinetics: Catalysis & Control*, Elsevier, Academic Press, Amsterdam, Boston, Heidelberg, **2010**, 1–892.
177. M. Kumar, W.-P. Lu, L. Liu, S. W. Ragsdale, *J. Am. Chem. Soc.* **1993**, *115*, 11646–11647.
178. C. M. Gorst, S. W. Ragsdale, *J. Biol. Chem.* **1991**, *266*, 20687–20693.
179. J. Seravalli, M. Kumar, S. W. Ragsdale, *Biochemistry* **2002**, *41*, 1807–1819.
180. E. L. Maynard, C. Sewell, P. A. Lindahl, *J. Am. Chem. Soc.* **2001**, *123*, 4697–4703.
181. X. S. Tan, C. Sewell, P. A. Lindahl, *J. Am. Chem. Soc.* **2002**, *124*, 6277–6284.
182. X. Tan, I. V. Surovtsev, P. A. Lindahl, *J. Am. Chem. Soc.* **2006**, *128*, 12331–12338.
183. X. Tan, H.-K. Loke, S. Fitch, P. A. Lindahl, *J. Am. Chem. Soc.* **2005**, *127*, 5833–5839.
184. J. Xia, Z. Hu, C. V. Popescu, P. A. Lindahl, E. Münck, *J. Am. Chem. Soc.* **1997**, *119*, 8301–8312.
185. M. R. Bramlett, A. Stubna, X. Tan, I. V. Surovtsev, E. Münck, P. A. Lindahl, *Biochemistry* **2006**, *45*, 8674–8685.
186. X. Tan, C. Sewell, Q. Yang, P. A. Lindahl, *J. Am. Chem. Soc.* **2003**, *125*, 318–319.
187. C. M. Gregg, S. Goetzl, J.-H. Jeoung, H. Dobbek, *J. Biol. Chem.* **2016**, *291*, 18129–18138.
188. H. Dobbek, in *Encyclopedia of Inorganic and Bioinorganic Chemistry*, Ed. R. A. Scott, John Wiley & Sons, Ltd, Chichester, UK, **2017**, *544*, 1–11.
189. M. R. Leach, D. B. Zamble, *Curr. Op. Chem. Biol.* **2007**, *11*, 159–165.

12

Zinc Fingers

Gaetano Malgieri, Luigi Russo, Gianluca D'Abrosca, Ilaria Baglivo, Paolo V. Pedone, Roberto Fattorusso, and Carla Isernia

Dipartimento di Scienze e Tecnologie Ambientali, Biologiche e Farmaceutiche, Università della Campania "L. Vanvitelli", I-81100 Caserta, Italy
<carla.isernia@unicampania.it>

Metal Ions in Life Sciences, Volume 20 Guest Editors: Martha E. Sosa Torres and Peter M. H. Kroneck
Series Editors: Astrid Sigel, Eva Freisinger, and Roland K. O. Sigel

https://doi.org/10.1515/9783110589757-012

Abstract: Zinc finger (ZF) domains, that represent the majority of the DNA-binding motifs in eukaryotes, are involved in several processes ranging from RNA packaging to transcriptional activation, regulation of apoptosis, protein folding and assembly, and lipid binding. While their amino acid composition varies from one domain to the other, a shared feature is the coordination of a zinc ion, with a structural role, by a different combination of cysteines and histidines. The classical zinc finger domain (also called Cys_2His_2) that represents the most common class, uses two cysteines and two histidines to coordinate the metal ion, and forms a compact ββα architecture consisting in a β-sheet and an α-helix. GAG-knuckle resembles the classical ZF, treble clef and zinc ribbon are also well represented in the human genome.

Zinc fingers are also present in prokaryotes. The first prokaryotic ZF domain found in the transcriptional regulator Ros protein was identified in *Agrobacterium tumefaciens*. It shows a Cys_2His_2 metal ion coordination sphere and folds in a domain significantly larger than its eukaryotic counterpart arranged in a βββαα topology. Interestingly, this domain does not strictly require the metal ion coordination to achieve the functional fold.

Here, we report what is known on the main classes of eukaryotic and prokarotic ZFs, focusing our attention to the role of the metal ion, the folding mechanism, and the DNA binding. The hypothesis of a horizontal gene transfer from prokaryotes to eukaryotes is also discussed.

Keywords: DNA binding · domain folding · horizontal gene transfer evolution · xenobiotic metal ions · zinc fingers

1. INTRODUCTION

The zinc-finger (ZF) motif represents the majority of the DNA-binding domains [1, 2], including 3 % of the human genes [3, 4] and, besides its DNA recognition function, is involved in many diverse processes: RNA packaging, transcriptional activation, regulation of apoptosis, protein folding and assembly, and lipid binding [5]. While the amino acid composition can vary from one domain to the other, a common ZF feature is the coordination of a structural Zn(II) ion by a different combination of cysteine and histidine: the most common is the Cys_2His_2 or classical ZF but also representative is the Cys_4 or $Cys_2HisCys$ zinc coordination sphere.

More recent is the discovery of the prokaryotic Cys_2His_2 ZF domain found for the first time by Chou and coworkers in 1998 [6] in the Ros protein from *Agrobacterium tumefaciens*. Since then, Ros homologues have been identified in many other bacteria, all bearing some intriguing new features that identify a family of domains with important differences to the eukaryotic domain. Here we will review differences and analogies of the two domains, summarizing the recent literature and offering significant insights into their structure/function relationships.

2. ZINC FINGER DOMAINS IN EUKARYOTES

Different in sequences and binding modes, a ZF domain is a solid scaffold that has been used in numerous and various specialized functions beside its originally found DNA-binding role: zinc sensing, protein folding, chromatin remodelling,

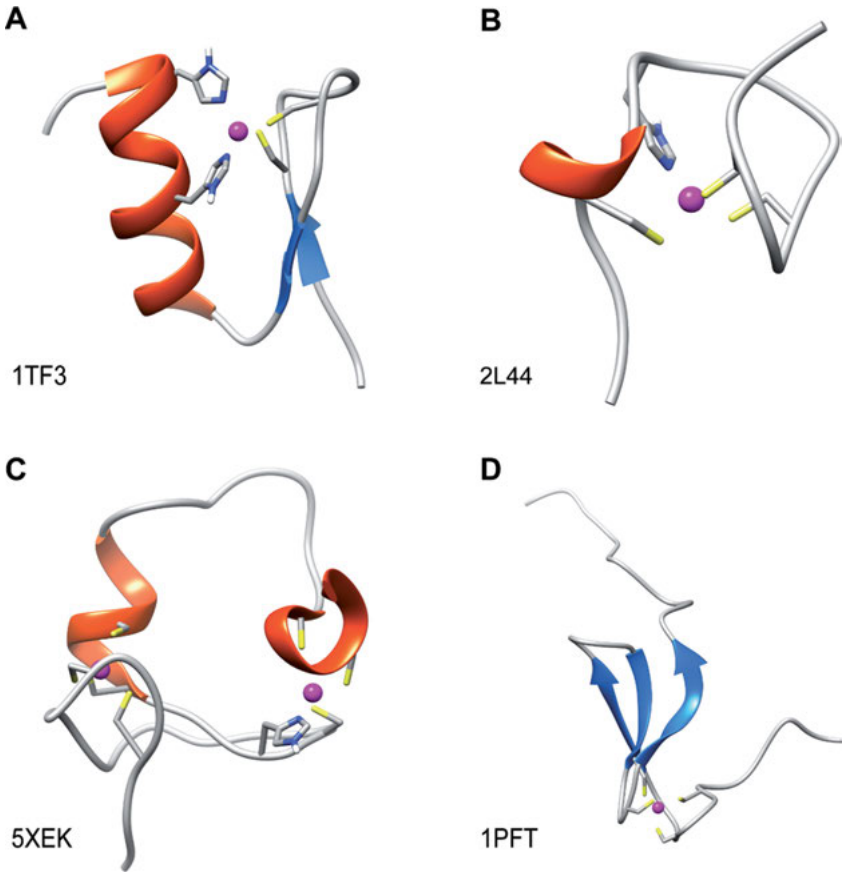


Figure 1. Ribbon drawing of the different classes of zinc fingers. (A) First finger of TFIIIA; (B) C-terminal zinc knuckle of the HIVNCp7; (C) C-terminal zinc finger of RING finger protein 141 (treble clef finger); (D) N-terminal domain of TFIIB (zinc ribbon). The PDB codes are reported in the figure.

cytoskeleton organization, epithelial development, gene transcription, translation, mRNA trafficking, and cell adhesion [2].

Each domain shows a different amino acid sequence which determines its binding abilities and is generally found in assemblies of more than one finger. The functional properties of each assembly is also dependent upon the number of domains, the length and composition of the spacer between the ZFs and higher order structures.

