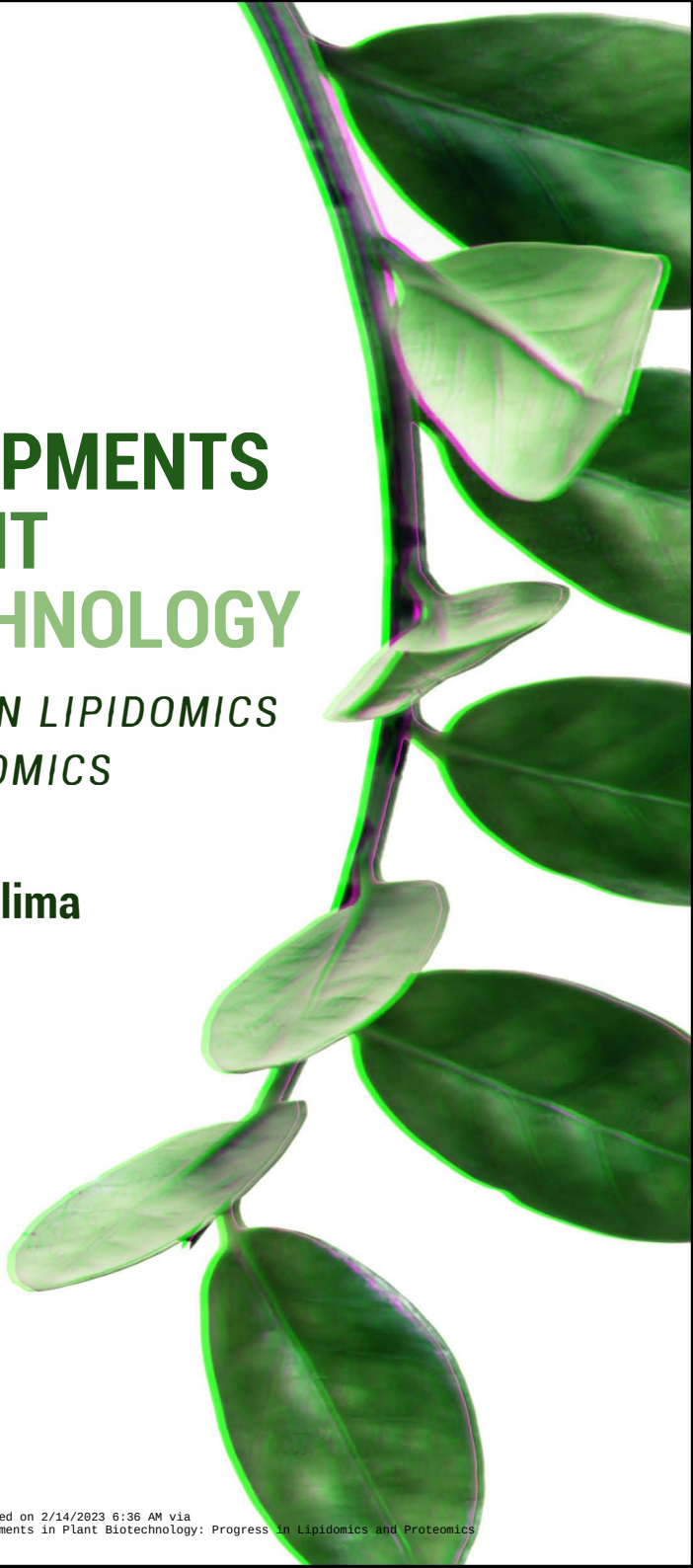
RECENT DEVELOPMENTS IN PLANT BIOTECHNOLOGY

PROGRESS IN LIPIDOMICS AND PROTEOMICS

Nihed Ben Halima



Recent Developments in Plant Biotechnology

Recent Developments in Plant Biotechnology:

Progress in Lipidomics and Proteomics

By

Nihed Ben Halima

Cambridge Scholars Publishing

Recent Developments in Plant Biotechnology:
Progress in Lipidomics and Proteomics

By Nihed Ben Halima

This book first published 2019

Cambridge Scholars Publishing

Lady Stephenson Library, Newcastle upon Tyne, NE6 2PA, UK

British Library Cataloguing in Publication Data
A catalogue record for this book is available from the British Library

ISBN (10): 1-5275-3770-6
ISBN (13): 978-1-5275-3770-5

TABLE OF CONTENTS

CHAPTER ONE

PROMISING INSIGHTS INTO PROTEOMIC AND LIPIDOMIC FEATURES IN SOME CEREALS

Abstract

Cereals constitute the main source of food; they are also nutraceutical with potential therapeutic effects as they contain numerous functional nutrients including essential proteins and lipids.

Although, cereals are often exposed to abiotic stresses such as heavy metal (e.g., cadmium (Cd)) and salt (NaCl) contamination, identified stress-related genes could be explored using omic tools to mitigate such problems. Characterization of novel genes from halophyte/glycophyte cereal encoded proteins with interesting structure and function such as enzymes are among the recent advances in cereal biotechnology. Identification of these novel genes would be done by genomic approaches, i.e., global transcriptomic analysis and production of expressed sequence tags (ESTs), as well as by functional proteomics.

In addition to proteins, cereal lipids are also of increasing interest in the biotechnology of cereals due to their multiple applications such as in medicine, cosmetic and food technology. Oiled cereal seeds could be potential sources of antioxidants, dietary polyunsaturated fatty acids, and biomolecules such as eicosanoids (prostaglandins and lipoxins) indispensable for health.

The chapter covered an important emerging field of research concerning the use of cereal to produce compounds of high value, i.e., proteins and lipids with applications for both human health and agricultural purposes.

Keywords: *Avena sativa*, lipids, enzymes, proteins, stress, valorization, biotechnological applications.

1. Introduction

Cereals are among the most staple foods worldwide. They are edible and have versatile nutrient compounds that could be used as either foods or nutraceuticals. Cereals are recognized to be useful for a healthier world. The assessment of the vital functions of cereal components is important to satisfy targeted objective uses. Oils, enzymes and other biomolecules found in cereals could be valorized in many fields, e.g., agronomics, nutraceuticals, food, and biotechnological applications. To this end, oat, for instance, which is undervalued in comparison with the other staple cereals (wheat, barley, and rice), has an interesting biomolecular composition especially lipids, proteins, fibers and antioxidants indispensable for health. The valorization aspects of oat (*Avena sativa*) are highlighted in many reports of Ben Halima et al. (2014a, b; 2015a, b; 2016; 2019 and unpublished data). From these reports, oat lipids potentially exhibit protective effects against deltamethrin (a pesticide frequently used in Tunisian agricultural areas), that causes reproduction damage in animal models, i.e., male mice (Ben Halima et al., 2014a). Moreover, oat enzymes mainly amylases and chitinases were identified by functional proteomics (Ben Halima et al., 2016; Ben Halima, 2019). In addition, we found an oat fraction with an effective amylolytic activity that could be used as a potential additive in bread to ameliorate its characteristics (Ben Halima et al., 2015b). For more information on oat lipids and nutriment compounds valorization for potential uses in industrial applications, refer to Ben Halima et al. (2015a).

Oat could be valorized in agronomics as it may tolerate abiotic stresses such as from salt (NaCl) and cadmium (Cd) contamination (Ben Halima et al., 2014b; unpublished data). In fact, cereals like all the other living organisms could be affected directly or indirectly by environmental pollution and contamination.

Abiotic stresses imposed by salinity and heavy metals pose serious threats to the growth and productivity of crops worldwide. It is ideal to develop plants with salinity and heavy metals tolerance.

Identification, characterization, and isolation of novel genes responsive to environmental stresses in economically important crops such as oat would be promisingly done by genomic approaches to understand their tolerance and regulatory mechanisms. Among the promising approaches, global transcriptome analysis and the production of expressed sequence tags (ESTs) are considerable, as they are relatively cost-effective, efficient and

rapid for identifying and discovering novel genes involved in the regulation of abiotic stresses (Sreenivasulu et al., 2007).

Structural insights of uncharacterized proteins (RNA Recognition Motif, annexin, enzymes, etc.) from glycophyte/halophyte plants may be considered as a predicted model to better characterize these proteins in cereals.

An attempt to summarize some own reports on recent developments in cereal biotechnology and in particular proteomic and lipidomic feature of oat (*Avena sativa*) is the focus of the current chapter.

2. Proteomics features in some cereals

2.1 Mitigation of abiotic stresses

2.1.1. Wheat Annexins

Wheat is one of the most important cereals worldwide as it is the most widely consumed and grown crop. The annual world production of wheat is estimated at 700 million tons (USDA, 2014). Environmental stresses, especially abiotic ones, negatively affect all cereal production including wheat (Lobell and Field, 2007).

Annexins are proteins with suggested functions in response to environmental stresses and signaling during plant growth and development. They can be considered as an evolutionary conserved multigene family of calcium-dependent phospholipid binding proteins that play important roles in stress resistance and plant development (Xu et al., 2016).

A recent report on the characterization of two annexin genes identified from Tunisian durum wheat varieties (*Triticum turgidum* L. *subsp. Durum*, cultivar Mahmoudi) has been achieved by *in silico* approach (Harbaoui et al., 2018). The primary and secondary structures of the resulting proteins (annexins) were analyzed, as well as their homology 3D-model structures. Multiple sequence alignments and the evolutionary relationships between the two identified annexins and their homologs were also studied. Insights into conserved protein motifs analysis and subcellular location prediction were performed to understand their eventual molecular function. In fact, some of the studied annexin motifs were found to be associated with calcium-dependent phospholipid binding/calcium ion binding function. Furthemore, the two newly annexins isolated from durum wheat might

play an active role in modulating plant cell responses to abiotic stress. These annexins could confer tolerance to various abiotic stresses when tested on yeast cells (Harbaoui et al., 2018).

2.1.2. *Aeluropus littoralis* stress-related genes/proteins

Halophytes are plants that are able to cope with high concentrations of salt so that they can complete their life cycle under at least 200 mM NaCl (Flowers et al., 2015). These plants that adapt to grow in saline environments offer important genetic resources (novel genes that code for functional proteins) involved in the salt stress adaptation. These particular novel genes can be isolated and successfully transferred y to glycophytes (plants growing in non-saline soils).

Aeluropus littoralis is a halophyte monocotyledonous plant belonging to the same family as wheat. A novel stress tolerance gene from *A. littoralis* designated as *AlSRG1* (*A. littoralis* Stress-Related Gene) has been studied by Ben Saad et al. (2018). *AlSRG1,* encoding a small RNA-binding protein could confer salt and drought tolerance in transgenic tobacco (Ben Saad et al., 2018). *AlSRG1* is a single copy gene with unknown function, isolated from an SSH (Suppression Subtractive Hybridization) cDNA library prepared using RNA extracted from the root of salt-stressed (300mM NaCl during 15 days) *A. littoralis* plants (Zouari et al., 2007). The characterization of the novel gene, *AlSRG1,* was performed based on bioinformatics analysis. In fact, phylogenetic and conserved domain prediction identified the protein *AlSRG1* as an uncharacterized protein of unknown function. The conserved domain database revealed that *AlSRG1* contains an RRM-SF superfamily domain. The RNA Recognition Motif (RRM), also known as RNA binding domain (RBD) or Ribonucleoprotein domain (RNP) was highlighted in *AlSRG1* primary, secondary and tertiary structures with the two conserved motifs RNP1 and RNP2 being part of it. These two conserved motifs contain the active site for RNA-binding. In addition, a 3D-model of the conserved domain could be built from known X-ray structures and sequence alignments. RRM is a highly abundant domain in eukaryotes found in proteins involved in post-transcriptional gene expression processes including mRNA and rRNA processing, RNA export, and RNA stability. The analysis with quantitative RT-PCR revealed higher *AlSRG1* transcript accumulation under salt stress and PEG-induced osmotic stress (Ben Saad et al., 2018). Taken together, the results of this study (Ben Saad et al., 2018) showed that this novel stress-tolerance gene might play an important positive modulation role in abiotic stress tolerance.

2.1.3. Effect of Cd and NaCl stresses on oat

The effects of cadmium (Cd) and salt (NaCl) stress on seed germination, seedling growth and antioxidant system of oat (*Avena sativa* L.) have been studied by Ben Halima et al. (2014b; unpublished data).

In fact, Cd is a major heavy metal pollutant, which is highly toxic to cereals. Vast agricultural areas worldwide could be contaminated with Cd from industrial wastes. In addition, abiotic stresses imposed by drought and salinity pose serious threats to the growth and productivity of crop plants worldwide.

Therefore, it is ideal to develop plants tolerant to Cd and salinity. Oat (*Avena sativa*) is an important crop in North America and northern Europe. The potential of *Avena sativa* plant may be reached as a genetic resource to improve salt and Cd tolerance in plants (Ben Halima et al., 2014b; unpublished data).

When *Avena sativa* plants were exposed to various concentrations of Cd and NaCl, they were found to express Cd and NaCl tolerance and accumulation *in planta*. Although high concentrations of Cd and NaCl could inhibit plant growth and reduce chlorophyll content, low and moderate concentration of such stresses could be tolerated by *Avena sativa* due to the efficiency of its antioxidant system. Indeed, abiotic stresses are the cause of imbalance in the oxidant/antioxidant system of plants. Measurements of the root and leaf endogenous Na^+, K^+ and Cd levels in *Avena sativa* plants stressed by Cd and NaCl compared to control conditions, showed an evident higher Na^+ and Cd accumulation in roots. Furthemore, the steady-state levels of transcripts of three-related genes (catalase, manganese superoxide dismutase and ascorbate peroxidase) were found at a significantly higher level in plants stressed by Cd or NaCl than control plants.

Taken together, these results of Ben Halima et al. (2014b; unpublished data) showed that oat (*Avena sativa*) could be a potentially useful candidate gene for engineering Cd and salt tolerance in cultivated plants.

3. Case study: Oat (*Avena sativa* L.)

Oat (*Avena sativa* L.) is a member of the *Poaceae* family (*Gramineae*), is one of the most cultivated crops, and is a promising crop plant worldwide. Oats are cultivated for grain, fodder, straw, and feed (Särkijärvi and

Saastamoinen, 2006) over more than 9 million hectares globally (FAO, 2014). Wild oats are diploid, but those cultivated (*Avena sativa*) are hexaploid with an estimated 1C genome size of 13.23pg, corresponding to about 13000 Mbp (Bennett and Smith, 1976). The common oat (*Avena sativa* L.) is an allohexaploid (2n=6x=42) crop species. Thus, oats have a large genome, which remains to be assessed. Complex polyploidy genomes such as oats need substantial contributions to the bioinformatics pipeline and drafting of the genome.

3.1. Glycoside hydrolases

3.1.1. Generalities

The abundance and the diversity of oligo- and polysaccharides provides a wide range of biological roles attributed either to these carbohydrates or to their relevant enzymes, i.e., the glycoside hydrolases (GHs). Carbohydrates could be in the form of mono-, di-, oligo- and polysaccharides and could be subjected to biocatalysis by GHs, enzymes hydrolyzing these sugars.

The biocatalysis by this family of enzymes is highly attractive for the generation of products used in potential application, e.g., pharmaceuticals and food industries. It is thus very important to extract and characterize such enzymes, particularly from plant tissues. Generally, GHs hydrolyze glycosidic bonds through the acid-base–assisted catalysis, deploying either double- or single-displacement mechanisms, leading to retention or inversion of anomeric configuration, respectively.

Obviously, starch is the most abundant grain component in most cereal cultivars such as wheat, barley, and oat, but it is also known as a storage polysaccharide present in the other major starch-containing crops worldwide like in pulses such as peas and beans, or in roots and tubers, such as cassava and potato. The starch that serves as an important reserve of energy in plants has provided dietary carbohydrates for animals and man for a long time. β-amylase catalyzes the liberation of maltose from the non-reducing ends of α-1,4-linked glucans and is, therefore, required for starch breakdown particularly during plant germination. The β-amylase isoforms conserved two Glu residues assigned as the "putative" catalytic residues, which would act as an acid and base pair in the catalytic process. A similar core $(\beta/\alpha)_8$-barrel architecture was found in the predicted oat β-amylases with a specific location of the active site in a pocket-like cavity structure made at one end of this core $(\beta/\alpha)_8$-barrel domain (Ben Halima et al., 2016). This suggests accessibility of the non-reducing end of the

substrate towards the β-amylase isoforms and thus confirming that they are exo-acting hydrolases (Ben Halima et al., 2016; Ben Halima, 2019).

Chitin is a polysaccharide, which is an important structural component of insects, and other invertebrates such as crustaceans, mollusks, and nematodes. Chitinases are, thus, an interesting type of glycoside hydrolytic enzymes that need to be further characterized in many organisms, particularly in cereals. Certainly, chitinases have found promising applications in diverse areas ranging from academic to industrial scales such as in medicine, agriculture and food industries. One promising feature of these applications includes growth inhibition of pathogenic fungi in industrial food processing. The catalytic domain of predicted oat chitinases adopts an α-helix-rich fold, stabilized by 3 disulfide bridges in accordance with other plant class I chitinases (Ben Halima, 2019; unpublished data).

3.1.2. β-amylases and chitinases

Since the amino acid sequence of oat β-amylase is not available in the protein data bank, an attempt was made to identify the catalytic amino acid residues as well as the 3-D structure of the enzyme by bioinformatic tools to define/predict biochemical properties of a soluble oat extract. The primary aim of the study by Ben Halima et al. (2016) was to identify a new sequence of exo-amylase through mass spectrometry and bioinformatic tools and to get insight into the oat genome. The characterization of glycoside hydrolases, for instance, GH14 family β-amylases and GH19 family chitinases, in oat seedling extract (Ben Halima et al., 2016; Ben Halima, 2019; unpublished data) may be considered as a further step in the characterization of these enzymes in plants.

A predicted β-amylase from oat seedling extract cDNA (1074 bp) corresponds to a 357 amino acid residue protein of a theoretical molecular weight of 41 kDa and with a theoretical isoelectric point of 4.84 (Ben Halima, 2019).

Moreover, a predicted oat chitinase cDNA (888 bp) corresponds to a 295 amino acid residue protein of a theoretical molecular weight of the protein of 31 kDa and with a theoretical isoelectric point of 9.10 (unpublished data). These two proteins are named AsBAMY1 and AsChi1, for *Avena sativa* β-amylase and chitinase, respectively (Ben Halima, 2019; unpublished data).

The sequencing of the *A. sativa* genome (drafting as well as complete sequencing) will certainly accelerate the identification of other catalytic enzymes with applications in biotechnology.

3.2. Lipidomics

In its large-scale concept, lipidomics refers to the study of the structure and function of the complete set of lipids in biological systems as well as their interaction with other lipids and metabolites including proteins (Wenk, 2005; Rolim et al., 2015; Gugiu, 2017). Indeed, lipidomics is a relatively new field, related to the advancement of high technologies of separation such as liquid chromatography (LC) and mass spectrometry (MS). Therefore, lipidomics (Wenk, 2005; Rolim et al., 2015) is developed later than proteomics along with the newly developed technologies (MS and other relevant technologies that allow the investigation of structural and biological diversity of lipids) (Fahy et al., 2011). There is a so-called Lipid Maps Structural Database (LMSD), which is the main database of lipids that contain the structure and annotation of biologically relevant lipids (Sud et al., 2007). In fact, lipids are biomolecules with complicated structural variability and can be classified into polar or neutral (nonpolar) lipids. The polar lipids include phospholipids, glycolipids, and sphingolipids. The nonpolar lipids include triacylglycerols and cholesterols. Lipids could be involved in many processes such as in nutrition and storage of energy and are of relevance for signal transduction processes, cell differentiation, and phagocytosis. Moreover, lipids could be classified based on their chemical structure; thus, there are so many classes of lipids with precise nomenclature (Fahy et al., 2005; Fahy et al., 2009). These lipids could be either abundant lipids or degraded lipid metabolites such as fatty acids (FA) and mono/di-glyceride (MG, DG) (Han and Gross, 2005).

Untargeted lipidomics aims to focus on measuring lipids whatever their classes maybe using omic workflows similar to proteomics (Washburn et al., 2001) and genomics (Yadav, 2007). However, it is worth to noting that targeted methods are normally used for certain lipids, mainly those of metabolites, specifically because of their low abundance (Han and Gross, 2005; Gugiu et al., 2006). Hence, targeted or untargeted MS analysis of lipids could be understood through lipidomics using either LC (Castro-Perez et al., 2010) or shotgun (Han and Gross, 2005) approaches coupled with tandem MS. Gugiu (2017) described the LC/MS methodology for lipid identification, which is similar to established omics workflows, e.g., proteomics (Washburn et al., 2001) or genomics (Yadav, 2007).

3.2.1. Characterization of oils and fats

Cereals constitute the main source of food. Cereal lipids are of increasing interest because of their multiple roles, e.g., in food technology and as potential sources of dietary polyunsaturated fatty acids (PUFAs) in significant amounts (Price and Parsons, 1974). Yet, oilseeds represent many versatile commodities due to their use essentially in food, feed, and medicine. Nevertheless, the total lipid content in cereals varies considerably. Furthemore, different parts of cereal seeds have different lipid contents. In wheat, for example, the germ contains about 25-30 % lipid while its endosperm has only 1 % lipid. However, considerable amounts of germ lipids of wheat are lost during milling (Galliard and Barnes, 1980).

3.2.2. Oat lipidomics

Compared to other cereals, the specific feature of *Avena sativa* L. grains is the high oil content, which could range from 3% to 18% of the grain weight in different cultivars (Banas et al., 2007; Frey and Holland, 1999; Peterson and Wood, 1997), and deposited mainly in the grain endosperm tissues (Price and Parsons, 1979; Youngs et al., 1977). However, most oat cultivars comprise about 5-6% of oil and 55-60% of the starch in the grain (Banas et al., 2007; Welsh, 1995; Doehlert et al., 2001), and high-lipid oat remains a potential oil crop (Heneen et al., 2009). Oats contain a wide range of active compounds, including avenanthramides, starches, hydrocolloid β-D-glucan, vitamins, saponins and other antioxidants (mostly phenolic esters) and a relatively high content of total lipids with a high percentage of UFAs (Kurtz and Wallo, 2007; Sur et al., 2008). Total lipids could reach 18% (Frey and Holland, 1999; Peterson and Wood, 1997), and about 41% of the groat lipids are triglycerides, while 5% are free fatty acids (Youngs, 1978). In fact, oats, apart from maize, comprise a high oil content used for a wide range of beneficial purposes. To this end, Ben Halima et al. (2015a) have reviewed oat lipids. On the other hand, oat lipid fraction was proven to be a protective agent against pesticide, e.g., deltamethrin causing reprotoxicity in male mice (Ben Halima et al., 2014a).

Studying lipolytic enzymes such as phospholipases from cereals would be very promising as these enzymes might play important roles in plant growth, development and stress response. In addition, these enzymes might find external applications such as in medicine and pharmaceuticals. Phospholipases D (PLD) and C (PLC) are among lipolytic enzymes that

hydrolyze the phosphodiester linkages of the head group of membrane phospholipids (Hong et al., 2016).

As mentioned above, oats are unique among the common cereal grains since oats have a high lipid content and their lipolytic enzymes are 10–15 times more active than those of wheat (Matlashewski et al., 1982). Identifying and characterizing oat phospholipases are further research axis undertaken by Ben Halima et al. (unpublished data).

4. Conclusions

Proteins and lipids are important constituents of all living organisms in particular cereals. Tremendous progress has been made in recent years in various aspects of proteomes and lipidomes; however, our understanding of these proteins and lipids in cereals is far from complete. For example, oats are unique among the other cereal grains due to the high lipids content and their promising enzymatic activities, in particular, amylases, chitinases, and phospholipases. The investigation and understanding of the relationship of proteomes and lipidomes would be an outstanding feature in the recent development in cereal biotechnology. The mechanism of action of major phospholipases and their roles in lipid signaling in cereals remains to be elucidated. Further studies with bioinformatics analysis would be an interesting key to accelerate discovering novel proteins as well as lipids in cereals.

5. Acknowledgments

The author thanks the Tunisian Ministry of Higher Education and Scientific Research for facilities. The author is very grateful to Erasmus Nchuaji Tang, a researcher at the Department of Biochemistry, University of Yaounde I-Cameroon, for his help reading the manuscript.

6. References

Banas, A., Dexbski, H., Banas, W., Heneen, W.K., Dahlqvist, A., Bafor, M., Gummeson, P.O., Marttila, S., Ekman, A., Carlsson, A.S. and Stymne, S. (2007). Lipids in grain tissues of oat (*Avena sativa*): differences in content, time of deposition, and fatty acid composition. J. Exp. Botany. 58, 2463–2470.

Ben Halima, N. (2019). Prediction and analysis of GH14 family β-amylases in oat seedling extract: structure and function insights using *in silico* approaches. International Journal of Biological Macromolecules. 125, 361-369.

Ben Halima, N. New insights into phospholipases in oat (*Avena sativa*) from bioinformatic analysis. Unpublished data.

Ben Halima, N. Glycoside hydrolases in oat (*Avena sativa*) seedling extract: new insights from bioinformatic analysis into GH19 family chitinases. Unpublished data.

Ben Halima, N., Ben Saad, R., Dabbous, A., Chaabene, Z. and Elleuch, A. Changing in morphological and biochemical behavior of oat (*Avena sativa*) to cope with excess of cadmium. Unpublished data.

Ben Halima, N., Ben Saad, R., Khemakhem, B., Fendri, I. and Abdelkafi, S. (2015a). Oat (*Avena sativa* L.): Oil and nutriment compounds valorization for potential use in industrial applications. J. Oleo. Sci. 64, 915-932.

Ben Halima, N., Borchani, M., Fendri, I., Khemakhem, B., Gosset, D., Baril, P., Pichon, C. Ayadi, M.A. and Abdelkafi, S. (2015b). Optimized amylases extraction from oat seeds and its impact on bread properties. Inter. J. Biol. Macromol. 72, 1213-1221.

Ben Halima, N., Ben Slima, A., Moalla, I., Fetoui, H., Pichon, C., Gdoura, R. and Abdelkafi, S. (2014a). Protective effects of oat oil on deltamethrin-induced reprotoxicity in male mice. Food Funct. 5, 2070-2077.

Ben Halima, N., Ben Saad, R., Ben Slima, A., Khemakhem, B., Fendri, I. and Abdelkafi, S. Effect of salt stress on stress-associated genes and growth of *Avena sativa* L. (2014b). ISESCO J. Sci. Technol. 10, 73-80.

Ben Halima, N., Khemakhem, B., Fendri, I., Ogata, H., Baril, P., Pichon, C. and Abdelkafi, S. (2016). Identification of a new oat β-amylase by functional proteomics. Biochimica et Biophysica Acta. 1864, 52–61.

Ben Saad, R., Ben Halima, N., Ghorbel, M., Zouari, N., Ben Romdhane, W., Guiderdoni, E., Al-Doss, A., and Hassairi, A. (2018). *AlSRG1*, a novel gene encoding an RRM-type RNA-binding protein (RBP) from *Aeluropus littoralis*, confers salt and drought tolerance in transgenic tobacco. 150, 25-36.

Bennett, M.D. and Smith, J.B. (1976). Nuclear DNA amounts in angiosperms. Philos. Trans. R. Soc. Lond. B 274, 227–274.

Castro-Perez, J.M., Kamphorst, J., DeGroot, J., Lafeber, F., Goshawk, J., Yu, K., Shockcor, J.P., Vreeken, R.J. and Hankemeier, T. (2010). Comprehensive LC-MS E lipidomic analysis using a shotgun approach

and its application to biomarker detection and identification in osteoarthritis patients. J. Proteome Res. 9, 2377–2389.

Doehlert, D.C., McMullen, M.S. and Hammond, J.J. (2001). Genotypic and environmental effects on grain yield and quality of oat grown in North Dakota. Crop Sci. 41, 1066–1072.

Fahy, E., Cotter, D., Sud, M. and Subramaniam, S. (2011). Lipid classification, structures and tools. Biochim Biophys Acta. 1811, 637–647.

Fahy, E., Subramaniam, S., Brown, H., Glass, C., Merrill, J.A., Murphy, R., Raetz, C., Russell, D., Seyama, Y., Shaw, W., Shimizu, T., Spener, F., van Meer, G., Vannieuwenhze, M., White, S., Witztum, J. and Dennis, E.A. (2005). A comprehensive classification system for lipids. J. Lipid Res. 46, 839–861.

Fahy, E., Subramaniam, S., Murphy, R., Nishijima, M., Raetz, C., Shimizu, T., Spener, F., van Meer, G., Wakelam, M. and Dennis, E.A. (2009). Update of the LIPID MAPS comprehensive classification system for lipids. J. Lipid Res. 50, S9–S14.

FAO. Food and Agriculture Organization of the United Nations: FAOSTAT. http://faostat.fao.org/ (accessed April 2014).

Flowers, T.J., Munns, R. and Colmer, T.D. (2015). Sodium chloride toxicity and the cellular basis of salt tolerance in halophytes. Ann Bot. 115, 419-431.

Frey, K.J. and Holland, J.B. (1999). Nine cycles of recurrent selection for increased grain-oil content in oat. Crop Sci. 39, 1636–1641.

Galliard, T. and Barnes, P.J. (1980). The biochemistry of lipids in cereal crops. In Biogenesis and Function of Plant Lipids, Eds, Mazliak P, Benveniste P, Costes C, Douce R. Elsevier, Amsterdam; pp. 191-198.

Gugiu, G.B. (2017). Lipid identification by untargeted tandem mass spectrometry coupled with ultra-high-pressure liquid chromatography. Sanjoy K. Bhattacharya (ed.), Lipidomics: Methods and Protocols, Methods in Molecular Biology, vol. 1609, DOI 10.1007/978-1-4939-6996-8_8, © Springer Science+Business Media LLC.

Gugiu, B.G., Mesaros, C.A., Sun, M., Gu, X., Crabb, J.W. and Salomon, R.G. (2006). Identification of oxidatively truncated ethanolamine phospholipids in retina and their generation from polyunsaturated phosphatidylethanolamines. Chem. Res. Toxicol. 19, 262–271.

Han, X. and Gross, R.W. (2005). Shotgun lipidomics: electrospray ionization mass spectrometric analysis and quantitation of cellular lipidomes directly from crude extracts of biological samples. Mass Spectrom Rev. 24, 367–412.

Harbaoui, M., Ben Saad, R., Ben Halima, N., Choura, M. and Brini, F. (2018). Structural and functional characterization of two novel durum wheat arnexin genes in response to abiotic stress. Functional Plant Biology. 45, 542-552.

Heneen, W.K., Banas, A., Leonova, S., Carlsson, A.S., Marttila, S., Debski, H. and Stymne, S. (2009). The distribution of oil in the oat grain. Plant Signal Behav. 4, 55-56.

Hong, Y., Zhao, J., Guo, L., Kim, S.C., Deng, X., Wang, G., Zhang, G., Li, M. and Wang, X. (2016). Plant Phospholipases D and C and their diverse functions in stress responses. Prog. Lipid Res. 62, 55-74.

Kurtz, E.S. and Wallo, W. (2007). Colloidal oatmeal: history, chemistry and clinical properties. J. Drugs Dermatol. 6, 167–170.

Lobell, D.B. and Field, C.B. (2007). Global scale climate-crop yield relationships and the impacts of recent warming. Environ. Res. Lett. 2: 014002.

Matlashewski, G.J., Urquhart, A.A., Sahasra-dudhe, M.R. and Altosaar, I. (1982). Lipase activity in oat flour suspensions and soluble extracts. Cereal Chem. 59, 418-422.

Peterson, D.M. and Wood, D.F. (1997). Composition and Structure of High-Oil Oat. J. Cereal Sci. 26, 121–128.

Price, P.B. and Parsons, J. (1979). Distribution of lipids in embryonic axis, bran-endosperm, and hull fractions of hulless barley and hulless oat grain. J. Agric. Food Chem. 27, 813–815.

Price, P.B. and Parsons, J.G. (1974). Lipids of six cultivated barley (*Hordeum vulgare*) varieties. Lipids. 9, 560-566.

Rolim, A.E.H., Henrique-Araújo, R., Ferraz, E.G., de Araújo Alves Dultra, F.K. and Fernandez, L.G. (2015). Lipidomics in the study of lipid metabolism: current perspectives in the omic sciences. Gene 554, 131–139.

Särkijärvi, S. and Saastamoinen, M. (2006). Feeding value of various processed oat grains in equine diets. Livestock Sci. 100, 3–9.

Sreenivasulu, N., Sopory, S.K. and Kavi Kishor, P.B. (2007). Deciphering the regulatory mechanisms of abiotic stress tolerance in plants by genomic approaches. Gene 388, 1–13.

Sud, M., Fahy, E., Cotter, D., Brown, A., Dennis, E.A., Glass, C.K., Merrill, A.H., Murphy, R.C., Raetz, C.R.H., Russell, D.W. and Subramaniam, S. (2007). LMSD: LIPID MAPS structure database. Nucleic Acids Res. 35(Database issue), D527–D532.

Sur, R., Nigam, A., Grote, D., Liebel, F. and Southall, M.D. (2008). Avenanthramides, polyphenols from oats, exhibit anti-inflammatory and anti-itch activity. Arch. Dermatol. Res. 300, 569–574.

USDA. 2014. http://www.worldwheatproduction.com/

Washburn, M.P., Wolters, D. and Yates, J.R. (2001). Large-scale analysis of the yeast proteome by multidimensional protein identification technology. Nat. Biotechnol. 19, 242–247.

Welsh, R.W. (1995). The chemical composition of oats. In: Welsh RW (ed.). The oat crop, London: Chapman & Hall; p. 279–320.

Wenk, M.R. (2005). The emerging field of lipidomics. Nat. Rev. Drug Discov. 4, 594–610.

Xu, L., Tang, Y., Gao, S., Su, S., Hong, L., Wang, W., Fang, Z., Li, X., Ma, J., Quan, W., Sun, H., Li, X., Wang, Y., Liao, X., Gao, J., Zhang, F., Li, L. and Zhao, C. (2016). Comprehensive analyses of the annexin gene family in wheat. BMC Genomics. 17: 415.