A ZF uses a different combination of cysteines and histidines to bind a Zn(II) ion in a cluster that stabilizes the three-dimensional structure of this small domain. The number of cysteines and histidine residues was originally used to classify ZFs while now they are grouped on the basis of the overall fold of the domain backbone, with classical, treble clef, GAG knuckle, and zinc ribbon representing the most common classes (Figure 1) [7]. In general, ZF, as defined

by Laity et al. [5], is a small functional domain that needs to coordinate at least one zinc ion to achieve its functional structure.

The metal ion has generally only a structural role, by increasing the conformational and thermal stability of the domain, and it is not involved in the interactions with the domain targets.

2.1. Classical Zinc Fingers

The most common class or fold group uses two cysteines and two histidines and is called Cys_2His_2 or classical ZF. The first classical ZF domain was discovered as a repeated unity in the transcription factor TFIIIA of *Xenopus laevis* [8, 9]. Since then, increasing evidences have supported crucial roles for ZFs in physiological processes from prokaryotes to eukaryotes, including the plant kingdom [10–16]. Cys-X_4-Cys-X_{12}-His-X_3-His is a general sequence for ZFs in the eukaryotic kingdom [17]. The Zn(II) ion is tetrahedrally coordinated and, together with a small hydrophobic core, helps to form a compact ββα architecture consisting of a β-sheet and an α-helix [18]. The α-helix is at the C-terminus of the domain and contains the two coordinating histidines while the two-stranded β-sheet includes the two coordinating cysteines and occurs at the N-terminus; in a few cases [19, 20], the β-sheet bears a third antiparallel β-strand. The ββα architecture allows the formation of a small compact hydrophobic core that favors the formation of the structure and further stabilizes it. Apart from these residues essential for the domain stability, non-conserved amino acids participate in tertiary interactions and molecular target recognition [21] conferring to each ZF different binding specificities and affinities [22].

The α-helix harbors generally polar and charged residues and it is typically engaged in contacts with base pairs of the DNA major groove conferring binding specificity to the domain.

2.2. GAG-Knuckle and Treble Clef

Although shorter (only 20 amino acids instead of 30), the GAG-knuckle fold group resembles the classical ZF: a zinc knuckle that reminds of the Cys_2His_2 β-hairpin, connects two short β-strands that are followed by a short α-helix or a loop. The motif bears a $Cys_2HisCys$ sequence that binds the Zn(II) ion: the two cysteines can be found on the knuckle while the other two ligands are found on the loop or at both ends of a short helix [7]. Motifs found in many proteins including retroviral nucleocapsid (NC) protein from HIV and other related viruses and RNA polymerase II have been assigned to this fold group. The NC zinc-knuckle shows a one-turn α-helix, while the RNA polymerase II shows a loop [7, 21, 23].

Treble clef fingers can be found in a variable group of proteins that often do not share similar sequences and functions. They all share a fold at their core [24] that consists of a β-hairpin at the N-terminus, a variable loop region, and an α-

helix at the C-terminus. Two of the structural Zn(II) ligands come from the knuckle while the other two are both located in the last turn of the α-helix. Treble clef fingers have been recognized in multidomain proteins which contain tandem or overlapping motifs; examples are RING and PHD zinc finger domains [7].

2.3. The Zinc Ribbon

The zinc ribbon is a very common ZF fold represented by two β-hairpins forming two structurally similar sub-sites (N- and C-terminal) that bind the Zn(II) ion [7, 21, 25].

The β-strands in the first β-hairpin are typically made by two to four residues. Most of the second β-hairpins bear a three-stranded antiparallel β-sheet whose β-strands can vary in length, but often are longer than those of the first subdomain. The two sub-sites can be at different distances and sometimes additional domains can be found between the two. This fold is found in proteins that show a discrete structural variability that is reflected by their limited sequence similarity, generally restricted only to the zinc ligands and the zinc binding motifs. Topoisomerase I, the regulatory β subunits of the kinase CK2β [7, 21, 25–27] and transcription factors TFIIS, TFIIB, as well as some ribosomal proteins are examples of zinc ribbon proteins.

Additionally, several other proteins have been indicated as belonging to the zinc ribbon class on the bases of similarity of the Zn(II) ligands or of the zinc knuckle motif instead of sequence similarity [7].

3. METAL BINDING AND FOLDING PROPERTIES

3.1. The Zn(II) Metal Ion

Proteins in general use Zn(II) as a cofactor for acid-base catalysis or to stabilize their structure. This latter is the case of the small ZF domains in which the Zn(II) ion has only a structural role and it is not involved in the function of the domain; the functions come from the other parts of the structure. However, the wide array of functions that can be performed by this domain heavily relies on the stability provided by the coordination of Zn(II) that allows ZF to be a stable structural template capable to perform functions more diverse than larger domains.

Structural, sequential, and chemical features of both the ZFs and the Zn(II) ion render the binding of Zn(II) to these domains highly favored. The binding is also dependent upon factors such as pH, ionic strength, temperature, and Zn(II) concentration. Free Zn(II) concentration in the cell can vary by two orders of magnitude, with an average value of around 10^{-10} M [28, 29]. Natural ZFs coordinate Zn(II) tetrahedrally with an affinity that can vary from low nanomolar to subfemtomolar [30, 31], typically using cysteinyl sulfurs and histidine

imidazole nitrogens. Such affinity mainly depends upon the acidity and spatial orientation of the cysteine thiol residues, as at physiological pH the imidazole ring is uncharged and will bind Zn(II) with no need for deprotonation. On the other hand, the cysteine thiol group is relatively basic with a pK_a between 8 and 9 in most structural sites [32] thus making cysteine deprotonation a critical event for Zn(II) binding [33–35]. The pK_a value is influenced by neighboring groups: those which are positively charged will favor ionization and will decrease the pK_a of thiols, while those which are negatively charged will cause an increase of their pK_a [36–39]. Sequential and structural factors have influences on these values. Approximately 50 % of ZFs present a Cys-X-X-Cys motif involved in β-turn formation. Acid-base properties of thiols differ between each specific metal binding motif with differences also between the two cysteines within the same motif. Acidity of cysteines changes in the case of ZFs with a higher number of them and in the case of motifs with more residues X between two binding cysteines in the β-turn [21].

3.2. Folding Mechanism

Molecular modeling studies have shown that the classical eukaryotic ZF folding process starts with the metal coordination by two cysteine residues that induces the formation of the β-hairpin; next the N-terminal histidine is engaged thus prompting the construction of a small hydrophobic core that stabilizes the ββα fold [40, 41]. Finally, the C-terminal histidine joins the metal coordination sphere allowing the complete formation of the helix that terminates the folding process [42, 43].

Zn(II) stabilizes the native structure and leads the whole process. Circular dichroism (CD) experiments have demonstrated that, upon mutation of the C-terminal histidine, the domain assumes a stable conformation while the mutation of the first histidine does not give a stable form [44, 45].

The stabilizing effects, resulting from hydrophobic interactions and hydrogen bonds present in the 3D structure of each ZF, determines the lower or higher affinity for Zn(II). In fact, removal of hydrophobic residues reduces the stability by more than one order of magnitude regardless of cysteine and histidine coordination [21].

3.3. Metal Replacement by Xenobiotic Ions

Naturally designed to have a high binding affinity towards their physiological structural Zn(II) ion, ZFs can also bind other metal ions with either lower or higher affinity. This opens up the possibility that certain metal ions under particular conditions (pH, temperature, ionic strength) might substitute the native structural Zn(II) ion in its coordination sphere. Such displacement can have a negative impact on the structure or the function of a ZF leading to toxic effects and carcinogenesis [46, 47]. Zn(II) is the best ion to grant structural stability to

ZFs as, under physiological conditions, it favors only the Zn(II) ($3d^{10}$) oxidation state and does not undergo redox chemistry. These features are key for ZFs to perform their functions without introducing unwanted reactivity. Cysteine oxidation, formation of mixed complexes, and incomplete metal ion coordination [16, 46–48] are the common observed structural problems shown upon Zn(II) substitution. Yet, isostructural substitution can also occur. It has been proposed that metal ions, such as Co(II), Ni(II), Pb(II), and Cd(II), can compete with Zn(II) for different ZF metal-binding sites. Co(II), because of its coordination properties similar to those of Zn(II), does not change the geometry of the binding site or the overall ZF structure. This feature, together with its very interesting spectroscopic and paramagnetic properties due to its incompletely filled $3d^7$ shell, has led to the common use of the Co(II) ion as a spectroscopic probe for the study of zinc binding sites in biomolecules [21, 49].

Also Ni(II), with its incompletely filled $3d^8$ shell, is very often used to investigate the metal binding properties [50] of Zn(II)-binding sites and ZFs having a Ni(II) ion can influence the tetrahedral geometry of the Zn(II)-binding site by rendering it more planar and thus sensible to oxidation [51].