Yadav, S.P. (2007). The wholeness in suffix-omics, -omes, and the word om. J Biomol Tech: JBT 18: 277.

Youngs, V.L. (1978). Oat lipids. Cereal Chem. 55, 591-597.

Youngs, V.L., Püskülcü, M. and Smith, R.R. (1977). Oat lipids I. Composition and distribution of lipid components in two oat cultivars. Cereal Chem. 54, 803–812.

Zouari, N., Ben Saad, R., Legavre, T., Azaza, J., Sabau, X., Jaoua, M., Masmoudi, K. and Hassairi, A. (2007). Identification and sequencing of ESTs from the halophyte grass Aeluropus littoralis. Gene 404, 61–69.

CHAPTER TWO

GLYCOSIDE HYDROLASES IN OAT (*AVENA SATIVA*) SEEDLING EXTRACT: NEW INSIGHTS FROM BIOINFORMATIC ANALYSIS INTO GH14 FAMILY *BETA*-AMYLASES AND GH19 FAMILY CHITINASES

Abstract

Oat (*Avena sativa* L.) seedling extract exhibited a high degree of catalytic activities. Bioinformatic tools were used to identify β-amylases and class I chitinases, the highly abundant enzymes in the oat seedling extract. These newly discovered oat enzymes are a member of the GH14 and GH19 family, respectively. Proteins in the *Avena sativa* seedling extract were separated by SDS-PAGE. These materials were digested with trypsin and the amino acid sequences of the tryptic peptides were determined by LC/ESI/MS/MS and database searches. These sequences were used to identify cDNAs from expressed sequence tags (ESTs) and Transcripts Shotgun Assembly (TSAs) of *Avena sativa*. Based upon EST and TSA sequences, at least 6 and 4 predicted enzymes were identified and designated as β-amylases and chitinases, respectively. Insights into the structural characterization of each of the oat predicted enzymes were analyzed using *in silico* approaches.

The results provide a detailed view of the main components involved in catalysis in these two enzymes.

Keywords: *Avena sativa*; Glycoside hydrolases; Bioinformatic analysis.

1. Introduction

The use of additives in industrial food processing is of high interest as they could improve or preserve food products. The promising additives are enzymes especially glycoside hydrolases (GHs), for instance, amylases

and chitinases which could have beneficial impacts on glycoside substrates enriched foods such as bread flour and its derivatives, i.e., starch, maltose, β-dextrin, etc. These enzymes are also potential candidates for anti-pathogen agents, e.g., anti-insects and antifungal agents. Fungi are a prominent source of contamination of all groups of food including cereals, fruits, vegetables, meat, milk, and their products. They are involved in food spoilage by secreting highly poisonous mycotoxins (Filtenborg et al., 1996).

The classification of GHs is commonly based on substrate specificity, mode of action or amino acid sequence similarities (Henrissat and Davies, 1997). O-glycoside hydrolases are classified, for instance, on the basis of substrate specificities as recommended by The International Union of Biochemistry and Molecular Biology (IUBMB) and are expressed in the EC number with the given code EC 3.2.1.x, where x generally represents the substrate specificity (Henrissat and Davies, 1997).

Chitin, an insoluble polysaccharide of β-(1,4)-linked N-acetyl-D-glucosamine residues (Muzzarelli, 1999), is the main constituent of the cell walls of many fungal plant phytopathogens. It can be decomposed by chitinases (EC 3.2.1.14) and herein plant chitinases are expressed during plant growth as well as plant and phytopathogen interactions. Therefore, plant chitinases have a major role as pathogenesis-related (PR) proteins that are involved in defense responses of a plant against its pathogens (Kasprzewska, 2003). Most characterized chitinases are clustered into families 18 and 19 of the GHs based on primary structures similarities of their catalytic domains including class III and V and class I, II, IV, VI and VII chitinases, respectively (Henrissat, 1991; Henrissat and Bairoch, 1993). However, some chitinases have also been identified as belonging to families' GH23 and GH48 (Arimori et al., 2013; Fujita et al., 2006).

The glycan metabolism involved many kinds of carbohydrate-active enzymes (CAZymes), which are grouped into sequence-based families on the CAZY database (Lombard et al., 2014), and the structural fold, as well as the catalytic mechanism, are highly conserved within these families. The most important CAZymes that depolymerize carbohydrate polymers are GHs (Munoz-Munoz et al., 2017). GH18 and GH19 chitinases are extensively characterized. Those from GH18 were exemplified to adopt the retaining mechanism, producing β-anomers after hydrolysis, while, GH19 commonly adopt the inverting mechanism, producing α-anomers after hydrolysis (Ohno et al., 1996; Udaya Prakash et al., 2010; Junges et al., 2014).

Extracts from different higher plants, especially from cereals, were proven to have class I chitinase isoforms and those from oat seed extracts were demonstrated to be more effective towards *Penicillium roqueforti*, a major contaminating fungal species in food, as opposed to extracts of other cereal seeds (Sørensen et al., 2010).

Starch, a polymer of glucose consists of two primary components, amylose, and amylopectin. Amylose is a linear polymer of glucose with α-1,4 linkages, while amylopectin is a branched glucose polymer in which linear chains of α-1,4 are interlinked with α-1,6 linkages (Muralikrishna and Nirmala, 2005).

Current thinking is that starch degradation in plants would certainly require the concerted activity of GHs, especially amylases that break down cereal starch into glucose, which is then taken up as the substrate for the plant germination and growth.

β-amylase (α-1,4-glucan maltohydrolase; EC 3.2.1.2; CAZY family GH14) is an inverting GH (Bourne and Henrissat, 2001; Cantarel et al., 2009), which is distributed throughout many organisms such as in prokaryotes or in eukaryotes, i.e., higher plants including cereals. β-amylase is a crucial enzyme that contributes to the complete degradation of starch into metabolizable or fermentable sugars during the germination or malting of cereal grains. However, different isoforms of β-amylase may exist in the same plant extract; thus, distinct structures and functions are ascribed to β-amylases (Fulton et al., 2008). To this, Daussant et al. (1994) have highlighted the diversity of cereal β-amylase isozymes when determined their activity in the kernels of 29 cultivars and inbred lines of seven kinds of cereal belonging to four tribes of the *Gramineae* such as *Hordeum vulgare* L. and *Avena sativa* L.

The common oat (*Avena sativa* L.) is a member of the *Poaceae* family (*Gramineae*) and is a promising plant for the future. It is edible and beneficial thanks to its nutritional, medicinal and pharmaceutical uses and, hence, recognized to be useful for a healthier world (Ben Halima et al., 2015a). The cultivated oat (*A. sativa* L.) is an allohexaploid ($2n = 6\times = 42$) species with a large genome which is still to be fully assessed.

Complex polyploidy genomes such as oats need substantial contributions to the bioinformatics pipeline and drafting of the genome. The oat genome has not yet been fully sequenced and the majority of GHs, especially

sequences of amylases and chitinases from oat and other cereals, has not been well characterized.

One important oat (*Avena sativa* L.) β-amylase sequence has already been characterized by functional proteomics in our previous study (Ben Halima et al., 2016).

Though Sørensen et al. (2010) have tried to characterize oat class I chitinase, this chitinolytic enzyme has not been subjected to further biochemical characterization. In this study, extracts of oat seedling were identified as a potential food additive through the catalytic activity of highly abundant proteins from GH14 and GH19 family. The proteins in the oat seedling extracts are isolated and characterized. Aiming to understand the mechanisms of hydrolytic action of oat β-amylases (AsBAMYs) and chitinases (AsChiIs), the sequencing of peptides resulting from tryptic digestion allowed identifying EST and TSA sequences from that the AsBAMYs and AsChiIs genes were compared by computational bioinformatics analysis.

Since an amino acid sequence of one oat chitinase is available in the protein data bank, an attempt has been made to more characterization and identification of the catalytic amino acid residues as well as the 3-D structure of the enzyme by bioinformatics tools to define/predict biochemical properties of soluble oat seedling extracts.

The primary aim of this work is to identify new insights into oat GH14 and GH19 family based on mass spectrometry and bioinformatic tools and so far, to get insights about the oat genome.

2. Materials and methods

2.1. Plant material and chemicals

Oat (*Avena sativa* L.) was used as plant material in which seeds extract containing GH family, especially GH14 and GH19 families, were the subject of this study. *A. sativa* seeds were purchased from a local market in Sfax, a coastal town in South-East Tunisia.

Chemicals used in this study were of reagent grade and were supplied by Invitrogen and Sigma Chemical Co. (St Louis, France).

2.2. Extraction of GHs from oat seedling

Oat seeds (*A. sativa*) germinated on wet tissue paper in a plastic box. They were grown in the dark for 5 days at room temperature. On day 10 after planting, oat seedlings were ground using mortar and pestle with 0.02 M sodium acetate buffer (pH 5.6), filtered through two layers of cheesecloth to remove large particles and the supernatant obtained was centrifuged at 15,000 × g for 20 min. The supernatant was used as an oat crude extract of GHs as well as start material for purification procedure.

As a crude enzyme, acetone was added to the oat crude extract (2:1; v:v), sample centrifuged at 14,000 × g for 15 min and the supernatant discarded. The partially delipidated acetone powder was resuspended in water. The mixture was stirred for 20 min at 4°C, sonicated for 5 min and finally centrifuged at 14,000 × g for 5 min before collecting the supernatant fraction, which was used as oat fraction enriched in GH activity.

For the purification procedure, the oat proteins were extracted as described above following a purification procedure of some steps of a novel oat β-amylase of 25 kDa according to the report of Uno-Okamura et al. (2004).

2.3. In-gel tryptic digestion and protein identification by mass spectrometry

Bands of interest were manually excised from gels and automated tryptic digestion was conducted as previously described (Abdelkafi et al., 2009, 2012; Fendri et al., 2009) or manually treated as follows. Gel bands were manually excised in a sterile laminar flow hood, transferred individually to 1.5 mL microtubes and then cut into cubes of roughly 1 mm^3. Gel cubes were destained for 1 h and 30 min at 4 °C using a solution of 45% acetonitrile and 55 mM ammonium bicarbonate. After gel cubes washing and in-gel trypsin proteolysis of proteins, the peptides produced were extracted onto Poros beads and purified with ZipTips (Millipore, France) as previously described (Beaufour et al., 2012).

Extracted proteolytic peptides were analyzed by nanoUltraHPLC–nanoESI UHR–QTOF MS. Experiments were performed using an UltiMate™ 3000 NanoRSLC System (Dionex, Sunnyvale, CA) connected to a Bruker MaXis UHR-QTOF 2 GHz mass spectrometer equipped with an online nano-ESI ion source. The LC–MS setup was controlled by Bruker Hystar™ software version 3.2. Peptides were pre-concentrated online on a Dionex Acclaim PepMap100 C18 reverse-phase precolumn (inner

diameter 100 μm, length 2 cm, particle size 5 μm, pore size 100 Å), and separated on a nanoscale Acclaim Pepmap100 C18 column (inner diameter 75 μm, length 25 cm, particle size 2 μm, pore size 100 Å) at a flow rate of 450 nL/min using a 2–35% gradient of acetonitrile in 0.1% formic acid. Peaks with the three highest intensities and a minimum of 400 ion counts were selected for CID MS/MS fragmentation using an isolation window of 3–9 Da depending on the m/z value.

Acquired MS/MS spectra were searched against the UniProtKB/Swiss-Prot/TrEMBL (database version 51.6; 257,964 sequence entries), non-redundant NCBI (http://www.ncbi.nlm.nih.gov) and the EST *A. sativa* L. database containing 25,400 entries (AM071411-CN180783) using the Mascot identification engine (version 2.3, Matrix Science, France). The search was conducted allowing for a maximum of two missed cleavages, 5 ppm tolerance for precursor ions and 0.04 Da for fragment ions, respectively. Methionine oxidation was allowed.

Moreover, a total of 17,776 ESTs were downloaded from GenBank for the oat cultivar (cv) CDC Dancer and a description of such EST generation can be accessed at the GrainGenes database site (Beattie et al., 2008). In fact, all CDC Dancer ESTs are from the seed and there are also 7097 oat ESTs from cv Ogle in GenBank, but are from a range of tissues and do not include enough seed ESTs to attempt seed protein sequence assemblies.

ESTs were assembled with the Seqman module of the Lasergene suite (DNAstar, Inc.) and all resulting contigs composed of three or more ESTs were annotated for the closest matching sequence of monocots (Anderson, 2014). Since contaminations from human (mainly keratins) origin could be present in the samples analyzed, the search in databases was restricted to plant species using UniProtKB/Swiss-Prot/TrEMBL, 49,887 sequence entries; NCBI nr, 551,056 sequence entries. In the case of peptides matching to multiple members of a protein family, the presented protein was selected based on both the highest score and the highest number of matching peptides.

2.4. *In silico* analysis

2.4.1. Retrieval of protein sequences

The amino acid sequences from the GHs that permitted comparison with the *de novo* sequencing of oat GH14 and GH19 proteins families were retrieved from the protein database of National Center for Biotechnology

Information (NCBI, http://www.ncbi.nlm.nih.gov/protein/). The sequences were saved in FASTA format. The *in silico* approach steps followed in this study consisted essentially on analysis of proteins from oat seedling by LC/MS/MS, MASCOT Search and Swiss-Prot Database as well as EST and TSA_*Avena sativa* Databases. Then, *de novo* sequencing of GH14 and GH19 proteins families from oat (*Avena sativa*) seedling were realized with structural characterization. This involved the prediction of primary and secondary structures through comparison with the retrieved protein sequences of GH14 and GH19 families and homology modeling analysis with some selected oat enzymes.

2.4.2 Sequence analysis

Bioinformatic analysis of the *A. sativa* peptide sequences, ESTs, genomic sequences and deduced protein sequences were performed using the following tools; Multiple sequence alignment was performed using the ClustalW algorithm (Thompson et al., 1994). The peptide sequences were compared with the NCBI (National Center for Biotechnology Information, USA) non-redundant sequence databases, the Transcriptome Shotgun Assembly (TSA) *A. sativa* database (GAJE01000001-GAJE01050182) and the Expressed Sequence Tag (EST) *A. sativa* database that contains 25,400 entries (AM071411-CN180783) using BLAST (Altschul et al., 2005).

Primary structure analysis was performed using the ExPASy Proteomics tools. The Translate tool (web.expasy.org/translate/) was used to translate DNA sequences to protein sequences, whereas the Compute pI/Mw tool (web.expasy.org/compute_pi/) was used to compute the theoretical isoelectric point (pI) and molecular mass (Bairoch et al., 2005; Gasteiger et al., 2003). The BioEdit software package (Hall, 1999) was used to manipulate, edit and compare DNA and amino acid sequences.

The prediction of the signal peptide sequence was performed using the signal P 4.1 application (Petersen et al., 2011).

To predict N- and O-glycosylation sites, the servers NetNGlyc 1.0 (www.cbs.dtu.dk/services/NetNGlyc/) and NetOGlyc 4.0 (www.cbs.dtu.dk/services/NetOGlyc/) (Steentoft et al., 2013) were used.

Phylogenetic analyses were performed using Molecular Evolutionary Genetics Analysis (MEGA) package version 7 (Kumar et al., 2016). The program MUSCLE (Edgar, 2004), implemented in MEGA7 package, was used to perform multiple alignments of amino acid sequences of AsBAMY

and AsChi and their homologous for phylogenetic analysis. The evolutionary history was inferred using either the Neighbor-Joining method (Saitou and Nei, 1987) or the unweighted pair group method with arithmetic mean (UPGMA) method (Sneath and Sokal, 1973). The evolutionary distances were computed using the JTT matrix-based method (Jones et al., 1992) and were in the units of the number of amino acid substitutions per site. All positions containing gaps and missing data were eliminated. The robustness of the inferred tree was evaluated by bootstrap (1000 replications) (Felsenstein, 1985).

2.4.3. Conserved protein motifs analysis and subcellular location prediction

Conserved protein motifs of the protein sequences from oat were analyzed using Multiple Expectation Maximization for Motif Elicitation (MEME) v.4.11.4 (Bailey and Elkan, 1994; Bailey and Gribskov, 1998; Bailey et al., 2015) (http://meme-suite.org) with the number of different motifs as 10, motif sites distribution as zero or one occurrence per sequence, and motifs width as 6 (minimum) and 50 (maximum).

The functional annotations of these motifs were analyzed by InterProScan (http://www.ebi.ac.uk/Tools/pfa/iprscan/) (Finn et al., 2014). The mapping between Pfam (http://pfam.xfam.org) analysis and Gene Ontology (GO) is provided by InterPro (Sangrador-Vegas et al., 2016).

The prediction on subcellular localization of oat protein was carried out using the CELLO v.2.5 server (http://cello.life.nctu.edu.tw/) (Yu et al., 2006).

2.4.4 Secondary structure prediction

The prediction of the protein secondary structure was performed using The PSIPRED Protein Sequence Analysis Workbench (http://bioinf.cs.ucl.ac.uk/psipred/).