Ni(II) appears to be able to substitute the native Zn(II) ion in ZFs only at really high molar excess. The measured K_d value for this ion has been found 6,700 times the one for Zn(II) in the case of Sp1–3 [52, 53] while 2,300 times higher [51, 53] in the case of XPA-ZF. In the first case, substitution of Zn(II) by Ni(II) affects protein-DNA interaction by changing the preferred DNA sequence recognized by Sp1–3. This unveils an impact of Ni(II)-induced folding on the mechanism of sequence specific recognition by ZF proteins [54], thus suggesting small differences in the structure of the ZF domains. In the case of XPA-ZF, Ni(II) induces a change from the tetrahedral structure to a square planar complex [51] thus disrupting the overall ZF structure.

Lead(II), a borderline Lewis acid, binds ZFs with K_d values that decrease depending on the number of cysteines present in the coordination site. A certain number of studies have reported that Pb(II) influences the DNA binding activity of ZF proteins TFIIIA, Sp1, and Egr-1 [53, 55–59]. In Cys_4 ZFs, Pb(II) shows an affinity higher than Zn(II) [42, 53] but it is not capable to stabilize the proper fold of the protein. When the coordination sphere bears histidine residues, Pb(II) does not interact with them strongly. Moreover, it has been proved using model ZFs that significant differences occur in sulfur-rich sites between the Pb(II) coordination geometry (trigonal pyramidal) and that of Zn(II) (tetrahedral). Such a difference has been proposed as a mechanism to influence Pb(II) binding on the structure and function of ZFs thus contributing to the problems associated with lead toxicity [53, 60].

Cadmium(II) is a highly toxic transition metal also at low concentrations [61–64] characterized by a high chemical similarity to Zn(II). For these reasons, effects of replacing Zn(II) by Cd(II) in different ZFs have been extensively studied, with the main focus on how this metal modulates the DNA-binding activity [65–71].

The specific domain sequence under consideration determines the effects of this substitution. Cd(II)-TFIIIA is not capable to specifically bind its cognate DNA [69, 70]. Cd(II)-SUP37 shows an overall domain fold similar to Zn(II)-

SUP37 but side chains involved in the interaction with DNA change their orientation, proposing a reduced binding affinity for its target DNA [63]. In the case of Sp1, Cd(II) does not alter its structure and the quantitative binding to its cognate DNA [71], while in the case of Cd(II)-Tramtrack a loss of helical character and a 10-fold decrease in DNA-binding affinity is observed [72].

4. DNA-BINDING MECHANISM

ZF binding to its DNA target juxtaposes a few amino acids of the α-helix to three base pairs in the DNA major groove. The identity of these amino acidic residues determines the specificity of each ZF (Figure 2); their variability gives the zinc finger domain its functional modular feature [73, 74]. This DNA interaction mode is shared by the classical ZF domain and by the treble clef finger [7]. Nature has exploited this recognition mechanism by creating proteins containing multiple ZF motifs: the cooperative action of a certain number of ZFs arranged in tandem gives to each protein sequence specificity and high affinity for DNA. Linker regions between each ZF usually form loops between each domain. Proteins comprehending multiple ZFs usually require a minimum of two domains for high-affinity DNA binding [75, 76]. Occasionally, a single ZF domain is sufficient for high-affinity DNA binding. In fact, it has been proved that the single ZFs of the Drosophila GAGA transcription factor [77, 78] and of the *Arabidopsis thaliana* SUPERMAN protein [79] are capable of sequence-specific DNA binding. In these single ZF proteins, basic residues flanking the domain provide additional contacts required for sequence specific DNA binding [79–82]. ZF-DNA complex formation generally slightly enlarges the DNA major groove [75].

The first accurate description of ZF-DNA interaction came in 1991, when the crystal structure of the complex formed by three ZFs from Zif268, a mouse transcription factor, and the GC-rich DNA target was solved [83]. The α-helix of each finger enters the DNA major groove and forms hydrogen bonds and hydrophobic interactions allowing the protein to wrap around the oligonucleotide in one turn. Hydrogen bonds are made between those α-helix residues at

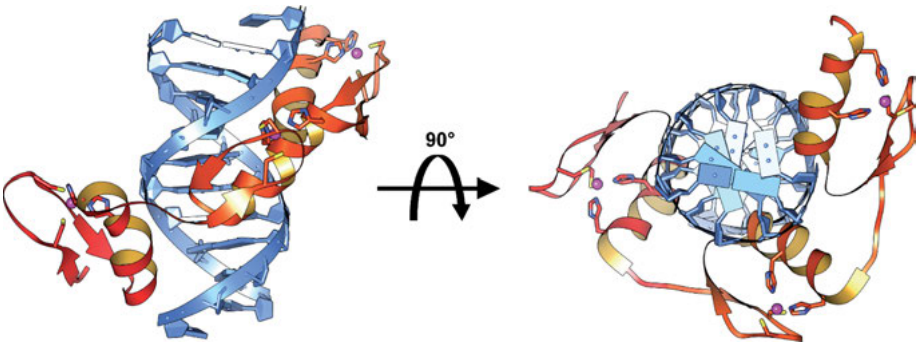


Figure 2. The three ZFs of Zif268 bound in the major groove of DNA (PDB code: 1AAY).

positions 1, 2, and 6 with three bases of the primary DNA strand while the residue at position 3 binds to a fourth base in the complementary strand [83]. The interaction is further stabilized by contacts made by lysine residues from the linker region (between the fingers) with the phosphate backbone. This mechanism appears to be common for many eukaryotic proteins [84–86]. A different mechanism has been shown to occur by Pedone et al. [77, 79, 87] in 1996: the single Cys_2His_2 ZF from the Drosophila protein GAGA binds DNA in the GA(CT)-rich region of the promoter with a good affinity. The protein exploits two N-terminal regions containing arginine and lysine residues (BR1 and BR2) to form a stable complex by making contacts with the minor groove; BR2 residues generate an α-helix. Without these additional interactions GAGA cannot generate adequate contacts to make a stable complex.

5. ZINC FINGERS IN PROKARYOTES

In the prokaryotic ZF domain, the eukaryotic ββα topology is enlarged with a third β-strand and a second α-helix forming a βββαα globular domain that is held together by the Zn(II) ion and by significant hydrophobic cores.

The first classical ZF-containing protein in bacteria was identified in 1998 by Chou et al. the *Agrobacterium tumefaciens* transcriptional regulator named Ros (Rough outer surface) [6]. This protein is involved in the horizontal transfer of genes from *A. tumefaciens* to a host plant infected by it [88]. The sequence [88–90] Cys-X_2-Cys-X_9-His-X_3-His-His occurs at residues 79–97 and shows two interesting peculiarities, namely three histidine residues and a nine-amino-acid spacer (eukaryota have 12-amino-acid spacers) that separates the first cysteine from the first histidine [91]. The deletion mutant of the first 55 amino acids (named Ros87 and numbered indicating Ala57 as residue 1) was proved to be the minimal region needed for DNA binding [92]. The Ros87 solution structure was solved by NMR evidencing the structural singularity of the prokaryotic Cys_2His_2 ZF: a globular domain of 58 residues (9–66 region) with three β-strands and two α-helices, with a 15 residues hydrophobic core and in which two cysteines and two histidines (the first and the third) coordinate a Zn(II) ion in a tetrahedral fashion. His41 is not involved in metal ion coordination but shields the structural ion from the bulk water [89]. Prokaryotic and eukaryotic Cys_2His_2 ZFs are compared in Figure 3. The βββα region of both domains are very similar and contain the Zn(II) coordination sphere with the two cysteines located on the β-hairpin and the two histidines located at the middle and at the C-terminus of the α-helix [89] which in Ros87 is one turn shorter. In Ros87 two hydrogen bonds anchor a second α-helix to the β-hairpin and bend over the βββα region with an axis nearly orthogonal to the first α-helix.

This arrangement allows the formation of the hydrophobic core composed by residues contained within each of the secondary structure elements; the core further stabilizes the globular fold [89]. The disordered and flexible N-terminus and C-terminus (residues 1–8 and 67–87, respectively) do not interact with the βββαα domain.