The secondary structures were also determined by using the self-optimized prediction method (SOPMA) software (http://npsa-pbil.ibcp.fr/cgibin/npsa_automat.pl? page=/NPSA/npsa_sopma.html) (Geourjon and Deleage, 1995). The parameters of similarity threshold and window width were set to 8 and 17 and all the other parameters were taken as default.

2.4.5. Molecular and homology modeling

PDB file of oat protein was generated by Swiss-Model server (http://swissmodel.expasy.org/). In order to build a model of a protein domain, multiple sequence alignments were performed between full-length chitinase or β-amylase protein sequence and another protein domain sequence in these databases. To build the model of the oat protein with more homology, high-resolution structure of protein model in Swiss model server was selected as a template.

3. Results and discussion

3.1. Results

A previous study has demonstrated the presence of both activities of chitinases and glucanases in the apoplastic compartment of oat (*Avena sativa* L.) primary leaves of 10-day old plants (Fink et al., 1988). Furthermore, Uno-Okamura et al. (2004) have shown the presence of a novel type of β-amylase from oat (*Avena sativa*) seedlings. Their results, especially on the catalytic properties of the purified enzyme (β-amylase) and the molecular size (approximately 25 kDa) are certainly different from those of known amylases obtained from *Gramineae* caryopses.

Taken together, these findings as well as the fact that oat seed extracts have previously been denoted for their catalytic activity of highly abundant class I chitinases (Sørensen et al., 2010), have proven the presence of many sequences of chitinases (GH19) and β-amylases (GH14) in 10-day old oat seedlings extract. Through LC/MS/MS technique and bioinformatics tools, novel amino acid sequences of both activities of oat β-amylases and chitinases could be reconstructed, in spite of the only one previously deposited sequence of oat chitinase in GenPept (P86181.1).

3.1.1. Extraction and identification of *A. sativa* seedling proteins from GH14 and GH19 families

Oat (*Avena sativa* L.) seedlings from 10-day old plants were used as starting material for extracting proteins from GH14 and GH19 families, i.e., β-amylases and chitinases. In fact, this extract was also enriched in amylolytic activity such as β-amylases as described by previous reports (Ben Halima et al., 2016; Ben Halima et al., 2015b). An aliquot of this extract was analyzed by SDS-PAGE followed a Coomassie blue staining

step and a number of protein bands were excised from the preparative gel (**Figure 1**).

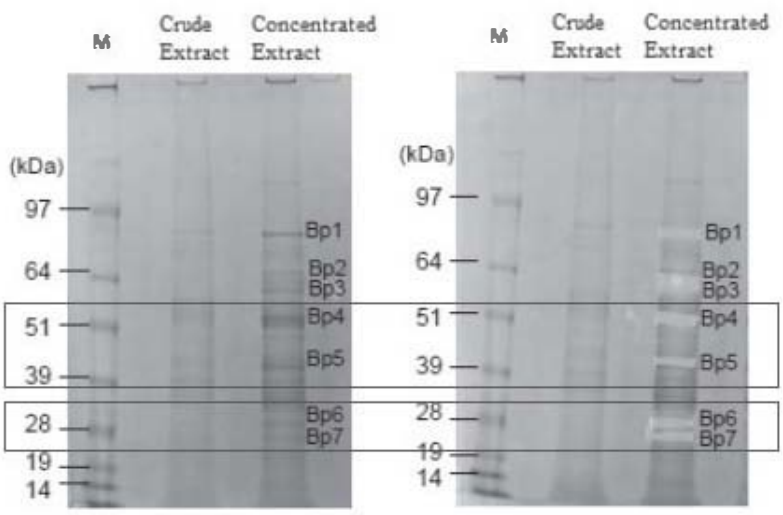


Figure 1. *The SDS-PAGE analysis of the 10-day old oat (Avena sativa) seedling extracts. Bp, Band of protein and M, molecular mass markers. The gel was stained with Coomassie blue. Bp 4 and 5 were matched to β-amylase according to Swiss-Prot database after in situ trypsin digestion and LC/MS/MS analysis. Bp 6 and 7 were matched to chitinase according to Swiss-Prot database after in situ trypsin digestion and LC/MS/MS analysis.*

An aliquot of the same oat extract was subjected to purification procedures of a glycoside hydrolase. The glycoside hydrolase activity recovered from oat seedlings was purified by precipitation with ammonium sulfate, and by chromatography on a gel filtration column (Superdex-75pg) in the FPLC system. To detect starch-degrading activity, the iodine method (Jones and Varner, 1967) was used and the activity was determined by monitoring the decrease in absorbance at 700nm of the starch-iodine complex and expressed as a relative starch-degrading activity. On the Superdex-75 column, a single peak of amylase activity was detected (data not shown). After this final purification step, SDS-PAGE with Coomassie blue staining revealed that the protein preparation migrated as a single band. Two aliquots of the pooled peak from Superdex-75 elution was analyzed by SDS-PAGE followed by a Coomassie blue staining step and the two resulting bands of proteins were excised from the preparative gel (data not

shown). All the excised proteins bands from the preparative gels were digested with trypsin and analyzed by LC/ESI/MS/MS.

The amino acid sequences of these peptides were determined either by manual interpretation of the collision-induced spectra of the major peptide ion or by computer-aided fragment-matching algorithms. The majority of the protein bands excised from SDS-PAGE were identified as glycoside hydrolases (**Table 1**), some of these bands corresponding to several proteins. A high score was obtained for the match between the six studied bands (bands 4 and 5, as well as bands 6 and 7 from crude extract and, bands 8 and 9 from partial extraction procedures) and β-amylases and chitinases in the Swiss-Prot database (**Table 1**).

The bands 6, 7, 8 and 9 have, particularly, been matched to an endochitinase (fragment) from *Avena sativa* (**Table 1**). This later partial sequence of oat seed endochitinase is previously deposited in Swiss-Prot/TrEMBL under the accession number P86181.1 (Sørensen et al., 2010). Interestingly, band 9 that was produced as partly extracting protein has shown to match with a high score and high matched peptides to the *Avena sativa* endochitinase (data not shown).

de novo sequence peptides were identified for band 4, 5, 6, 7, 8 and 9 corresponding to β-amylases (Bands 4 and 5) and chitinases (Bands 6, 7, 8 and 9).

Table 1. Proteins separated by SDS-PAGE (bands 4, 5, 6, 7, 8 and 9) and directly identified by LC/ESI/MS/MS after tryptic digestion according to Swiss-Prot database

Band number on SDS-PAGE	Protein: Species origin	Score	Number of unique matched peptides	Sequence coverage (%)	Theoretical molecular mass (kDa)
4	***Beta*-amylase**: *Triticum aestivum*	284.3	7	18.5	56.6
5	***Beta*-amylase**: *Triticum aestivum*	49.6	1	3.6	56.6
6	**Endochitinase** (Fragments)/ *Avena sativa*	137.9	3	19.0	21.7
6	*Alpha*-amylase inhibitor/**endochitinase** (Fragments): *Coix lachryma-jobi*	78.6	1	12.0	14.3
6	**Chitinase**: *Oryza sativa*	51.4	2	2.9	35.6
7	**Endochitinase** (Fragments): *Avena sativa*	285.5	6	28.0	21.7
7	Alpha-amylase inhibitor/**endochitinase** (Fragments): *Coix lachryma-jobi*	56.1	1	0.0	14.3
8	**Endochitinase** (Fragments): *Avena sativa*	272.8	7	39.5	21.7
9	**Endochitinase** (Fragments): *Avena sativa*	356.4	9	44.0	21.7

The peptide sequences obtained were then used to screen for *A. sativa* EST/genomic sequences dataset (AM071411-CN180783; GAJE01000001-GAJE01050182). Interestingly, 7 genomic scaffolds (TSA_*A. sativa*: GAJE01008919.1-GAJE01008925.1) as well as 5 ESTs_*A. sativa* (GO591196.1, GO595108.1, GO595264.1, GO591125.1, and GO589909.1) corresponding to the peptide sequences of bands 4 and 5 using TBLASTN (http://blast.ncbi.nlm.nih.gov) (Altschul et al., 1997) were identified. In addition, 10 other genomic scaffolds (TSA_*A. sativa*: GAJE01021162.1-GAJE01021171.1) as well as an EST_*A. sativa* (GO586051.1) corresponding to the peptide sequences of bands 6, 7, 8 and 9 were also identified using TBLASTN (http://blast.ncbi.nlm.nih.gov) (Altschul et al., 1997). These genomic scaffolds are useful tools to the identification of 7 oat β-amylases as well as 4 oat chitinases, respectively.

We could then predict the structure of the identified genes by comparing the oat genomic scaffolds with related plant proteins (β-amylases and chitinases) using BLAST analysis (Altschul et al., 1997).

Based on these analyses, the proteins isolated from the *A. sativa* seedling extract that correspond to bands 4 and 5 were identified as β-amylases and were named AsBAMY_x (where x is the number of the predicted enzymes; in this study, we predicted at least 6 sequences of oat β-amylases [AsBAMY1 to 6] different from the previously reported AsBAMY sequence (Ben Halima et al., 2016)). Moreover, the proteins from the *A. sativa* seedling extract that correspond to bands 6, 7, 8 and 9 were identified as chitinases and were named AsChi_y (where y is the number of the predicted enzymes; in this study, we predicted 4 oat chitinases different from the deposited sequence in the GenPept database with the accession number of P86181.1).

3.1.2. Sequence analysis of oat proteins: β-amylases and chitinases

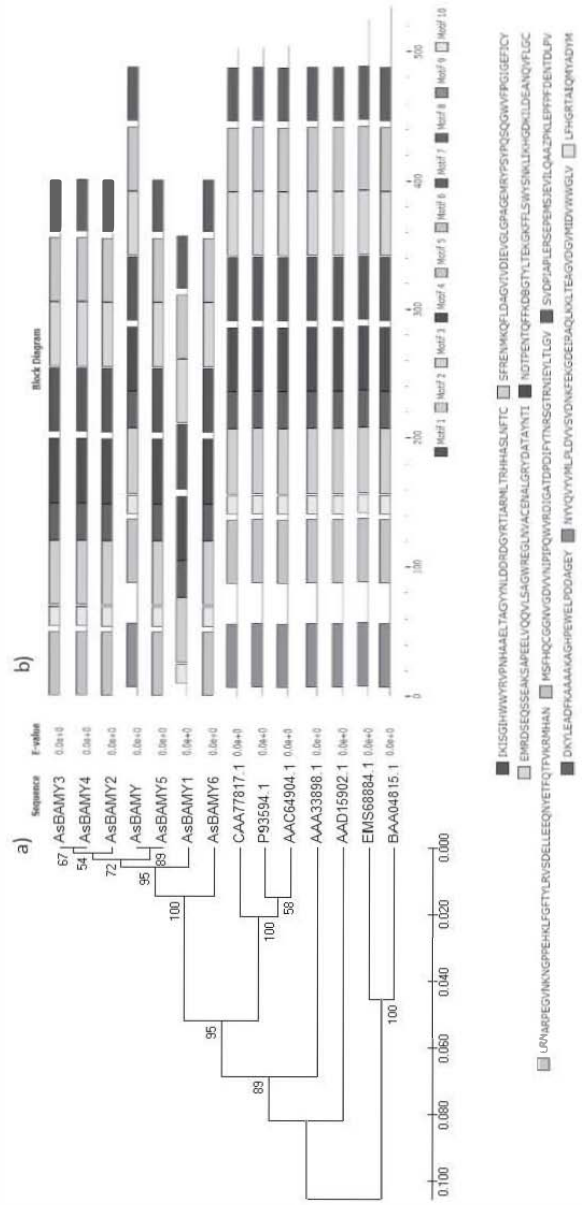
For sequence alignments of the 7 sequences of oat β-amylases, we have chosen 7 homologs in amino acid sequences, which are retrieved from the *Poaceae* family. These plant β-amylase homologs are as follows: *Triticum aestivum* (P93594.1), *Oryza sativa* (AAA33898.1), *Zea mays* (AAD15902.1), *Triticum urartu* (EMS68884.1), *Secale cereale* (CAA77817.1), *Hordeum vulgare* (BAA04815.1) and *Hordeum vulgare* (AAC64904.1). In fact, the two conserved Glu residues (E184 and E378 in AsBAMY) assigned as the "putative" catalytic residues that would act as an acid and base pair in the catalytic process are highlighted in the oat β-amylase sequences (Ben Halima, 2019). The two PROSITE consensuses of the β-amylase active

site are conserved in AsBAMY and AsBAMY1 to 6 (except for AsBAMY1 that contains only the second PROSITE consensus) and are highlighted in the oat β-amylase sequences (Ben Halima, 2019). The LC/MS/MS oat peptides matched to bands 4 and 5 are highlighted in red rectangles (Ben Halima, 2019).

GO term prediction of the highly conserved motif related to β-amylase (**Figure 2**) denoted the presence of (GO:0000272) highlighting the polysaccharide catabolic process as well as (GO:0016161) highlighting β-amylase activity (GH14 family).

Whereas for phylogenetic evolution of oat β-amylases, we have chosen homologs from monocots and eudicots groups. The monocots are those from the *Poaceae* family used in multiple sequence alignments and 3 others from the same family (*Sorghum bicolor* (XP_021310050.1), *Bromus sterilis* (CCW36754.1) and *Brachypodium distachyon* (XP_003562967.2)). The monocots non-*Poaceae* used are *Elaeis guineensis* (XP_010931493.1) and *Phoenix dactylifera* (XP_008793743.1) from the *Arecaceae* family. The eudicots used are *Glycine max* (AAZ38831.1), *Medicago truncatula* (XP_013455524.1), *Cajanus cajan* (XP_020232390.1), *Vigna angularis* (XP_017433505.1) and *Cicer arietinum* (XP_004513548.1) from the *Fabaceae* family. Indeed, the five nearest homologous (orthologous) of oat β-amylases are β-amylases from the *Poaceae* family and are as follow: *Bromus sterilis* (CCW36754.1), *Triticum aestivum* (P93594.1), *Secale cereale* (CAA77817.1), *Hordeum vulgare* (AAC64904.1) and *Brachypodium distachyon* (XP_003562967.2) (**Figure 3**).

Figure 2 (next page). Phylogenetic analysis and predicted structure of β-amylase proteins in A. sativa (AsBAMY and AsBAMY 1 to 6), T. aestivum (P93594.1), O. sativa (AAA33898.1), Z. mays (AAD15902.1), T. urartu (EMS68884.1), S. cereale (CAA77817.1), H. vulgare (BAA04815.1) and H. vulgare (AAC64904.1). ***a)*** *Evolutionary relationships of taxa related to the cereal β-amylases. The evolutionary history was inferred using the UPGMA method. The optimal tree with the sum of branch length = 0.51875970 is shown. The percentages of replicate trees in which the associated taxa clustered together in the bootstrap test (1000 replicates) are shown next to the branches. The tree is drawn to scale, with branch lengths in the same units as those of the evolutionary distances used to infer the phylogenetic tree. The analysis involved 14 amino acid sequences. There were a total of 357 positions in the final dataset.* ***b)*** *Conserved motifs of the plant β-amylases obtained by the MEME 4.11.4 software (Ben Halima, 2019).*



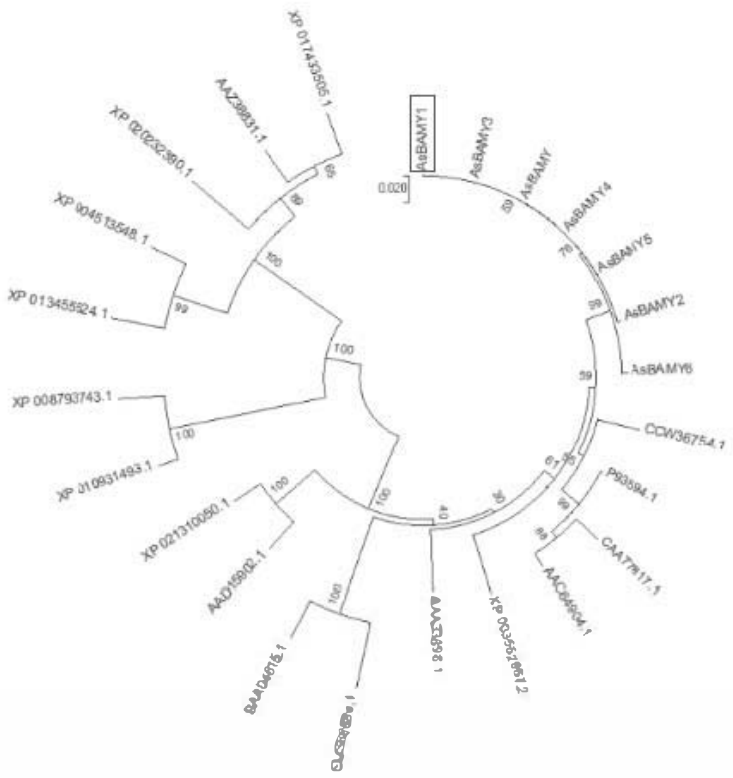


Figure 3. Evolutionary relationships of taxa related to oat β-amylases. *The evolutionary history was inferred using the Neighbor-Joining method. The optimal tree with the sum of branch length = 1.09736821 is shown. The percentages of replicate trees in which the associated taxa clustered together in the bootstrap test (1000 replicates) are shown next to the branches. The tree is drawn to scale, with branch lengths in the same units as those of the evolutionary distances used to infer the phylogenetic tree. The analysis involved 24 amino acid sequences. There were a total of 284 positions in the final dataset (Ben Halima, 2019)*

For sequence alignments of the 5 oat chitinases, we have chosen 10 homologs in amino acid, which are retrieved from monocots and especially from the *Poaceae* family like the target plant (*Avena sativa*). These plants chitinases homologues are as follow: *Avena sativa* (P86181.1), *Triticum aestivum* (AHY24793.1), *Triticum aestivum* (Q8W427), *Triticum aestivum*

(Q41539), *Hordeum vulgare* (BAJ89873.1), *Aegilops tauschii* (XP_020147158.1), *Brachypodium distachyon* (XP_003569604.1), *Secale cereale* (Q9FRV1.1), *Secale cereal* (Q9AXR9), *Zea mays* (AAT40015.1) and *Oryza sativa* (XP_015643569.1). In fact, in contrast to the previously deposited sequence of oat chitinase (P86181.1), the two PROSITE consensuses of the catalytic domain are conserved in the AsChi 1 to 4 and are highlighted in the oat chitinase sequences (data not shown). In addition, the chitin binding domain is presented in these oat chitinase sequences (AsChi1 to 4) and not in (P86181.1) with the conserved Cys residues (data not shown). The PROSITE consensus pattern for chitin bind domain located in the N-terminus of the four oat chitinases (and not in P86181) is highlighted in the oat sequences (data not shown). The LC/MS/MS oat peptides matched to bands 6, 7, 8 and 9 are highlighted in red rectangles (data not shown). One example is given in **Figure 4**, which shows the fragment ion spectrum of the double charged precursor ion $(M + 2H)^{2+}$ at m/z 858.4242 corresponding to GPIQISYNYNYGAAGK peptide.