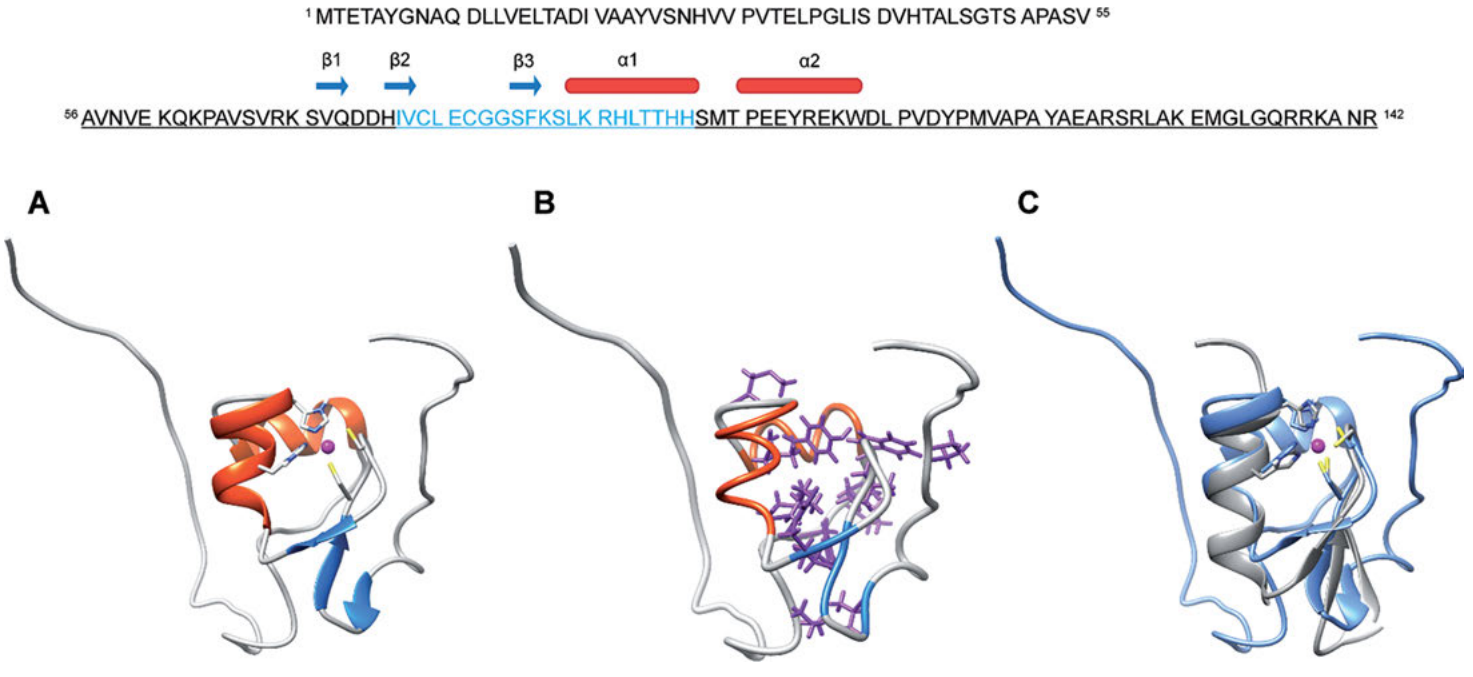


Figure 3. The prokaryotic zinc finger. **Upper part**: Ros sequence; the amino acids deleted in Ros87 are in the first line. The second line represents the Ros87 sequence showing in cyan the putative Cys_2His_2 ZF region and with the secondary structure elements drown on the sequence. **Lower panel**: The solution structure of Ros87 (PDB code: 2JSP). (**A**) The zinc ion is colored in magenta and the four coordinating side chains are shown. (**B**) The hydrophobic core of Ros87; the sidechains of the 15 residues are shown in violet. (**C**) Superposition of Ros87 (in light blue) with the first ZF domain of the eukaryotic Tramtrack protein (PDB code 2DRP, in grey), obtained by aligning the four zinc coordinating residues.

5.1. Ros/MucR Family

The Ros protein is a transcription repressor found in *A. tumefaciens.* It binds the *ros* box, a 40 bp sequence in the promoter of virC and virD operons and in the promoter of the *ros* gene itself. Ros belongs to the Ros/MucR family of proteins whose members have been recognized in different bacteria, mostly α-proteobacteria [93–96], symbionts, and pathogens of mammals and plants [97]. The *ros*/*mucR* gene is very conserved in proteobacteria and is central for host-bacterium interactions [98]. In the plant symbiont *Sinorhizobium meliloti* MucR regulates motility and the production of exopolysaccharides [99]. A MucR orthologue has been identified in *Brucella* [98, 100]. *Brucella* [97] is the animal pathogen that causes a widespread zoonosis known as brucellosis that affects mammals, including livestock and humans [101]. MucR has been demonstrated to be indispensable for the virulence of *Brucella* [97]: the *B. melitensis muc*R mutant strain appears to be a vaccine candidate against brucellosis [101] and the *B. abortus* CC092 strain lacking the *mucR* gene turned out to be attenuated in a mouse model [98]. MucR regulates its own transcription, the transcription of a great amount of genes involved in polysaccharide biosynthesis, genome plasticity, transcription regulation, cell envelope integrity and iron homeostasis [98]. Other homologues have been found in *Mesorhizobium loti* and in the archaea *Halobacterium sp.* NRC-1 and *Haloferax mediterranei* [2, 90], and in different members of the *Rhizobiaceae.* Generally, the characterized orthologues are transcriptional repressors that control host–bacterial interaction [102–105], various cell envelope modifications, quorum sensing, and motility [98, 102, 106–112].

Lately, a role for MucR from *Caulobacter crescentus* in the regulation of the bacterial cell cycle has been proposed [94]. Numerous *muc*R orthologues from α-proteobacteria have been proved to be interchangeable and can functionally substitute MucR in *C. crescentus* or *Sinorhizobium meliloti* [100].

In this big class of bacteria, the role of MucR proteins in controlling cell cycle regulation appears conserved in the α-proteobacteria thus suggesting that the basic architecture of the regulatory network has been preserved among them [95]. Indeed, it has been suggested that MucR proteins together with other virulence regulators primarily control G1-phase transcription, rendering expression of target genes related to virulence and/or symbiosis in tune with the cell cycle [94].

5.2. DNA Binding and Functional Domain

The DNA recognition mechanism of the prokaryotic ZF appears rather peculiar and different if compared to what is observed in the eukaryotic domain [93, 113]. NMR, molecular dynamics and biochemical studies performed on Ros87, the minimal DNA binding domain of Ros from *A. tumefaciens*, show that residues belonging to the first α-helix contact the major groove bases and backbone phosphate. The C-terminal tail adopts a helical conformation upon DNA recognition, is enveloped around the DNA double-helix and makes essential contacts with

bases positioned both at the 5′ and at the 3′ ends. Ros87 is the deletion mutant of the first 55 residues of the protein Ros (Figure 3).

The N-terminal region of Ros/MucR family members has been recently demonstrated to be responsible for oligomerization of these proteins: the full-length proteins Ml1 and Ml2 from *Mesorizhobium loti* and MucR from *B. abortus* form high-order oligomers through the N-terminal domain; MucR1 and MucR2 from *C. crescentus* form heterodimers [114]. Oligomerization of MucR in *B. abortus* is essential to get the proper expression regulation of the target genes involved in the infectious process [115].

The preferred DNA target site recognized by Ml1, Ml2, and by *B. abortus* MucR is an AT-rich sequence that contains a T-A step. These proteins make essential contacts in the minor groove of DNA [114–116]. As AT-rich sequences containing a T-A step are present in the vir Box recognized by Ros [2] as well, this suggests that also the protein Ros binds DNA using a similar mechanism.

This DNA recognition mode together with the mechanism used by this protein to regulate gene expression suggests that the studied members of the Ros/MucR protein family behave as H-NS-like protein. The histone-like nucleoid-structuring class of proteins (H-NS) is a family of small proteins of about 90 amino acids that bind DNA and behave as histone-like proteins; also H-NS proteins form high-order oligomers through their N-terminal region [114–116].

Just as H-NS proteins, members of the Ros/MucR family can recognize AT-rich sequences via the C-terminal DNA-binding domain and extend their presence along the nucleoid via oligomerization through N-terminal region.

5.3. Coordination Sphere and Loss of the Structural Metal Ion

Amino acids composing the coordination sphere of the structural ion are poorly conserved in most of the proteins of the Ros/MucR family albeit their sequence identity can be very high [93]. As an example, Figure 4 reports the alignment of the Cys_2His_2 ZF domain from Ros with the corresponding regions of 10 proteins from *M. loti*; the sequence identity is above 50 % and becomes higher if only the metal coordinating region (residues 24–42) is evaluated.

Three of the four metal coordinating residues of Ros are poorly conserved in the *M. loti* proteins [91]. Only the histidine in the third coordinating position is conserved; in three proteins the first cysteine is substituted by a serine, the possible fourth zinc coordinating position bears different combinations of residues [93] while all the shown proteins have an aspartic acid in the possible second coordinating position [93]. Aspartic acid, serine, tyrosine, phenylalanine, leucine, and glycine have never been reported as Zn(II)-coordinating residues in ZF motifs and for the phenylalanine, leucine, and glycine a Zn(II)-binding capability is hard to imagine. The high sequence identity of the shown *M. loti* proteins suggests a similar three-dimensional structure in which the metal ion has a completely new Zn(II) coordination modality or a secondary role in protein folding.

Metal-coordinating proteins frequently have distant relatives lacking the corresponding metal binding site [7, 117, 118]; however, only in a few cases proteins

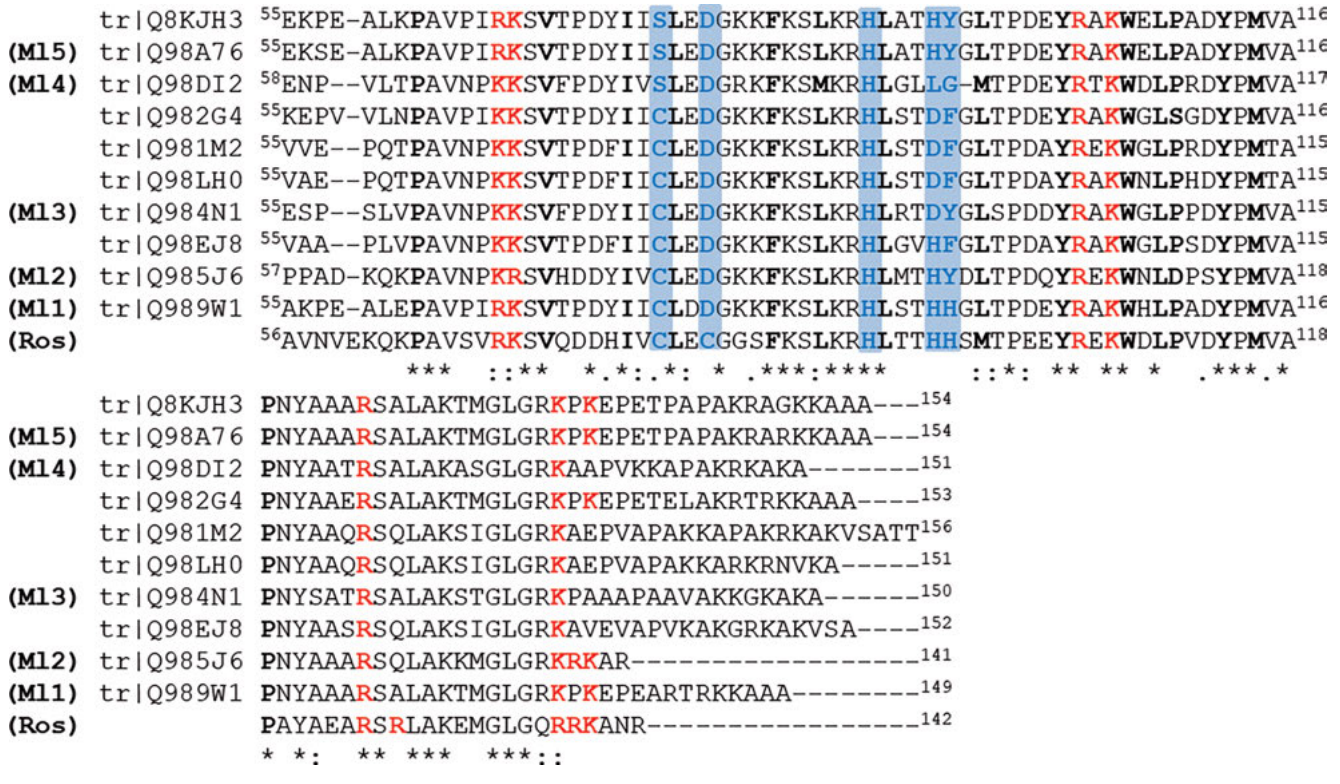


Figure 4. Sequence alignment with the Ros87 sequence of 10 proteins of *Mesorhizobium loti*. The accession numbers of the proteins are reported. The positions corresponding to the coordination residues in Ros are shown in light blue (His96 and His97 are able to function alternatively as the fourth coordinating residue and are both highlighted). Amino acids constituting the Ros hydrophobic core are in bold. Basic residues important for Ros DNA binding are in red. The asterisks indicate conserved residues; a colon indicates a conservative substitution; a period indicates a semiconservative substitution.

with more than 35 % sequence identity differ in whether they bind a structural metal [7, 117, 118] and in rare cases, the proteins are able to properly fold in the absence of their structural metal ion [119–121]. N-terminal deletion mutants corresponding to Ros87 of Ml1, Ml2, Ml3, Ml4, and Ml5 have been shown to be able to specifically bind the Ros87 DNA target, suggesting a similar fold of the DNA-binding domains and a common DNA-binding modality. Ml1, Ml2, and Ml3 bind Zn(II) while Ml4 and Ml5 do not bind either the zinc ion or any other metal. NMR, bioinformatics, and biochemical data demonstrated that these latter two proteins, though lacking the Zn(II) ion, assume a 3D structure similar to that of Ros87 and that the residues replacing the Zn(II) ligands form a network of hydrogen bonds and hydrophobic interactions that surrogates the Zn(II) coordinating role in stabilizing the correct and functional fold [122].

Furthermore, Ser → Cys mutation transforms the metal-lacking protein Ml5 into a metal coordinating protein while Cys → Ser mutation renders Ml2 a metal-lacking protein [93]. The structure of the prokaryotic ZF domain seems extremely adaptable as it preserves the DNA-binding function either with the Zn(II) ion bound via a really mutable set of coordinating residues or lacking Zn(II) [123–125]. It is remarkable that an eukaryotic stable ββα ZF domain without the structural Zn(II) has been obtained [126] on the basis of a natural occurring ZF by replacing the Zn(II) ligands by two phenylalanines that provide a further hydrophobic contribute to a well-packed core.

In eukaryota, the effect of Zn(II) binding directs the associated domain folding in the correct functional ZF influencing the magnitude of the experimentally determined binding affinities [49, 127, 128]. In the case of the prokaryotic ZF Ros87, the protein functionally binds Cd(II) with an affinity that is comparable to Zn(II) [129]. When in the presence of Hg(II) and Pb(II) [53], the structuring effect of the hydrophobic core is not sufficient to allow the protein to fold, causing significantly higher dissociation constants. It is interesting to underline that the ionic radii of both Hg(II) and Pb(II) are significantly larger than that of Zn(II), and this may influence the folding impairing the hydrophobic interactions that stabilize the domain tertiary structure. On the other hand, the Ni(II) ion with a radius slightly smaller than that of Zn(II), allows the protein to fold in a structure which however is incapable to bind DNA. Overall, the prokaryotic ZF seems capable to structurally tolerate metal ions with varied ionic radii; the Cd(II) seems to be the maximum limit in terms of possible dimensions to obtain a functional domain.

5.4. Folding Mechanism

The folding pathway of a the prokaryotic ZF has been explored by using a combination of calorimetry, CD, and solution NMR spectroscopy. Two different mechanism of folding have been described for this family of proteins depending on whether a Zn(II)-binding or a Zn(II)-free ZF is considered. The metal-binding Ros87 and the metal-lacking Ml4 have been studied as examples of the two cases [130]. On the one hand, the unfolding pathway of Ml4 can be described

as fully cooperative: the protein experiences a classical two-state cooperative unfolding and a quite low melting temperature of 306 K [130]. On the other hand, the presence of the structural Zn(II) ion in Ros87 appears to modify the folding mechanism of this protein that shows a two-transition unfolding process. Ros87 shows evidences of a bipartite folding with a well-structured intermediate. Such an intermediate is mainly constituted by the metal-binding site in which the two cysteines bind the Zn(II) ion and stabilize the β-turn while the rest of the protein is mainly unstructured and dynamic. The histidine in the fourth coordinating position (in the eukaryotic ZF it is the residue in the third coordinating position) contributes to the formation of a proto-hydrophobic core. The intermediate converts in the native state through a barrierless downhill mechanism [130, 131]. Finally, the second histidine in the third coordinating position binds the metal so that the helix undergoes a complete fold.

5.5. Evolution

How are the prokaryotic and eukaryotic ZF domains evolutionarily related?

Evolution plays with proteins introducing progressive variations in terms of sequences and structures [132]: single amino acid substitutions, insertions, and deletions are the most common tools to create new protein folds [133]. The distinctive structural features of the eukaryotic ZF domain have allowed evolution to create a plethora of transcription factors by simply combining and varying ZFs finely regulating affinity for their specific targets by means of point substitutions [134]. Initially, it was suggested that eubacteria acquired ZFs by horizontal gene transfer (HGT) from eukaryotes [135]. As a matter of fact, the first identified bacterial protein containing a ZF domain (Ros from *A. tumefaciens*) is itself involved in HGT [135]. Later, Ros homologues have been found in bacteria belonging to the α-subdivision of proteobacteria from which mitochondria originated leading to the hypothesis that eukaryotes had classical ZFs because of mitochondrial endosymbiosis [96].