The sequences used in this study functionally associated to chitinase activity (GH19) possessed the glutamate residue, which acts as an acid catalyst, and another glutamate residue capable of acting as a base (data not shown). All these sequences were shown to have at least a part of the highly conserved motif [FHY]-G-R-G-[AP]-x-Q-[IL]-[ST]-[FHYW]-[HN]-[FY]-NY (Huet et al., 2008). Obviously, motif 3 (**Figure 5, Figure 6**) contained this part of the highly conserved motif ('FGRGPIQISYNYNY') found to be functionally associated with GH19 chitinase superfamily proteins. The above-mentioned motif forms the substrate-binding region of GH19 proteins. Thus, the sequences used in this dataset hit the criteria for GH19 proteins that contained both the catalytic and the substrate binding regions.

GO term prediction of the highly conserved motif related to chitinase (**Figure 5**) denoted the presence of (GO:0006032) for chitin catabolic process and (GO:0016998) for cell wall macromolecule catabolic process, as well as (GO:0004568) for chitinase activity (GH19 family).

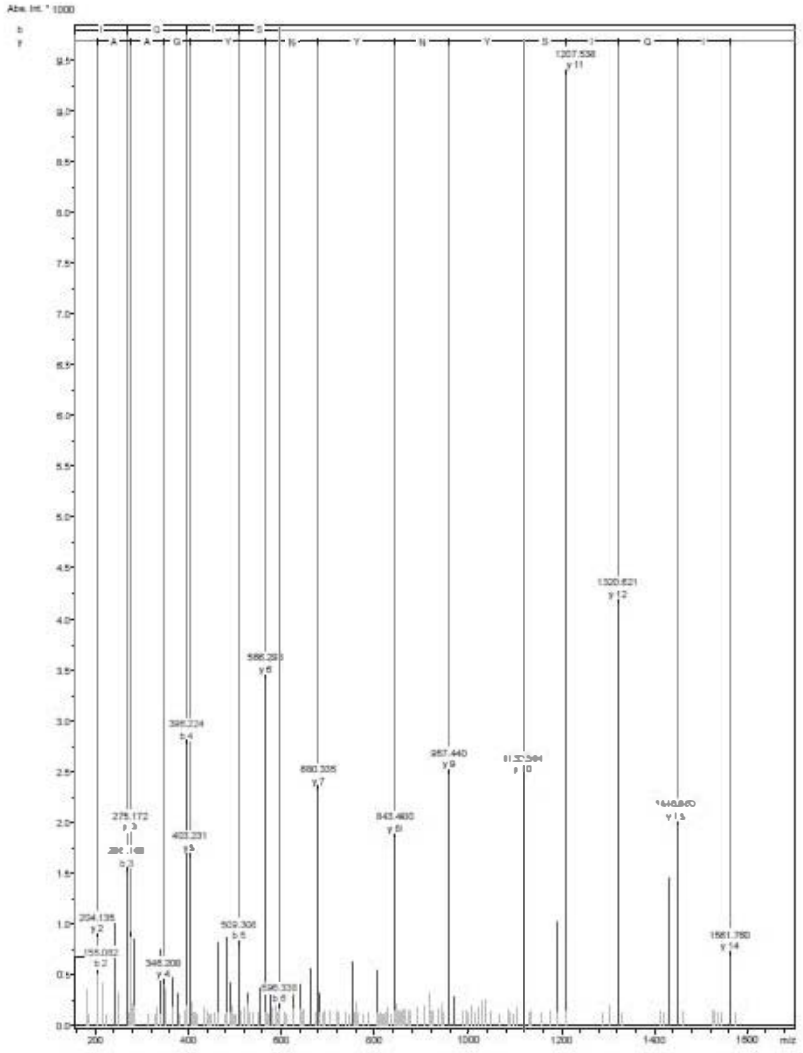


Figure 4. *MS/MS (fragment ion) spectrum of the double charged precursor ion (M + 2H)$^{2+}$ at m/z 858.4242 corresponding to the GPIQISYNYNYGAAGK peptide.*

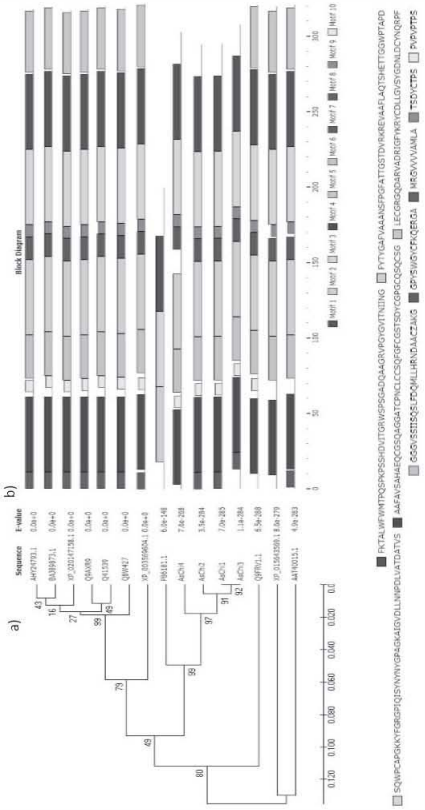


Figure 5. Phylogenetic analysis and predicted structure of chitinase proteins in A. sativa (AsChi 1 to 4 & P86181.1), T. aestivum (AHY24793.1), T. aestivum (Q8W427), T. aestivum (Q41539), H. vulgare (BAJ89873.1), A. tauschii (XP_020147158.1), B. distachyon (XP_003569604.1), S. cereale (Q9FRV1.1), S. cereal (Q9AXR9), Z. mays (AAT40015.1) and O. sativa (XP_015643569.1). a) *Evolutionary relationships of taxa related to the cereal chitinases. The evolutionary history was inferred using the UPGMA method. The optimal tree with the sum of branch length = 0.80359393 is shown. The percentages of replicate trees in which the associated taxa clustered together in the bootstrap test (1000 replicates) are shown next to the branches. The tree is drawn to scale, with branch lengths in the same units as those of the evolutionary distances used to infer the phylogenetic tree. The analysis involved 15 amino acid sequences. There were a total of 179 positions in the final dataset.* ***b)*** *Conserved motifs of the cereal chitinases obtained by the MEME 4.11.4 software. The 1, 2 and 3 motifs were found to be the highly conserved motifs among the tested proteins functionally associated with chitinase activity.*

Figure 6. *The conserved sequence motif 3 (represented as web logos) present across the GH19 chitinase family proteins.*

Distribution of GO terms in the biological process category in oat chitinase could also reveal the (GO:0005975) for carbohydrate metabolic process.

Distribution of GO terms in the molecular function category in chitinase could also reveal the (GO:0008061) for chitin binding.

Moreover, for phylogenetic evolution of oat chitinases, we have chosen 25 homologs from monocots and eudicots groups. The monocots are those from the *Poaceae* family used in multiple sequence alignments and 3 others from the same family (*Bromus inermis* (BAG12896.1), *Festuca arundinacea* (ACJ23248.1) and *Poa pratensis* (AAF04454.1)). The monocots non-*Poaceae* used are *Elaeis guineensis* (XP 010941404.1) and *Phoenix dactylifera* (XP 008812110.1) from the *Arecaceae* family. The eudicots used are from 5 different origin species: *Theobroma cacao* (XP 007046549.2) and *Gossypium raimondii* (XP 012452524.1) from the *Malvaceae* family, *Capsicum annuum* (XP 016560402.1) and *Solanum tuberosum* (NP 001305536.1) from the *Solanaceae* family, and *Carica papaya* (3CQL A) from the *Caricaceae* family. Indeed, the nearest homologous (orthologous) of oat chitinase sequences are chitinases from the *Poaceae* family and are as follow: *Poa pratensis* (AAF04454.1), *Festuca arundinacea* (ACJ23248.1), *Bromus inermis* (BAG12896.1), *Triticum aestivum* (Q41539), *Secale cereal* (Q9AXR9), *Aegilops tauschii* (XP 020147158.1), *Hordeum vulgare* (BAJ89873.1), *Triticum aestivum* (AHY24793.1) and *Triticum aestivum* (Q8W427) (**Figure 7**).

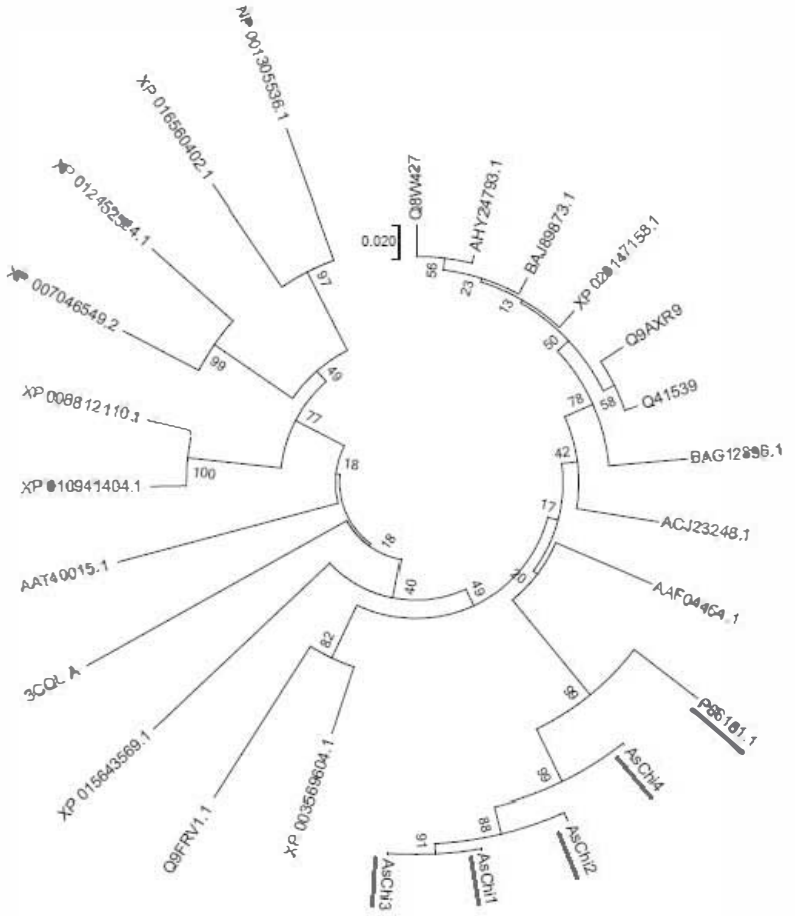


Figure 7. Evolutionary relationships of taxa related to oat chitinases. *The evolutionary history was inferred using the Neighbor-Joining method. The optimal tree with the sum of branch length = 1.62384419 is shown. The percentages of replicate trees in which the associated taxa clustered together in the bootstrap test (1000 replicates) are shown next to the branches. The tree is drawn to scale, with branch lengths in the same units as those of the evolutionary distances used to infer the phylogenetic tree. The analysis involved 25 amino acid sequences. There were a total of 174 positions in the final dataset.*

Two sequences from each of the two predicted types of oat proteins, i.e., β-amylases and chitinases (AsBAMY and AsBAMY1 for oat β-amylases

and P86181.1 and AsChi1 for oat chitinases) were selected to further enlighten the results and interpretations.

3.1.3. Structural features of the selected oat proteins

The predicted AsBAMY cDNA (1464 bp), which corresponds to a 488 amino acid residue protein has already been presented in our previous report (Ben Halima et al., 2016).

The predicted AsBAMY1 cDNA (1071 bp) corresponds to a 357 amino acid residues protein (**Figure 8**) of a theoretical molecular weight of the protein of 41 kDa and with a theoretical isoelectric point (pI) of 4.84.

Figure 8. *Open reading frame and amino acid sequences of AsBAMY1. The amino acid sequence is shown in single code letters beneath the cDNA sequence. Bold italics mark the start codon (ATG) and an asterisk indicates the translation stop codon (TAA). Amino acid sequences obtained by mass spectrometry are shown in bold red.*

The previously deposited sequence of oat chitinase (P86181.1) that corresponds to 200 amino acid residues has already been reported in the study of Sørensen et al. (2010).

The predicted **AsChi1** cDNA (885 bp) corresponds to a 295 amino acid residues protein (**Figure 9**) of a theoretical molecular weight of the protein of 31 kDa and with a theoretical isoelectric point (pI) of 9.10.

Figure 9. *Open reading frame and amino acid sequences of AsChi1. The amino acid sequence is shown in single code letters beneath the cDNA sequence. Bold italics mark the start codon (ATG) and an asterisk indicates the translation stop codon (TGA). Amino acid sequences obtained by mass spectrometry are shown in bold red. The putative signal peptide is underlined, with the cleavage site, AHA/QQ indicated by an arrow.*

The subcellular localization of AsBAMY and AsBAMY1 is mainly cytoplasmic with a reliability of 3.105 and 3.485, respectively (data not shown); while the subcellular localization of oat chitinase_P86181.1 and AsChi1 is mainly extracellular with a reliability of 2.606 and 3.844, respectively (data not shown).

A search against the Conserved Domain Database (CDD) (Marchler-Bauer et al., 2017) (http://www.ncbi.nlm.nih.gov/Structure/cdd/cdd.shtml), revealed that the deposited amino acid sequence of oat chitinase (P86181.1) possesses a conserved domain highly homolog (E-value: 7.30e-117) to Glyco_hydro_19 superfamily (accession cl27735) that described the chitinase class I. The four newly identified oat chitinases possess this conserved domain (accession cl27735) with another accession (pfam00187) relative to chitin recognition protein. Herein, for instance, AsChi1 possesses the highly conserved domain, which is highly homolog (E-value: 9.43e-133) to the Glyco_hydro_19 superfamily (accession cl27735) and the second domain with accession (pfam00187) homolog (E-value: 4.32e-20) to chitin_bind_1 (**Figure 10**).

HMMER (https://www.ebi.ac.uk/Tools/hmmer) and Pfam (http://pfam.xfam.org) analysis revealed that oat β-amylases possess a conserved domain relative to Glyco-hydro_14 (accession: PF01373.16). Herein, AsBAMY and AsBAMY1 possess this conserved domain highly homolog (E-value: 1e-120 and E-value: 6e-74, respectively) to GH14 family (**Fig. 10**).

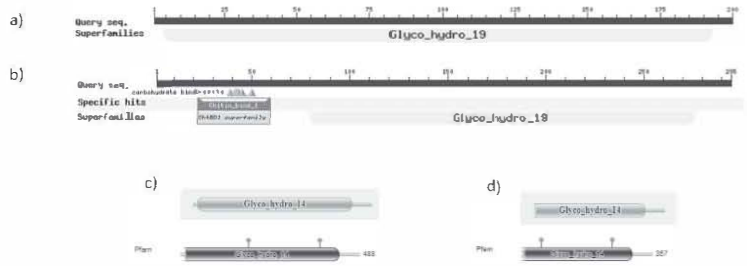


Figure 10. Putative conserved domains in oat chitinases (P86181.1) (a) and (AsChi1) (b) as detected by the Conserved Domain Database (http://www.ncbi.nlm.nih.gov/Structure/cdd/cdd.shtml) and those detected by Pfam (http://pfam.xfam.org/) for oat β-amylase (AsBAMY) (c) and (AsBAMY1) (d).

A 20-residue signal peptide was predicted using the Expasy SignalP V4.1 program and the N-terminal sequence of the mature AsChi1 is expected to start at residue Q21. Whereas, no signal peptide residues were predicted in P86181.1, neither in AsBAMY nor in AsBAMY1 using the Expasy SignalP V4.1 program.

The NetNGlyc 1.0 Server predicted four possible N-glycosylation sites at residues 123–126 (NRSG), 237–240 (NDTP), 338–341 (NFTC) and 442–445 (NQSH) in AsBAMY, and three possible N-glycosylation sites at residues 106-109 (NDTP), 207-210 (NFTC) and 311-314 (NQSH) in AsBAMY1. Whereas, no sites predicted of N-glycosylation in both oat chitinases (P86181.1 and AsChi1) using the NetNGlyc 1.0 Server.

The NetOGlyc 4.0 Server predicted two possible O-glycosylation sites at residues 115 and 448 in AsBAMY and at residues 313 and 317 in AsBAMY1. 16 possible O-glycosylation sites are predicted in AsChi1 at residues 25, 57, 60, 62, 63, 66, 68, 70, 120, 172, 241, 251, 253, 254, 279 and 285; whereas, 7 possible O-glycosylation sites at residues 37, 65, 135, 141, 145, 147 and 148 in P86181.1 using the NetOGlyc 4.0 Server.

Oat proteins (β-amylases and chitinases) were accessed by predicting their secondary structures using SOPMA server software and PSIPRED online server.

These proteins show a large proportion of alpha helix and random coils. The oat chitinases proteins possessed a high percentage of alpha helix (27.50 % and 24.75%) and random coils (44.00% and 53.90%) for P86181.1 and AsChi1, respectively. Oat β-amylases also possessed a high percentage of alpha helix (38.93% and 39.50%) and random coils (34.63% and 36.13%) for AsBAMY and AsBAMY1, respectively.

Moreover, the secondary structure of oat proteins was analyzed by PSIPRED online server and showed that P86181.1 presents 8 helices, 1 stranded-sheet, and 10 coils, whereas AsChi1 presents 8 helices, 4 stranded-sheet and 13 coils. The secondary structure of AsBAMY showed 11 helices, 12 stranded-sheet and 24 coils and that of AsBAMY1 showed 9 helices, 6 stranded-sheet, and 16 coils.

The Swiss-Model server was used to predict the 3D structure of oat proteins based on known crystal structures of homologous proteins (**Figure 11**). The lack of a 3D structure for the majority of proteins from *Avena sativa* in PDB motivated us to construct the 3D model for each of the studied proteins. The most successful techniques for prediction of three-dimensional structures of proteins rely on aligning the sequence of a protein to a homolog of a known structure.

The crystal structure of barley β-amylase complexed with 2,3-epoxypropyl-alpha-D-glucopyranoside (PDB ID: 2XGB.1.A) was used as the template for 3D structure prediction of oat β-amylases (AsBAMY and

AsBAMY1) with 81.76% and 81.97% sequence identity, respectively (**Figure 11**).

The highest scoring and validated models for oat chitinases (P86181.1 and AsChi1) exhibit the greatest amino acid sequence identity with the crystal structure of a family GH-19 chitinase from rye seeds (PDB ID: 4DWX.1.A) (**Figure 11**). This template protein is 76.50% and 77.25% identical to P86181.1 and AsChi1, respectively (data not shown).

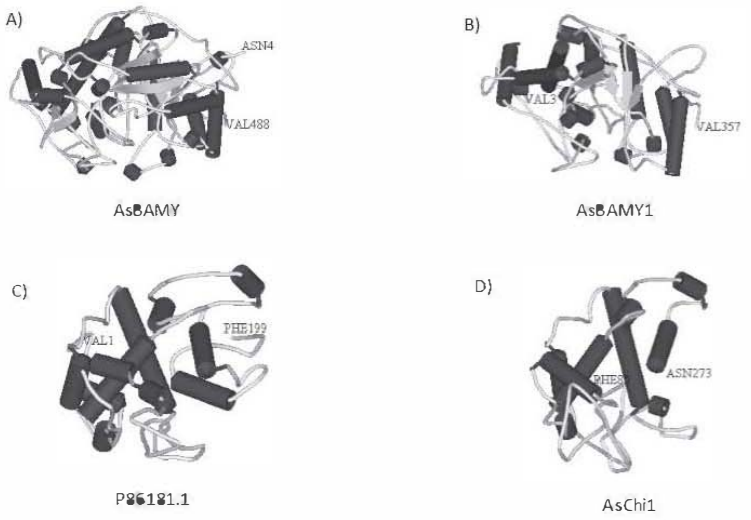


Figure 11. *3D structure prediction of GH proteins from Avena sativa (A, B, C, D). The N-terminal and the C-terminal of each 3D structure prediction were designed by corresponding amino acid. (red: Helix; blue: Sheet; green: Turn; grey: Coil).*

The secondary structures of the studied oat proteins were in agreement with the related 3D-structures, which revealed abundant alpha-helical structures in both oat β-amylases and chitinases.

3.2. Discussion

Identification of novel genes, especially in organisms with unknown genomic data, would be promisingly done by genomic approaches; among these approaches are global transcriptomic analysis and the production of expressed sequence tags (ESTs). ESTs technology, which has been proven to be a relatively cost-effective, efficient, rapid and powerful tool for

identifying and discovering novel genes, is also a commonly used strategy in identifying genes involved in specific functions (Gueguen et al., 2003).

The Transcriptome Shotgun Assembly (TSA) or the ESTs databases analyses provide valuable tools for elucidating information regarding gene expression as well as an opportunity to identify new genes involved in biological function. These databases have been developed in many glycophytic plant species and are time-consuming and laborious methods to analyze gene function in plants as well as for comparative genomics.

Our previous studies involved information on these genomic and ESTs approaches that have been successfully used to identify genes encoding related proteins with catalytic activity including β-amylase (Ben Halima et al., 2016) or lipolytic activities (Abdelkafi et al., 2009, 2012). The functional proteomics is also an important clue to identify such proteins. A tutorial made by Cottrell (2011) explained some features on the identification of proteins using MS/MS data.

The continuing initiative to find novel plant carbohydrate-active enzymes (CAZymes) by such functional proteomics and genomic approaches is very interesting for the valorization of plant biomass as a substrate for various products in many areas, e.g., food and medicine.

In the present study, seedling extract from the glycophytic oat (*Avena sativa*) was proven to be a potential source of proteins from GH19 and GH14 family and the focus was on the highly abundant class I chitinases (GH19) and β-amylases (GH14) using the strategies of functional proteomics and genomic approaches.

Although in our previous report, we have identified one oat β-amylase by mainly ESTs approach (Ben Halima et al., 2016), in the current study, we have identified 6 other oat β-amylases by the TSA approach. Thus, abundant catalytically active β-amylases were characterized from oat (*A. sativa* L.) seedlings extract. Biochemical analysis and *in silico* approach indicated that these GH14 enzymes were identical to previously characterized plant β-amylases with regard to its amino-acid sequence, catalytic activities, and molecular structure. Indeed, zymogram analysis of the oat seedling extract revealed at least 2 clear bands including bands around 42 and 53 kDa indicating enzymatic activity on starch used as a substrate (Ben Halima et al., 2016). The SDS-PAGE analysis of the corresponding fractions allowed the identification of seven major protein bands (**Figure 1**). After tryptic digestion and LC/ESI/MS/MS, the majority

of these protein bands, especially Bp 4, 5, 6, 7, 8 and 9 were identified as glycoside hydrolases according to the Swiss-Prot database (**Table 1**).

Band 4 (53 kDa) has yet been characterized as containing a β-amylase with seven *de novo* peptide sequencing (Ben Halima et al., 2016). Several other β-amylase sequences may be associated with band 4 with the same approach of the seven *de novo* peptide sequence hits that were used to screen *A. sativa* ESTs and TSAs. Based upon TSA and EST sequences as well as sequence homologies with other plant β-amylases, 7 sequences of cDNA could be constructed and encoded 7 different oat β-amylases (AsBAMY and AsBAMY1 to 6). In fact, the 7 resulted peptides obtained after in-gel tryptic digestion and LC/MS/MS analysis of band 4 all matched to β-amylase from *Triticum aestivum* as revealed by the Swiss-Prot database and are all conserved in AsBAMY, AsBAMY3, AsBAMY4, AsBAMY4, and AsBAMY6. These peptides are also conserved in AsBAMY2 with only one amino acid difference (Ben Halima, 2019). Interestingly, AsBAMY1 would, very likely, be associated with band 5 due to the correspondence of molecular weight and the one matched peptide (Ben Halima, 2019). Indeed, the theoretical molecular weight of AsBAMY1 is 41 kDa; and since the apparent molecular weight for band 5 estimated from SDS-PAGE was approximately 42 kDa, AsBAMY1 might be glycosylated (three N-glycosylation sites and two possible O-glycosylation sites are predicted from sequence analysis), but other post-translational modifications might also be involved and related to AsBAMY.

The majority of cereal β-amylases were reported to have molecular weights in the range of 53–64 kDa (Ziegler, 1999). The β-amylases from band 4 correlate well with this fact. Moreover, Wu et al. (2011) observed a similar size of approximately 56 kDa for β-amylase isozymes of Australian wheat.

The amino acid sequence of oat β-amylases (AsBAMY and AsBAMY1) deduced from the nucleotide sequences of cDNA did not have signal peptide; suggesting that these enzymes would probably be synthesized in the cytosol as mature protein (Ziegler, 1999). Moreover, the subcellular localization of these β-amylases was predicted to be cytoplasmic as revealed by the CELLO v.2.5 server (Ben Halima, 2019).

In all the tested plant β-amylases, sequence alignment has suggested the presence of the two Glu conserved catalytic residues in position 184 and 378 in AsBAMY (Ben Halima et al., 2016, Ben Halima, 2019). Indeed,

three highly conserved sequence regions would be found in all β-amylases. The first of these regions contains an aspartate, which would be involved in the catalytic mechanism. This region is located in the N-terminal sequence of the enzymes. The other regions would be centered around the conserved glutamate residues (E184 and E378 in AsBAMY) involved in the catalytic mechanism. Signature patterns form PROSITE used of the above regions as follows:

The first pattern (PS00506) is: H-x-C-G-G-N-V-G-D (where D is an active site residue) and the second pattern (PS00679) is: G-x-[SA]-G-E-[LIVM]-R-Y-P-S-Y (where E is an active site residue). These two PROSITE patterns are conserved in all the β-amylases aligned in this study with the exception of AsBAMY1, which lacks only the first pattern (Ben Halima, 2019).

The catalytic activity of soluble oat (*Avena sativa*) extract has been known to have potential versatile applications, especially in bread making (Ben Halima et al., 2015b) and anti-fungi spoilage bread (Sørensen et al., 2010). The important catalytic activity for cereal germination is, without doubt, amylolytic activity since cereals have a considerable amount of starch in their endosperm. Thus, β-amylases with the different isoforms are among the amylolytic activities of oats, which may be required during oat metabolism, germination and growth. It is worth noting that β-amylases have different structure and function in cereals, for instance, there are two distinct categories of cereal β-amylases (the 'endosperm-specific' and tissue ubiquitous' types of β-amylases) which reinforce the complex roles and features of these enzymes during cereal seed storage or germination (Ziegler, 1999).

In addition to the abundance of β-amylases in 10-day old oat seedling extract, chitinases are also abundant in this oat extract. The importance of this extract in the conservation of bread was proven in the report of Ben Halima et al. (2015b). They showed that the use of oat extract as an additive during bread making increased the conservation days. This, in addition to the amylase activity, maybe due to the abundance of chitinases in this oat extract, which are involved in defense against pathogens such as fungi and insects by destroying their chitin-containing cell wall.

Chitinase from GH19 (class I or class II) are enzymes involved in the hydrolysis of β-1,4- linked polysaccharides. Unlike class II chitinases, plant class I chitinases have a cysteine-rich N-terminal chitin-binding domain.

Several other studies reported the characterization of other plant chitinases such as those from *Limonium bicolor*, which are successfully expressed in a heterologous system that exhibit recombinant chitinase activities (Liu et al., 2013). *In silico* identification of coffee genome expressed sequences by the report of Mazzinghy Alvarenga et al. (2010) state coffee chitinases as potentially associated with resistance to diseases.

A possible mechanism of antifungal activity was suggested for chitinases in the report of Landim et al. (2017) who reported biochemical and structural features of a class I chitinase from cowpea (*Vigna unguiculata*) as well its hydrolytic action.

The study of Udaya Prakash et al. (2010) emphasis estimation of the pattern of evolution between bacteria and plant chitinases. They support the horizontal gene transfer theory, which states that GH19 chitinase genes are transferred from higher plants to bacteria (Udaya Prakash et al., 2010). In our point of view, we eliminate such transfer theory, as we believe that there is no common ancestor in the three major superkingdoms of life.

Oat seedlings of 10-day old extract is also enriched in chitinases activity as revealed by SDS-PAGE (bands 6, 7, 8 and 9) (Fig. 1, Table 1). Obviously, after purification procedure by ammonium sulfate and gel filtration (Superdex 75) of the oat extract, and instead of obtaining amylases from the purified fraction, the most significant match with a higher score was found with endochitinase *Avena sativa* (Accession no. P86181.1). This confirmed the high abundance of chitinases in oat seedling extract. The peptides matched are shown in bold red in **Figure 12**.

Further, protein sequence of the fragment oat endochitinase (P86181.1) was retrieved from the NCBI database as LC/ESI/MS/MS-based peptide mass fingerprint of our oat chitinases (Band 9) (MLLHR, SFPAFATTGSTDVR, GPIQISYNYNYGAAGK, AIGVDLLR, TALWFWMTPQSPKPSSHDVITGR, WSPSSTDK, GQESHVADR and IGYYK) were found to be conserved. In fact, 4 new sequences of oat chitinases were identified in this study that could be referred to either band 6 or band 7 as they didn't conserve all the 8 matched peptides (Figure 1; Table 1).

Cmpd.	No. of Cmpds.	m/z meas.	Δ m/z [ppm]	z	Rt [min]	Score	P	Range	Sequence	Modification
22	1	336.1969	-0.05	2	6.9	24.7	0	13-17	K.MLLHR.G	
9	1	343.1954	2.99	2	6.5	23.2	●	13-17	K.MLLHR.G	Oxidation: 1
114	1	486.2417	2.63	3	8.9	78.6	0	31-44	K.SFPAFATTGSTDVR.K	
104	1	858.4242	1.42	2	8.7	82.1	0	83-98	R.GPIQISYNYYGAAGK.A	
121	1	428.7663	0.24	2	8.9	37.1	0	99-106	K.AIGVDLLR.N	
126	2	665.3326	-1.81	4	9.0	42.4	0	121-143	K.TALWFWMTPQSPKPSSHDVITGR.W	Oxidation: 7
24	1	454.2124	2.20	2	7.0	25.1	0	144-151	R.WSPSSTDK.A	
3	1	499.7373	2.33	2	6.1	22.3	0	175-183	K.GQESHVADR.I	
32	1	322.1762	0.21	2	7.1	21.0	0	184-188	R.IGYYK.D	

Figure 12. *9 oat peptides from Bp 9 matched to chitinase from Avena sativa as revealed by the Swiss-Prot database. The conserved peptides matched on the basis of LC/ESI/MS/MS analysis of oat protein are shown in red.*

As detected by searches against the CDD, AsChi1 contains a type 1 chitin binding domain (ChBD1, pfam00187) and a GH19 catalytic domain (CatD, accession cl27735). The primary structure of the chitin-binding domain of AsChi1 ($AsChi1_{ChBD}$) contains 8 Cys residues in the same positions as those found in the alignment plant chitinase sequences (data not shown).

A central segment of $AsChi1_{ChBD}$ (^{32}CPNSLCCSQYGFCGSTNDYC^{51}) follows the consensus pattern C-x(4,5)-C-C-S-x(2)-G-x-C-G-x(3,4)-[FYW]-C (where the 5 C's are probably involved in disulfide bonds), which is the PROSITE signature for the ChtBD1 (PROSITE_PS00026).

Moreover, when the $AsChi1_{ChBD}$ amino acid sequence was aligned with the corresponding other plant chitinases, the 7 residues that presumptively compose its chitin-binding site were mapped. The residues are as follows: Ser^{39}, Tyr^{41}, Gly^{42}, Phe^{43}, Gly^{45}, Asp^{49} and Tyr^{50} which, are similar to the other aligned sequences (data not shown).