In the frame of a possible HGT from prokaryotes to eukaryotes, as postulated by Moreira and Rodriguez-Valera [96], the evolution of eukaryotic classical ZFs from the prokaryotic Ros-like proteins might have occurred simply by obtaining two sets of properly spaced metal-chelating ligands [134]. In fact, known the diverse lengths of the amino acid loop between the second cysteine and the first histidine, an increase to 12 (from 9) in the number of residues in this loop of the Ros protein appears enough to produce a structural transition from a Ros-like to a eukaryotic classical ZF-like structure [134] thus leading to the autonomous folding of the Cys_2His_2 portion of the protein. A simple three residues shift in the N-terminal direction of the protein frame or an insertion of three amino acids could lead to such conversion.

This latter hypothesis as a single mutational event appears evolutionary and structurally more probable [134]. Starting from a prokaryotic Ros-like ZF, a structural evolution to eukaryotic classical ZFs may have happened in a few steps: a three amino acids insertion, the loss of the adjacent parts of the globular

domain followed by tandem duplications and diversification of the domain thus obtained [134]. In this way, eukaryotic ZFs might have evolved multiple times from Ros homologues.

6. GENERAL CONCLUSIONS

ZFs have crucial roles in many cellular functions. The modularity and specificity of ZF proteins has generated in time the idea of creating synthetic ZF proteins by exploiting the fact that each domain can recognize a specific DNA triplet on the basis of the α-helix residues. Such mechanism was believed capable to allow the creation of a straightforward code to recognize extremely specific DNA sequences. The engineered ZF protein could in principle bind the DNA target sequence while a different domain of the same protein would introduce the desired biological function [136].

Frequent applications include ZF transcription factors and ZF nucleases, but other applications have also been described. Engineered ZF proteins usually bear 3 to 6 domains thus resulting in different lengths of target binding sites. Proteins with more domains are more attractive as their binding target site is longer and thus the chances of being unique are higher [137].

Moreover, the role of ZFs in cancer development, progression, and metastasis is emerging [74]. The specificity of function and expression of some ZFs for some tumors suggests the use of this class of proteins as prognostic factors.

In conclusion, because of their diverse nature, natural ZFs show considerably different properties if compared to the engineered domains, thus rendering the meticulous characterization of natural ZFs highly needed [21].

ACKNOWLEDGMENTS

This work was funded by the Italian MiUR (PRIN 20157WZM8A to G.M.)

ABBREVIATIONS AND DEFINITIONS

A	adenine
CD	circular dichroism
Cys	cysteine
C	cytosine
G	guanine
GAGA	drosophila protein
HGT	horizontal gene transfer
His	histidine
HIV	human immunodeficiency virus
H-NS	histone-like structuring class of proteins

mRNA	messenger ribonucleic acid
NC	nucleocapsid
NMR	nuclear magnetic resonance
Ros	rough outer surface (transcription regulator)
Ser	serine
T	thymine
ZF	zinc finger

REFERENCES

1. A. L. Lehninger, D. L. Nelson, M. M. Cox, *Principles of Biochemistry*, 2nd edn., Worth Publishers Inc., New York, 1993.
2. G. Malgieri, M. Palmieri, L. Russo, R. Fattorusso, P. V. Pedone, C. Isernia, *FEBS J.* **2015**, *282*, 4480–4496.
3. A. Klug, *Quart. Rev. Biophys.* **2010**, *43*, 1–21.
4. R. Tupler, G. Perini, M. R. Green, *Nature* **2001**, *409*, 832–833.
5. J. H. Laity, B. M. Lee, P. E. Wright, *Curr. Opin. Struct. Biol.* **2001**, *11*, 39–46.
6. A. Y. Chou, J. Archdeacon, C. I. Kado, *Proc. Natl. Acad. Sci. USA* **1998**, *95*, 5293–5298.
7. S. S. Krishna, I. Majumdar, N. V. Grishin, *Nucleic Acids Res.* **2003**, *31*, 532–550.
8. R. S. Brown, C. Sander, P. Argos, *FEBS Lett.* **1985**, *186*, 271–274.
9. J. Miller, A. D. McLachlan, A. Klug, *EMBO J.* **1985**, *4*, 1609–1614.
10. C. Bai, P. P. Tolias, *Nucleic Acids Res.* **1998**, *26*, 1597–1604.
11. L. M. Green, J. M. Berg, *Proc. Natl. Acad. Sci. USA* **1989**, *86*, 4047–4051.
12. C. He, J. C. Hus, L. J. Sun, P. Zhou, D. P. Norman, V. Dötsch, H. Wei, J. D. Gross, W. S. Lane, G. Wagner, G. L. Verdine, *Mol. Cell* **2005**, *20*, 117–129.
13. E. Layat, A. V. Probst, S. Tourmente, *Biochim. Biophys. Acta* **2013**, *1829*, 274–282.
14. A. T. Maynard, D. G. Covell, *J. Am. Chem. Soc.* **2001**, *123*, 1047–1058.
15. M. Nanami, T. Ookawara, Y. Otaki, K. Ito, R. Moriguchi, K. Miyagawa, Y. Hasuike, M. Izumi, H. Eguchi, K. Suzuki, T. Nakanishi, *Arterioscler. Thromb. Vasc. Biol.* **2005**, *25*, 2495–2501.
16. S. M. Quintal, Q. A. dePaula, N. P. Farrell, *Metallomics* **2011**, *3*, 121–139.
17. K. S. Eom, J. S. Cheong, S. J. Lee, *J. Microbiol. Biotechnol.* **2016**, *26*, 2019–2029.
18. S. A. Wolfe, L. Nekludova, C. O. Pabo, *Annu. Rev. Biophys. Biomol. Struct.* **2000**, *29*, 183–212.
19. L. Fairall, J. W. Schwabe, L. Chapman, J. T. Finch, D. Rhodes, *Nature* **1993**, *366*, 483–487.
20. R. N. Dutnall, D. Neuhaus, D. Rhodes, *Structure* **1996**, *4*, 599–611.
21. K. Kluska, J. Adamczyk, A. Krężel, *Coord. Chem. Rev.* **2018**, *367*, 18–64.
22. M. McCall, T. Brown, W. N. Hunter, O. Kennard, *Nature* **1986**, *322*, 661–664.
23. M. F. Summers, L. E. Henderson, M. R. Chance, J. W. Bess, T. L. South, P. R. Blake, I. Sagi, G. Perez-Alvarado, R. C. Sowder, D. R. Hare, *Protein Sci.* **1992**, *1*, 563–574.
24. N. V. Grishin, *Nucleic Acids Res.* **2001**, *29*, 1703–1714.
25. H. T. Chen, P. Legault, J. Glushka, J. G. Omichinski, R. A. Scott, *Protein Sci.* **2000**, *9*, 1743–1752.
26. X. Qian, S. N. Gozani, H. Yoon, C. J. Jeon, K. Agarwal, M. A. Weiss, *Biochemistry* **1993**, *32*, 9944–9959.
27. N. V. Grishin, *J. Mol. Biol.* **2000**, *299*, 1165–1177.
28. A. Krężel, W. Maret, *J. Biol. Inorg. Chem.* **2006**, *11*, 1049–1062