One stretch of amino acids within the AsChi1 sequence (^{97}CEAKGFYTYNAFLAAAKSFPAFA^{119}) matches the PROSITE consensus pattern 1 (PS00773) of the GH19 chitinases, C-x(4,5)-F-Y-[ST]-x(3)-[FY]-[LIVMF]-x-A-x(3)-[YF]-x(2)-F-[GSA]. A second segment of the primary structure of AsChi1 (^{223}VAFKTALWFWM^{233}) follows the PROSITE signature 2 (PS00774) of the GH19 chitinases, [LIVM]-[GSA]-F-x-[STAG](2)-[LIVMFY]-W-[FY]-W-[LIVM] (data not shown). However, the previous deposited sequence (P86181.1) had one difference in the PROSITE signature 2 (E118 vs A223 in AsChi1) (data not shown).

Besides, searches against the CDD allowed the identification of the AsChi1 residues presumed to be involved in catalysis (Glu141, Glu163, and Ser194) and sugar binding (Gln192, Tyr197, Asn198, Asn272 and Pro285) (**Figure 10b and data not shown**). Most of these 8 residues are conserved in other true GH19 chitinases from different sources (data not shown).

Altogether, these sequence analyses suggested that oat chitinases are likely functional enzymes, capable to bind and hydrolyze chitin.

Although several plant chitinases and β-amylase enzymes have been isolated, cloned and characterized, the knowledge on these enzyme families is still limited.

The results obtained here on the identification and biochemical properties of glycoside hydrolases from family 19 and 14 from *A. sativa* 10 day old extract are a further step in the characterization of these enzymes in plants. The physiological role of such enzymes remains, however, to be elucidated. The complete sequencing of the *A. sativa* genome will certainly accelerate the identification of other catalytic activities from *A. sativa* with applications in biotechnology.

4. Conclusions

Proteins from GH14 (*β*-amylases) and GH19 (chitinases) family were *de novo* sequenced from oat (*Avena sativa*) 10-day old seedling extracts. An attempt of good characterization of these proteins has been made in this study, providing insights into the importance of such enzymes in oat metabolism, defense against pathogens as well as biotechnological uses of these catalytic activities.

5. Acknowledgments

The author thanks the Tunisian Ministry of Higher Education and Scientific Research for facilities. The author would like to express gratitude to Dr. Guillaume Gabant from the "Plateforme de Spectrométrie de Masse et Protéomique du Centre de Biophysique Moléculaire" (Orleans, France) for mass spectrometry analysis as well as for his precious discussion. The author is very grateful to Erasmus Nchuaji Tang, a researcher at the Department of Biochemistry, University of Yaounde I-Cameroon, for his help reading the manuscript.

6. References

Abdelkafi, S., Ogata, H., Barouh, N., Fouquet, B., Lebrun, R., Pina, M., Scheirlinckx, F., Villeneuve, P., & Carrière, F. (2009). Identification and biochemical characterization of a GDSL motif carboxylester hydrolase from *Carica papaya* latex. Biochim. Biophys. Acta 1791, 1048–1056.

Abdelkafi, S., Abousalham, A., Fendri, I., Ogata, H., Barouh, N., Fouquet, B., Scheirlinckx, F., Villeneuve, P., & Carrière, F. (2012). Identification of a new phospholipase D in *Carica papaya* latex. Gene 499, 243–249.

Altschul, S.F., Madden, T.L., Schaffer, A.A., Zhang, J., Zhang, Z., Miller, W., & Lipman, D.J. (1997). Gapped BLAST and PSI-BLAST: a new

generation of protein database search programs. Nucleic Acids Res. 25, 3389–3402.

Altschul, S.F., Wootton, J.C., Gertz, E.M., Agarwala, R., Morgulis, A., Schaffer, A.A., & Yu, Y.K. (2005). Protein database searches using compositionally adjusted substitution matrices. FEBS J. 272, 5101–5109.

Anderson, O.D. (2014). The spectrum of major seed storage genes and proteins in oats (*Avena sativa*). PLoS One 9, e83569.

Arimori, T., Kawamoto, N., Shinya, S., Okazaki, N., Nakazawa, M., Miyatake, K., Fukamizo, T., Ueda, M., & Tamada, T. (2013). Crystal structures of the catalytic domain of a novel glycohydrolase family 23 chitinase from *Ralstonia* sp. A-471 reveals a unique arrangement of the catalytic residues for inverting chitin hydrolysis. J. Biol. Chem. 288, 18696-18706, http://dx.doi.org/10.1074/jbc.M113.462135.

Bailey, T.L., & Elkan, C. (1994). Fitting a mixture model by expectation maximization to discover motifs in biopolymers", *Proceedings of the Second International Conference on Intelligent Systems for Molecular Biology*, pp. 28-36, AAAI Press, Menlo Park, California.

Bailey, T.L., & Gribskov, M. (1998). Combining evidence using p-values: Application to sequence homology searches. Bioinformatics (Oxford, England). 14, 48–54.

Bailey, T.L., Johnson, J., Grant, C.E., & Noble, W.S. (2015). The MEME Suite. Nucleic Acids Res. 43(W1):W39–49.

Bairoch, A., Apweiler, R., Wu, C.H., Barker, W.C., Boeckmann, B., Ferro, S., Gasteiger, E., Huang, H., Lopez, R., Magrane, M., Martin, M.J., Natale, D.A., O'Donovan, C., Redaschi, N., & Yeh, L.S. (2005). The universal protein resource (UniProt). Nucleic Acids Res. 33, 154–159.

Beattie, A., Eckstein, P., Scoles, G., & Rossnagel, B. (2008). An oat (Avena sativa L.) cDNA derived from three developmental stages, 8th International Oat Conference, Minneapolis, MN, Available: http://wheat.pw.usda.gov/GG2/Avena/event/IOC2008/IOCposter/VI-1_Beattie.pdf.

Beaufour, M., Godin, F., Vallée, B., Cadene, M., & Bénédetti, H. (2012). Interaction proteomics suggests a new role for the Tfs1 protein in yeast. J. Proteome Res. 11, 3211–3218.

Ben Halima, N. (2019). Prediction and analysis of GH14 family β-amylases in oat seedling extract: structure and function insights using *in silico* approaches. International Journal of Biological Macromolecules. 125, 361-369.

Ben Halima, N., Ben Saad, R., Khemakhem, B., Fendri, I., & Abdelkafi, S. (2015a). Oat (*Avena sativa* L.): Oil and nutriment compounds valorization for potential use in industrial applications. J. Oleo Sci. 64, 915-932.

Ben Halima, N., Borchani, M., Fendri, I., Khemakhem, B., Gosset, D., Baril, P., Pichon, C., Ayadi, M.A., & Abdelkafi, S. (2015b). Optimized amylases extraction from oat seeds and its impact on bread properties. Int. J. Biol. Macromol. 72, 1213–1221.

Ben Halima N., Khemakhem B., Fendri I., Ogata H., Baril P., Pichon C., & Abdelkafi, S. (2016). Identification of a new oat β-amylase by functional proteomics. Biochimica et Biophysica Acta. 1864, 52–61.

Bourne, Y., & Henrissat, B. (2001). Glycoside hydrolases and glycosyltransferases: families and functional modules, Curr. Opin. Struct. Biol. 11, 593–600.

Cantarel, B.L., Coutinho, P.M., Rancurel, C., Bernard, T., Lombard, V., & Henrissat, B. (2009). The carbohydrate-active EnZymes database (CAZy): an expert resource for glycogenomics, Nucleic Acids Res. 37, D233–D238.

Cottrell, J.S. (2011). Protein identification using MS/MS data. J. Proteomics 74, 1842 – 1851.

Daussant, J., Sadowski, J., & Ziegler, P. (1994). Cereal β-amylases: diversity of the β-amylase isozyme status within cereals. J. Plant Physiol. 143, 585 - 590.

Edgar, R.C. (2004). MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res. 32, 1792–7.

Felsenstein, J. (1985). Confidence limits on phylogenies: An approach using the bootstrap. Evolution 39, 783-791.

Fendri, I., Tardif, C., Fierobe, H.P., Lignon, S., Valette, O., Page, S., & Perret, S. (2009). The cellulosomes from *Clostridium cellulolyticum*: identification of new components and synergies between complexes. FEBS J. 276, 3076–3086.

Filtenborg, O., Frisvad, J. C., & Thrane, U. (1996). International Journal of Food Microbiology. 33, 85–102. doi:10.1016/0168-1605(96)01153-1.

Fink, W., Liefland, M., & Mendgen, K. (1988). Chitinases and β-1,3-Glucanases in the apoplastic compartment of oat leaves (*Avena sativa* L.). Plant Physiol. 88, 270-275.

Finn, R.D., Bateman, A., Clements, J., Coggill, P., Eberhardt, R.Y., Eddy, S.R., Heger, A., Hetherington, K., Holm, L., Mistry, J., et al. (2014). Pfam: the protein families database. Nucleic Acids Res. 42(Database issue): D222–230.

Fujita, K., Shimomura, K., Yamamoto, K.-I., Yamashita, T., & Suzuki, K. (2006). A chitinase structurally related to the glycoside hydrolase family 48 is indispensable for the hormonally induced diapause termination in a beetle. Biochem. Biophys. Res. Commun. 345, 502-507, http://dx.doi.org/10.1016/j.bbrc.2006.04.126.

Fulton, D.C., Stettler, M., Mettler, T., Vaughan, C.K., Li, J., Francisco, P., Gil, M., Reinhold, H., Eicke, S., Messerli, G., Dorken, G., Halliday, K., Smith, A.M., Smith, S.M., & Zeeman, S.C. (2008). Beta-AMYLASE4, a noncatalytic protein required for starch breakdown, acts upstream of three active *beta*-amylases in *Arabidopsis* chloroplasts. Plant Cell 20, 1040–1058.

Gasteiger, E., Gattiker, A., Hoogland, C., Ivanyi, I., Appel, R.D., & Bairoch, A. (2003). ExPASy: the proteomics server for in-depth protein knowledge and analysis. Nucleic Acids Res. 31, 3784-3788.

Geourjon, C., & Deleage, G., (1995). SOPMA: significant improvements in protein secondary structure prediction by consensus prediction from multiple alignments. Computer applications in the biosciences: CABIOS 11, 681-684.

Gueguen, Y., Cadoret, J.P., Flament, D., Barreau-Roumiguie`re, C., Girardot, A.L., Gamier, J., Hoareau, A., Bache`re, E., & Escoubas, J.M. (2003). Immune gene discovery by expressed sequence tags generated from hemocytes of the bacteria-challenged oyster, *Crassostrea gigas*. Gene 303, 139–145.

Hall, T.A. (1999). BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nucleic Acids Symp. Ser. 41, 95-98.

Henrissat, B. (1991). A classification of glycosyl hydrolases based on amino acid sequence similarities. Biochem. J. 280, 309-316.

Henrissat, B., & Bairoch, A. (1993). The Biochemical Journal. 293, 781–788.

Henrissat, B., & Davies, G. (1997). Structural and sequence-based classification of glycoside hydrolases. Current Opinion in Structural Biology. 7, 637-644.

Huet, J., Rucktooa, P., Clantin, B., Azarkan, M., Looze, Y., Villeret, V., & Wintjens, R. (2008). X-ray structure of papaya chitinase reveals the substrate binding mode of glycosyl hydrolase family 19 chitinases. Biochemistry 47, 8283–8291.

Jones, D.T., Taylor, W.R., & Thomton, J.M. (1992). The rapid generation of mutation data matrices from protein sequences. Comput. Appl. Biosci. 8, 275–282.

Jones, R.L., & Vamer, J.E. (1967). The bioassay of gibberellins. Planta 72, 155–161.

Junges, A., Boldo, J.T., Souza, B.K., Guedes, R.L.M., Sbaraini, N., Kmetzsch, L., Thompson, C.E., Staats, C.C., de Almeida, L.G.P., de Vasconcelos, A.T.R., Vainstein, M.H., & Schrank, A. (2014). Genomic Analyses and Transcriptional Profiles of the Glycoside Hydrolase Family 18 Genes of the Entomopathogenic Fungus *Metarhizium anisopliae*. PLoS ONE 9(9): e107864. doi:10.1371/journal.pone.0107864.

Kasprzewska, A. (2003). Plant chitinases-regulation and function. Cellular and Molecular Biology Letters. 8, 809–824.

Kumar, S., Stecher, G., & Tamura, K. (2016). MEGA7: Molecular Evolutionary Genetics Analysis version 7.0 for bigger datasets. Molecular Biology and Evolution. 33, 1870-1874.

Landim, P.G.C., Correia, T.O., Silva, F.D.A., Nepomuceno, D.R., Costa, H.P.S., Pereira, H.M., Lobo, M.D.P., Moreno, F.B.M.B., Brandao-Neto, J., Medeiros, S.C., Vasconcelos, I.M., Oliveira, J.T.A., Sousa, B.L., Barroso-Neto, I.L., Freire, V.N., Carvalho, C.P.S., Monteiro-Moreira, A.C.O., & Grangeiro, T.B. (2017). Production in *Pichia pastoris*, antifungal activity and crystal structure of a class I chitinase from cowpea (*Vigna unguiculata*): Insights into sugar binding mode and hydrolytic action. Biochimie 135, 89-103.

Liu, Z., Huang, Y., Zhang, R., Diao, G., Fan, H., & Wang, Z. (2013). Chitinase Genes LbCHI31 and LbCHI32 from Limonium bicolor Were Successfully Expressed in Escherichia coli and Exhibit Recombinant Chitinase Activities. The Scientific World Journal Volume, Article ID 648382, 9 pages. http://dx.doi.org/10.1155/2013/648382.

Lombard, V., Golaconda Ramulu, H., Drula, E., Coutinho, P.M., & Henrissat, B. (2014). The carbohydrate-active enzymes database (CAZy) in 2013. Nucleic Acids Res. 42, D490–D495.

Marchler-Bauer, A. et al. (2017). CDD/SPARCLE: functional classification of proteins via subfamily domain architectures. Nucleic Acids Res. 45, 200-3.

Mazzinghy Alvarenga, S., Teixeira Caixeta, E., Hufnagel, B., Thiebaut, F., Maciel-Zambolim, E., Zambolim, L., & Sussumu Sakiyama, N. (2010). *In silico* identification of coffee genome expressed sequences potentially associated with resistance to diseases. Genetics and Molecular Biology. 33, 795-806.

Munoz-Munoz, J., Cartmell, A., Terrapon, N., Henrissat, B., & Gilbert, H.J. (2017). Unusual active site location and catalytic apparatus in a glycoside hydrolase family. PNAS 1-6.

Muralikrishna, G., & Nirmala, M. (2005). Cereal α-amylases – an overview. *Carbohyd. Polymers* 60, 163–173.

Muzzarelli, R.A.A. (1999). Native, industrial, and fossil chitins, in P. Jolles, R.A.A. Muzzarelli (Eds.), Chitin and Chitinases, Birkhauser, Basel, pp. 1-6.

Ohno, T., Armand, S., Hata, T., Nikaidou, N., Henrissat, B., Mitsutomi, M., & Watanabe, T. (1996). A modular family 19 chitinase found in the prokaryotic organism *Streptomyces griseus* HUT 6037. J. Bacteriol. 178, 5065–5070.

Petersen, T.N., Brunak, S., von Heijne, G., & Nielsen, H. (2011). SignalP 4.0: discriminating signal peptides from transmembrane regions. Nature Methods 8, 785-786.

Saitou, N., & Nei, M. (1987). The neighbor-joining method. A new method for reconstructing phylogenetic trees. Molecular Biology and Evolution. 4, 406-425.

Sangrador-Vegas, A., Mitchell, A.L., Chang, H.Y., Yong, S.Y., & Finn, R.D. (2016). GO annotation in InterPro: why stability does not indicate accuracy in a sea of changing annotation. Database. 1–8.

Sneath, P.H.A., & Sokal, R.R. (1973). Numerical Taxonomy. Freeman, San Francisco.

Sørensen, H.P., Madsen, L.S., Petersen, J., Andersen, J.T., Hansen, A.M., & Beck, H.C. (2010). Oat (*Avena sativa*) seed extract as an antifungal food preservative through the catalytic activity of a highly abundant class I chitinase. Appl. Biochem. Biotechnol. 160, 1573–1584. DOI 10.1007/s12010-009-8557-4.

Steentoft, C., Vakhrushev, S.Y., Joshi, H.J., Kong, Y., Vester-Christensen, M.B., Schjoldager, K.T.-B.G., Lavrsen, K., Dabelsteen, S., Pedersen, N.B., Marcos-Silva, L., Gupta, R., Bennett, E.P., Mandel, U., Brunak, S., Wandall, H.H., Levery, S.B., & Clausen, H. (2013). Precision mapping of the human O-GalNAc glycoproteome through SimpleCell technology. EMBO J. 32, 1478-1488, http://dx.doi.org/10.1038/emboj.2013.79.

Thompson, J.D., Higgins, D.G., & Gibson, T.J. (1994). ClustalW, improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res. 22, 4673-4680.

Udaya Prakash, N.A., Jayanthi, M., Sabarinathan, R., Kangueane, P., Mathew, L., & Sekar, K. (2010). Evolution, homology conservation, and identification of unique sequence signatures in GH19 family chitinases. J. Mol. Evol. 70, 466-478. DOI 10.1007/s00239-010-9345-z.