29. Y. Li, W. Maret, *Exp. Cell Res.* **2009**, *315*, 2463–2470.
30. A. Miłoch, A. Krężel, *Metallomics* **2014**, *6*, 2015–2024.
31. M. Sikorska, A. Krężel, J. Otlewski, *J. Inorg. Biochem.* **2012**, *115*, 28–35.
32. J. W. Nelson, T. E. Creighton, *Biochemistry* **1994**, *33*, 5974–5983.
33. C. A. Blasie, J. M. Berg, *Biochemistry* **2004**, *43*, 10600–10604.
34. A. R. Reddi, B. R. Gibney, *Biochemistry* **2007**, *46*, 3745–3758.
35. A. M. Rich, E. Bombarda, A. D. Schenk, P. E. Lee, E. H. Cox, A. M. Spuches, L. D. Hudson, B. Kieffer, D. E. Wilcox, *J. Am. Chem. Soc.* **2012**, *134*, 10405–10418.
36. G. Bulaj, T. Kortemme, D. P. Goldenberg, *Biochemistry* **1998**, *37*, 8965–8972.
37. H. Nakamura, *Quart. Rev. Biophys.* **1996**, *29*, 1–90.
38. M. F. Perutz, *Science* **1978**, *201*, 1187–1191.
39. A. Warshel, *Acc. Chem. Res.* **1981**, *14*, 284–290.
40. A. Fung, P. Li, R. Godoy-Ruiz, J. M. Sanchez-Ruiz, V. Muñoz, *J. Am. Chem. Soc.* **2008**, *130*, 7489–7495.
41. S. W. Englander, *Annu. Rev. Biophys. Biomol. Struct.* **2000**, *29*, 213–238.
42. W. Li, J. Zhang, J. Wang, W. Wang, *J. Am. Chem. Soc.* **2008**, *130*, 892–900.
43. W. Li, J. Wang, J. Zhang, W. Wang, *Curr. Opin. Struct. Biol.* **2015**, *30*, 25–31.
44. V. Lebrun, A. Tron, C. Lebrun, J. M. Latour, N. D. McClenaghan, O. Sénèque, *Chemistry* **2015**, *21*, 14002–14010.
45. A. Nomura, Y. Sugiura, *Inorg. Chem.* **2002**, *41*, 3693–3698.
46. Y. Shi, R. D. Beger, J. M. Berg, *Biophys. J.* **1993**, *64*, 749–753.
47. A. Hartwig, M. Asmuss, I. Ehleben, U. Herzer, D. Kostelac, A. Pelzer, T. Schwerdtle, A. Bürkle, *Environ. Health Perspect.* **2002**, *110, Suppl 5*, 797–799.
48. A. Witkiewicz-Kucharczyk, W. Bal, *Toxicol. Lett.* **2006**, *162*, 29–42.
49. V. Sivo, G. D'Abrosca, L. Russo, R. Iacovino, P. V. Pedone, R. Fattorusso, C. Isernia, G. Malgieri, *Bioinorg. Chem. Appl.* **2017**, *2017*, 1527247; https://doi.org/10.1155/2017/1527247.
50. R. G. Pearson, *Coord. Chem. Rev.* **1990**, *100*, 403–425.
51. W. Bal, T. Schwerdtle, A. Hartwig, *Chem. Res. Toxicol.* **2003**, *16*, 242–248.
52. M. C. Posewitz, D. E. Wilcox, *Chem. Res. Toxicol.* **1995**, *8*, 1020–1028.
53. V. Sivo, G. D'Abrosca, I. Baglivo, R. Iacovino, P. V. Pedone, R. Fattorusso, L. Russo, G. Malgieri, C. Isernia, *Inorg. Chem.* **2019**, *58*, 1067–1080.
54. M. Nagaoka, J. Kuwahara, Y. Sugiura, *Biochem. Biophys. Res. Commun.* **1993**, *194*, 1515–1520.
55. M. R. Basha, W. Wei, M. Brydie, M. Razmiafshari, N. H. Zawia, *Int. J. Dev. Neurosci.* **2003**, *21*, 1–12.
56. M. Razmiafshari, J. Kao, A. d'Avignon, N. H. Zawia, *Toxicol. Appl. Pharmacol.* **2001**, *172*, 1–10.
57. J. M. Ordemann, R. N. Austin, *Metallomics* **2016**, *8*, 579–588.
58. M. Razmiafshari, N. H. Zawia, *Toxicol. Appl. Pharmacol.* **2000**, *166*, 1–12.
59. J. S. Hanas, J. S. Rodgers, J. A. Bantle, Y. G. Cheng, *Mol. Pharmacol.* **1999**, *56*, 982–988.
60. J. S. Magyar, T. C. Weng, C. M. Stern, D. F. Dye, B. W. Rous, J. C. Payne, B. M. Bridgewater, A. Mijovilovich, G. Parkin, J. M. Zaleski, J. E. Penner-Hahn, H. A. Godwin, *J. Am. Chem. Soc.* **2005**, *127*, 9495–9505.
61. A. Hartwig, *Biometals* **2010**, *23*, 951–960.
62. D. Beyersmann, A. Hartwig, *Arch. Toxicol.* **2008**, *82*, 493–512.
63. M. Waisberg, P. Joseph, B. Hale, D. Beyersmann, *Toxicology* **2003**, *192*, 95–117.
64. T. S. Nawrot, J. A. Staessen, H. A. Roels, E. Munters, A. Cuypers, T. Richart, A. Ruttens, K. Smeets, H. Clijsters, J. Vangronsveld, *Biometals* **2010**, *23*, 769–782.

65. G. Malgieri, L. Zaccaro, M. Leone, E. Bucci, S. Esposito, I. Baglivo, A. Del Gatto, L. Russo, R. Scandurra, P. V. Pedone, R. Fattorusso, C. Isernia, *Biopolymers* **2011**, *95*, 801–810.
66. P. F. Predki, B. Sarkar, *J. Biol. Chem.* **1992**, *267*, 5842–5846.
67. A. Hartwig, *Antioxid. Redox Signal.* **2001**, *3*, 625–634
68. D. Krepkiy, F. H. Försterling, D. H. Petering, *Chem. Res. Toxicol.* **2004**, *17*, 863–870.
69. D. H. Petering, M. Huang, S. Moteki, C. F. Shaw, *Mar. Environ. Res.* **2000**, *50*, 89–92.
70. M. Huang, D. Krepkiy, W. Hu, D. H. Petering, *J. Inorg. Biochem.* **2004**, *98*, 775–785.
71. R. Kothinti, A. Blodgett, N. M. Tabatabai, D. H. Petering, *Chem. Res. Toxicol.* **2010**, *23*, 405–412.
72. G. Roesijadi, R. Bogumil, M. Vasák, J. H. Kägi, *J. Biol. Chem.* **1998**, *273*, 17425–17432.
73. Y. Peng, K. J. Clark, J. M. Campbell, M. R. Panetta, Y. Guo, S. C. Ekker, *Development* **2014**, *141*, 4042–4054.
74. M. Cassandri, A. Smirnov, F. Novelli, C. Pitolli, M. Agostini, M. Malewicz, G. Melino, G. Raschellà, *Cell Death Discov.* **2017**, *3*, 17071.
75. J. M. Berg, H. A. Godwin, *Annu. Rev. Biophys. Biomol. Struct.* **1997**, *26*, 357–371.
76. D. Xiong, Y. Wang, C. Deng, R. Hu, C. Tian, *Gene* **2015**, *562*, 169–179.
77. J. G. Omichinski, P. V. Pedone, G. Felsenfeld, A. M. Gronenborn, G. M. Clore, *Nature Struct. Biol.* **1997**, *4*, 122–132.
78. A. Klug, J. W. Schwabe, *FASEB J.* **1995**, *9*, 597–604.
79. P. V. Pedone, R. Ghirlando, G. M. Clore, A. M. Gronenborn, G. Felsenfeld, J. G. Omichinski, *Proc. Natl. Acad. Sci. USA* **1996**, *93*, 2822–2826.
80. N. Dathan, L. Zaccaro, S. Esposito, C. Isernia, J. G. Omichinski, A. Riccio, C. Pedone, B. Di Blasio, R. Fattorusso, P. V. Pedone, *Nucleic Acids Res.* **2002**, *30*, 4945–4951.
81. R. Dinkins, C. Pflipsen, A. Thompson, G. B. Collins, *Plant Cell. Physiol.* **2002**, *43*, 743–750.
82. B. W. Tague, H. M. Goodman, *Plant. Mol. Biol.* **1995**, *28*, 267–279.
83. N. P. Pavletich, C. O. Pabo, *Science* **1991**, *252*, 809–817.
84. D. S. Wuttke, M. P. Foster, D. A. Case, J. M. Gottesfeld, P. E. Wright, *J. Mol. Biol.* **1997**, *273*, 183–206.
85. R. T. Nolte, R. M. Conlin, S. C. Harrison, R. S. Brown, *Proc. Natl. Acad. Sci. USA* **1998**, *95*, 2938–2943.
86. J. M. Berg, *Proc. Natl. Acad. Sci. USA* **1992**, *89*, 11109–11110.
87. P. V. Pedone, J. G. Omichinski, P. Nony, C. Trainor, A. M. Gronenborn, G. M. Clore, G. Felsenfeld, *EMBO J.* **1997**, *16*, 2874–2882.
88. C. I. Kado, *Antonie Van Leeuwenhoek* **1998**, *73*, 117–126.
89. G. Malgieri, L. Russo, S. Esposito, I. Baglivo, L. Zaccaro, E. M. Pedone, B. Di Blasio, C. Isernia, P. V. Pedone, R. Fattorusso, *Proc. Natl. Acad. Sci. USA* **2007**, *104*, 17341–17346.
90. C. I. Kado, *Plasmid* **2002**, *48*, 179–185.
91. M. B. Cooley, M. R. D'Souza, C. I. Kado, *J. Bacteriol.* **1991**, *173*, 2608–2616.
92. S. Esposito, I. Baglivo, G. Malgieri, L. Russo, L. Zaccaro, L. D. D'Andrea, M. Mammucari, B. Di Blasio, C. Isernia, R. Fattorusso, P. V. Pedone, *Biochemistry* **2006**, *45*, 10394–10405.
93. I. Baglivo, L. Russo, S. Esposito, G. Malgieri, M. Renda, A. Salluzzo, B. Di Blasio, C. Isernia, R. Fattorusso, P. V. Pedone, *Proc. Natl. Acad. Sci. USA* **2009**, *106*, 6933–6938.

94. C. Fumeaux, S. K. Radhakrishnan, S. Ardissone, L. Théraulaz, A. Frandi, D. Martins, J. Nesper, S. Abel, U. Jenal, P. H. Viollier, *Nature Commun.* **2014**, *5*, 4081; doi: 10.1038/ncomms5081.
95. G. Panis, S. R. Murray, P. H. Viollier, *FEMS Microbiol. Rev.* **2015**, *39*, 120–133.
96. D. Moreira, F. Rodríguez-Valera, *Trends Microbiol.* **2000**, *8*, 448–450.
97. E. Moreno, E. Stackebrandt, M. Dorsch, J. Wolters, M. Busch, H. Mayer, *J. Bacteriol.* **1990**, *172*, 3569–3576.
98. C. C. Caswell, A. E. Elhassanny, E. E. Planchin, C. M. Roux, J. N. Weeks-Gorospe, T. A. Ficht, P. M. Dunman, R. M. Roop, *Infect. Immun.* **2013**, *81*, 1040–1051.
99. C. Bahlawane, M. McIntosh, E. Krol, A. Becker, *Mol. Plant. Microbe Interact.* **2008**, *21*, 1498–1509.
100. A. Mirabella, M. Terwagne, M. S. Zygmunt, A. Cloeckaert, X. De Bolle, J. J. Letesson, *J. Bacteriol.* **2013**, *195*, 453–465.
101. G. Pappas, P. Papadimitriou, N. Akritidis, L. Christou, E. V. Tsianos, *Lancet Infect. Dis.* **2006**, *6*, 91–99.
102. M. Janczarek, A. Skorupska, *Mol. Plant. Microbe Interact.* **2007**, *20*, 867–881.
103. M. A. Bittinger, J. Handelsman, *J. Bacteriol.* **2000**, *182*, 1706–1713.
104. M. Janczarek, J. Kutkowska, T. Piersiak, A. Skorupska, *BMC Microbiol.* **2010**, *10*, 284; https://doi.org/10.1186/1471-2180-10-284.
105. K. Rachwał, E. Matczyńska, M. Janczarek, *BMC Genomics* **2015**, *16*, 1111; https://doi.org/10.1186/s12864-015-2332-4.
106. P. A. Bertram-Drogatz, I. Quester, A. Becker, A. Pühler, *Mol. Gen. Genet.* **1998**, *257*, 433–441.
107. S. Rüberg, A. Pühler, A. Becker, *Microbiology* **1999**, *145*, 603–611.
108. M. McIntosh, E. Krol, A. Becker, *J. Bacteriol.* **2008**, *190*, 5308–5317.
109. A. Becker, S. Rüberg, H. Küster, A. A. Roxlau, M. Keller, T. Ivashina, H. P. Cheng, G. C. Walker, A. Pühler, *J. Bacteriol.* **1997**, *179*, 1375–1384.
110. M. Martín, J. Lloret, M. Sánchez-Contreras, I. Bonilla, R. Rivilla, *Mol. Plant. Microbe. Interact.* **2000**, *13*, 129–135.
111. C. Bahlawane, B. Baumgarth, J. Serrania, S. Rüberg, A. Becker, *J. Bacteriol.* **2008**, *190*, 3456–3466.
112. K. Mueller, J. E. González, *J. Bacteriol.* **2011**, *193*, 485–496.
113. L. Russo, M. Palmieri, J. V. Caso, G. D'Abrosca, D. Diana, G. Malgieri, I. Baglivo, C. Isernia, P. V. Pedone, R. Fattorusso, *Eur. J. Med. Chem.* **2015**, *91*, 100–108. Epub 2014 Sep 16; https://doi.org/10.1016/j.ejmech.2014.09.040.
114. I. Baglivo, L. Pirone, E. M. Pedone, J. E. Pitzer, L. Muscariello, M. M. Marino, G. Malgieri, A. Freschi, A. Chambery, R. M. Roop Ii, P. V. Pedone, *Sci. Rep.* **2017**, *7*, 15805.
115. L. Pirone, J. E. Pitzer, G. D'Abrosca, R. Fattorusso, G. Malgieri, E. M. Pedone, P. V. Pedone, R. M. Roop, I. Baglivo, *Sci. Rep.* **2018**, *8*, 17238.
116. I. Baglivo, L. Pirone, G. Malgieri, R. Fattorusso, R. M. Roop Ii, E. M. Pedone, P. V. Pedone, *FEBS Open Bio.* **2018**, *8*, 711–718.
117. W. Maret, *J. Trace Elem. Med. Biol.* **2005**, *19*, 7–12.
118. J. W. Torrance, M. W. Macarthur, J. M. Thornton, *Proteins* **2008**, *71*, 813–830.
119. Y. M. Lee, C. Lim, *J. Mol. Biol.* **2008**, *379*, 545–553.
120. E. H. Cox, G. L. McLendon, *Curr. Opin. Chem. Biol.* **2000**, *4*, 162–165.
121. D. S. Auld, *Biometals* **2001**, *14*, 271–313.
122. I. Baglivo, M. Palmieri, A. Rivellino, F. Netti, L. Russo, S. Esposito, R. Iacovino, B. Farina, C. Isernia, R. Fattorusso, P. V. Pedone, G. Malgieri, *Biochim. Biophys. Acta* **2014**, *1844*, 497–504.

123. M. Palmieri, L. Russo, G. Malgieri, S. Esposito, I. Baglivo, A. Rivellino, B. Farina, I. de Paola, L. Zaccaro, D. Milardi, C. Isernia, P. V. Pedone, R. Fattorusso, *J. Inorg. Biochem.* **2014**, *131*, 30–36.
124. G. D'Abrosca, L. Russo, M. Palmieri, I. Baglivo, F. Netti, I. de Paola, L. Zaccaro, B. Farina, R. Iacovino, P. V. Pedone, C. Isernia, R. Fattorusso, G. Malgieri, *J. Inorg. Biochem.* **2016**, *161*, 91–98.
125. A. Urbani, R. Bazzo, M. C. Nardi, D. O. Cicero, R. De Francesco, C. Steinkühler, G. Barbato, *J. Biol. Chem.* **1998**, *273*, 18760–18769.
126. B. I. Dahiyat, S. L. Mayo, *Science* **1997**, *278*, 82–87.
127. J. C. Payne, M. A. ter Horst, H. A. Godwin, *J. Am. Chem. Soc.* **1999**, *121*, 6850–6855.
128. B. A. Krizek, D. L. Merkle, J. M. Berg, *Inorg. Chem.* **1993**, *32*, 937–940.
129. G. Malgieri, M. Palmieri, S. Esposito, V. Maione, L. Russo, I. Baglivo, I. de Paola, D. Milardi, D. Diana, L. Zaccaro, P. V. Pedone, R. Fattorusso, C. Isernia, *Metallomics* **2014**, *6*, 96–104.
130. M. Palmieri, G. Malgieri, L. Russo, I. Baglivo, S. Esposito, F. Netti, A. Del Gatto, I. de Paola, L. Zaccaro, P. V. Pedone, C. Isernia, D. Milardi, R. Fattorusso, *J. Am. Chem. Soc.* **2013**, *135*, 5220–5228.
131. G. Malgieri, G. D'Abrosca, L. Pirone, A. Toto, M. Palmieri, L. Russo, M. F. M. Sciacca, R. Tatè, V. Sivo, I. Baglivo, R. Majewska, M. Coletta, P. V. Pedone, C. Isernia, M. De Stefano, S. Gianni, E. M. Pedone, D. Milardi, R. Fattorusso, *Chem. Sci.* **2018**, *9*, 3290–3298.
132. E. D. Scheeff, P. E. Bourne, *PLoS Comput. Biol.* **2005**, *1*, e49.
133. N. V. Grishin, *J. Struct. Biol.* **2001**, *134*, 167–185.
134. F. Netti, G. Malgieri, S. Esposito, M. Palmieri, I. Baglivo, C. Isernia, J. G. Omichinski, P. V. Pedone, N. Lartillot, R. Fattorusso, *Mol. Biol. Evol.* **2013**, *30*, 1504–1513.
135. N. Bouhouche, M. Syvanen, C. I. Kado, *Trends Microbiol.* **2000**, *8*, 77–81.
136. F. D. Urnov, E. J. Rebar, M. C. Holmes, H. S. Zhang, P. D. Gregory, *Nature Rev. Genet.* **2010**, *11*, 636–646.
137. Q. Liu, D. J. Segal, J. B. Ghiara, C. F. Barbas, *Proc. Natl. Acad. Sci. USA* **1997**, *94*, 5525–5530.

Subject Index

A

Metal Ions in Life Sciences, Volume 20 Guest Editors: Martha E. Sosa Torres and Peter M. H. Kroneck
Series Editors: Astrid Sigel, Eva Freisinger, and Roland K. O. Sigel

https://doi.org/10.1515/9783110589757-013

B

C

D

E

F

G

H

I

K

L

M

N

O

P

Q

R

S

T

U

V

W

X

Y

Z