Uno-Okamura, K., Soga, K., Wakabayashi, K., Kamisaka, S., & Hoson, T. (2004). Purification and properties of apoplastic amylase from oat (*Avena sativa*) seedlings. Physiologia Plantarum. 121, 117–123.

Wu, M.J., McKay, S., Howes, N., Hegedus, E., & Chin, J. (2011). Polymorphism and pedigree analysis of β-amylase isozymes in Australian wheat. Journal of Cereal Science. 53, 362-370.

Yu, C.S., Chen, Y.C., Lu, C.H., & Hwang, J.K. (2006). Prediction of protein subcellular localization. Proteins: structure, function, and bioinformatics. 64, 643-651.

Ziegler, P. (1999). Cereal *beta*-amylases. J. Cereal Sci. 29, 195–204.

CHAPTER THREE

PLANT LIPOLYTIC ENZYMES: A CASE STUDY OF PHOSPHOLIPASES A (PLAS) FROM OAT (*AVENA SATIVA*)

Abstract

The plant lipolytic enzymes discussed in this chapter are mainly phospholipases, which include diverse series of enzymes catalyzing polar lipids, i.e., glycerophospholipids, at different ester bonds. Plant phospholipases could be classified into three major types namely phospholipase A (PLA), phospholipase C (PLC) and phospholipase D (PLD); and within each class, there are multiple forms of phospholipases.

After an overview of some important characterizations of these plant enzymes, the study of sequences of PLA from oat (*Avena sativa*) is presented as a case study. In this way, insights into the structural characterization of such predicted enzyme are highlighted in this chapter.

Keywords: *Arabidopsis*; *Avena sativa*; Applications; Bioinformatics; Characterization; Functions; Genes; Identification; *in silico* approach; Lipids; Lipolytic enzymes; Valorization; Lipidomics; Phospholipids; Phospholipases; Physiology; Plants; PLAs; Prediction; Proteins.

1. Introduction

The plant lipid hydrolyzing enzymes are vital compounds of the plant. The plant glycerolipids are constantly subjected to an important turnover by relevant enzymes namely galactolipases and phospholipases.

The lipid hydrolyzing activities are found in many sources in living systems as these activities are omnipresent either in prokaryotes e.g., bacteria, or eukaryotes e.g., animals and plants.

The enzymes with lipolytic activities could be defined as esterases, i.e., biocatalysts of esters as substrates. Different lipolytic enzymes could be distinguished based on the substrate specificities and the positions of cleavage sites within the relevant lipid molecule.

Phospholipids are polar lipids that provide the backbone for membranes of living organisms. There are different types of phospholipids that could serve as rich sources of signaling messengers, and play important roles in lipid metabolism.

Phospholipases activities are involved in the structure and stability of cellular membranes, and in many other cellular functions. Many reports have dealt with the involvement of phospholipases from, in particular, animal systems, in the cell regulation and signaling processes (Rhee and Bae, 1997; Assmann and Shimazaki, 1999; Williams, 1999; Liscovitch et al., 2000). Phospholipases from plant systems have been the focus of many recent reports (Wang, 2001; Chen et al., 2011Wang et al., 2012; Ben Halima, 2019) due to their promising properties, cellular functions, and potential applications. Moreover, several reports have reviewed various aspects of phospholipids and phospholipases from plants in the cellular regulation in particular (Chapman, 1998; Munnik et al., 1998; Drobak et al., 1999; Wang, 1999, 2000; Stevenson et al., 2000).

This chapter will provide an overview of some important aspects of plant phospholipases. In fact, there are different classes of plant phospholipases: those that catalyze the cleavage of acyl groups (phospholipases A1 (PLA_1), A2 (PLA_2), B (PLB); lysophospholipases and acyl hydrolases) or those cleaving head group (phospholipases D (PLD) and C (PLC)) of phospholipids.

An attempt to summarize some knowledge of lipolytic enzymes in particular plant PLAs is the focus of the current chapter.

2. Background

The most frequent sources of plant lipolytic enzymes are seeds and oilseeds (oleaginous plants) or cereal seeds in particular. In addition, laticifers are a good source of enzymes of lipolytic activities. These enzymes are essential in various cellular and physiological processes in plants.

Some main features of lipolytic enzymes from different plant origins (oleaginous plants, cereals, and laticifers) are given in Table 1. In fact, Table 1 just gives an idea about the interesting biochemical properties and diverse applications of plant lipolytic enzymes, which contribute to their use in biocatalysis as potential industrial enzymes (Table 2).

Table 1. Some sources of plant lipolytic enzymes (oleaginous plants, cereals, and laticifers) with their biochemical characteristics and applications

Plant lipolytic enzymes				
Sources	**Biochemical characteristics**			**Applications**
	Optimum pH	**Optimum temperature**	**Selectivity** [a]	
Oleaginous plants				
Jatropha curcas	pH 7.5	37•C		Biodiesel synthesis
Castor bean	pH4.5	3●•C	R: *sn* -1 and *sn* -2, CL: short and medium, SD: unsaturated	Esterification of fatty acids and glycerol
Coconut	pH 8.5	3●-4●•C	R: *sn* -1,3	
Cereals				
Rice (bran)	pH 11	8●•C	R: nonspecific on TG, *sn* -2 on PC	Phospholipid hydrolysis
Corn			CL: Long, SD: unsaturated	TG hydrolysis
●at	9.●	75•C (EII isoenzyme) and 65•C (EIII isoenzyme)		●at shelf-life studies

Wheat	pH 5.5	32–37°C		Hydrolysis and esterification
Laticifers				
Euphorbia characias	pH 5	60°C	CL: short and medium	Hydrolysis of TG and synthetic monoesters
Babaco (*Carica pentagona)*	pH 8	50°C	a_w 0.38, R: *sn* -1,3, CL: short, SD: unsaturated	Alcoholysis of sunflower oil, Naproxen resolution
Carica papaya	pH 9	50°C	R: *sn*-3	Lipids modification, asymmetric resolutions

TG triglycerides, *PC* phosphatidylcholine
[a] *R* regioselectivity, *CL* chain-length preference, *SD* fatty acid saturation degree preference
Source: (Rivera et al., 2012; Barros et al., 2010; Ben Halima, 2019)

Lipid hydrolyzing activities are used in diverse industrial applications (Table 2). The flexibility of biochemical characteristics found in lipolytic enzymes leads to interesting applications either in degradation or synthesizing processes at industrial scale, such as the food industries and related areas (Table 2).

There are many books related to general aspects of lipolytic enzymes. For example, the book of Brockerhoff and Jensen (1974) unveiled details concerning the nature, nomenclature, kinetics of lipolysis, and properties of lipolytic enzymes. As well, many other available books that reported general information on lipids and lipolytic enzymes such as the books of Doolittle and Reue (1999), Müller and Petry (2004), Murphy (2005), Sandoval (2012), Plaxton and Lambers (2015), and Barh et al. (2015). Recently, the book of Ben Halima (2019) has focused on unique sequence signatures in plant lipolytic enzymes. Lipolytic enzymes, which include various types of enzymes with lipolytic activities such as plant phospholipases, would get much interest to be used in industrial applications due to their relevant properties.

Table 2. Some industrial applications of lipolytic enzymes

Lipolytic enzymes	
Industry	**Application**
Detergent	Lipid stain removal
Starch and fuel	Synthesis of lipase-catalyzed biodiesel
Food	Improvement of food texture
Baking	Dough stability and conditioning
	Dough stability and conditioning
	Dough strengthening, bread whitening
Pulp and paper	Pitch control, contaminant control
Fats and oils	Transesterification
	Degumming, lysolecithin production
Organic synthesis	Resolution of chiral alcohols and amides
Leather	Depickling
Environmental application	Degradation of lipid wastes
	Removal of solid and water pollution by hydrocarbons, oils, and lipids

Source: Ramnath et al. (2017); Ben Halima (2019)

It should be noted that there are several hydrolase activities acting on the *sn*-1 or *sn*-2 positions of phospholipids and glycolipids have been described in plant extracts. In fact, the isolation of the first genes encoding lipolytic enzymes as well as the completion of the sequencing of the genome of *Arabidopsis* have without doubt greatly advanced the understanding and knowledge of the roles of lipolytic enzymes in plants.

Phospholipases hydrolyze phospholipids and constitute an important group of biocatalysts. They are a complex group of enzymes that are involved in a wide range of plant processes mainly in the plant physiological processes

such as lipid biosynthesis, membrane homeostasis, cell signaling and stress responses (Chen et al., 2011; Ryu, 2004; Wang, 2001).

3. Plant Phospholipases

Plant phospholipases form groups of lipolytic enzymes with diverse interesting characteristics, which can be classified into phospholipases A (PLAs), C (PLCs) and D (PLDs) according to the sites of phospholipids hydrolysis (**Figure 1**).

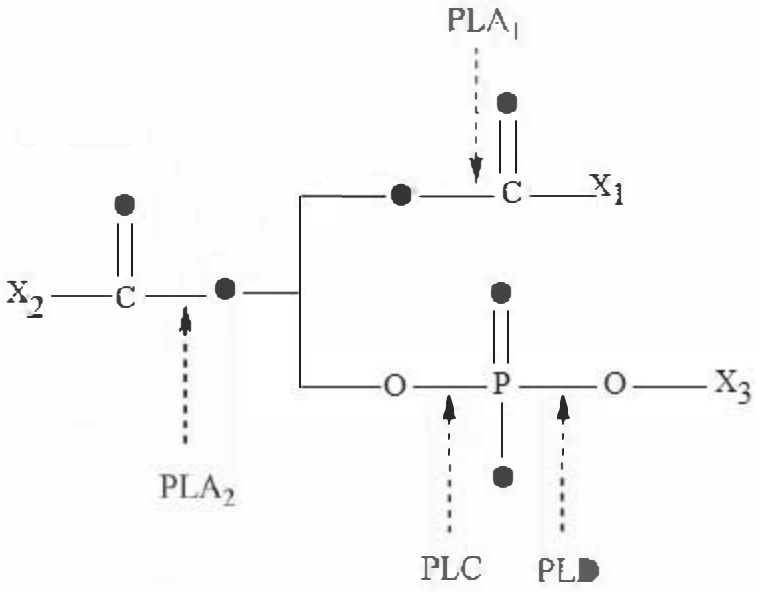


Figure 1: Sites of phospholipid hydrolysis by the major types of plant phospholipases namely PLD: phospholipase D, PLC: phospholipase C and PLA: phospholipase A. X1 and X2: represent fatty acid chains. X3: represents the head group of a phospholipid such as choline, ethanolamine, serine, glycerol, inositol or inositol 4,5-bisphosphate.

PLAs (PLA_1 and PLA_2) catalyze the hydrolysis of acyl groups from membrane lipids, whereas PLCs and PLDs catalyze the hydrolysis of the first and the second phosphodiesteric bond from the phospholipid head groups, respectively (**Figure 1**).

Phospholipases could have various localizations in plants. Table 3 delineates some classes of plant phospholipases with their names and localizations.

Table 3. Name and localization of some plant phospholipases

Type	Organism	Name	Localization	References
PLA_1	*Arabidopsis thaliana*	DAD1	Chloroplastic (anther)	- Ishiguro et al. (2001)
PLA_2	*Arabidopsis thaliana*, carnation, castor bean, citrus, durum wheat, elm, rice, soybean, tobacco	sPLA2 (Groups XIA and XIB)	Secreted, cytosolic (various tissues)	- Verlotta et al. (2013)
	Arabidopsis thaliana, rape, broad bean, cucumber, pepper, potato, tobacco, tomato	Patatin-like phospholipase	Cytosolic (various tissues)	- Matos and Pham-Thi (2009) - Scherer et al. (2010)
PLC	*Arabidopsis thaliana*, bean, cedar, cereals, *Lilium*, moss, potato, soybean, tobacco, tomato	PI-PLC	Cytosolic (various tissues)	- Pokotylo et al. (2014)

	Arabidopsis thaliana, avocado, peanut, rape, rice, soybean, tomato, wheat	NPC	Cytosolic (various tissues)	- Pokotylo et al. (2013)
PLD	*Arabidopsis thaliana*, cabbage, castor bean, grape, *Jatropha curcas*, mustard, peanut, poplar, poppy, strawberry, rice, sunflower, tobacco, tomato	C2-PLD (PLDα, PLDβ, PLDγ, PLDδ, PLDε), PXPH-PLD (PLDζ1, PLDζ2)	Secreted, cytosolic, membrane-bound (various tissues)	-Selvy et al. (2011) -Khatoon et al. (2015)

Source: Borrelli and Trono (2015); Ben Halima (2019)

Each class of phospholipases is further divided into different families or subfamilies of enzymes that could differ in sequences and biochemical properties (Wang, 2001; Wang et al., 2012).

The different types of phospholipases play important roles in cellular functions and physiological processes in plants (Wang, 2004). In fact, the roles of plant phospholipases could be grouped into three major categories namely, cellular regulation, lipid degradation and membrane lipid remodeling (Wang et al., 2012).

The different biochemical properties of plant phospholipases such as substrate specificity and reaction conditions could be explored for many applications. Some of the important industrial applications of plant phospholipases are related to the baking industries (e.g., for dough stability and conditioning) or to the fats and oils industries (e.g., for degumming and lysolecithin production).

It is also important to note that biochemical analyses and modern genetics revealed that plants possess multiple isoforms of phospholipases (e.g., PLA, PLC and PLD) in addition to the multiple types of these enzymes. Indeed, the multiple isoforms of plant phospholipases could differ in expression patterns, gene or protein structures, catalytic properties, and developmental or tissue-specific distribution. All of this complexity of plant phospholipases suggests diverse roles attributed to these enzymes, which have yet to be fully clarified (Chen et al., 2011).

Furthemore, there are many considerable strides, especially in the two past decades, which have been made in the isolation, identification, classification, characterization, and functional and biochemical analyses of plant phospholipases. Several plant phospholipases, in particular, plant PLDs, could be used in plant biotechnology to enhance, for instance, abiotic stresses tolerance (Zhang et al., 2008; Peng et al., 2010). Plant phospholipases could also be used in industries such as in industrial biotechnology and food applications due to their suitable molecular and catalytic diversity in such industrial applications (Mansfeld, 2009; Servi, 1999; De Maria et al., 2007).

This chapter will focus mainly on some characteristics of plant phospholipases. In fact, different classes of phospholipases are found in plants that catalyze the cleavage of acyl groups (e.g., phospholipases A1 (PLA_1) and A2 (PLA_2)) or head group (e.g., phospholipases C (PLC) and D (PLD)) of phospholipids.

Readers are referred to Wang (2001) and Chen et al. (2011) for a more comprehensive review and description of earlier developments relative to plant phospholipases.

3.1. Hydrolysis of acyl groups of phospholipids: plant phospholipases A (PLAs)

Phospholipases A (PLAs) are biocatalysts that hydrolyze acyl groups of phospholipids generating lysophospholipids and free fatty acids. Indeed, PLAs could hydrolyze glycerophospholipids at their *sn*-1 and/or *sn*-2 position to liberate a free fatty acid. The phospholipases A could have hydrolytic activity at both the *sn*-1 and *sn*-2 positions of the phospholipids (PLA) or be specific only for the *sn*-1 position (PLA_1) or be specific to only the *sn*-2 position (PLA_2) (Scherer et al., 2010). PLA_1 and PLA_2 release fatty acids by hydrolyzing acyl groups from the *sn*-1 and *sn*-2 positions of phospholipids, respectively (**Figure 1**).

In fact, PLAs from plants could be classified into four families based on biological properties and sequence data. These four families are as following: The phosphatidylcholine (PC)-hydrolyzing PLA_1 (PLA_1), the phosphatidic acid (PA)-preferring PLA_1 (PA-PLA_1), the secretory low-molecular-weight PLA_2 ($sPLA_2$), and the patatin-like PLA (pPLA) (Ryu, 2004).

Below are some enzymatic properties of the different families of plant PLAs as well as their cellular functions and potential applications.

3.1.1. Plant phospholipases A1: PLA_1 and PA-PLA_1 families

Phospholipases A1 (PLA_1s) catalyze the cleavage of the acyl groups of phospholipids from the *sn*-1 position, thus, releasing *sn*-2 lysophospholipid and free fatty acids.

3.1.1.1. PLA_1 family

3.1.1.1.1. PLA_1 family: enzymatic property

From the *Arabidopsis* genome database, twelve isoforms of PLA_1 have been identified (Ryu, 2004). It was also reported that the twelve isoforms of *Arabidopsis* PLA_1 (Ryu, 2004) could be classified into three classes (I, II, III) based on two criteria: the presence of particular N-terminal stretches in these proteins and sequence similarities in their catalytic region (Wang et al., 2012; Ryu, 2004; Ishiguro et al., 2001).

In the report of Dörmann (2005), and according to Kato et al. (2002), a mutant of *Arabidopsis* (*sgr2* for shoot gravitropism deficient) was isolated. This mutant carries a mutation in a gene orthologous to animals PLA_1. On the other hand, it was reported that there is only one PLA1-like gene locus, which is known in *Arabidopsis*, nevertheless, the activity of PLA1 of the corresponding protein has not been confirmed (Dörmann, 2005).

The PLA_1 family includes the two following *Arabidopsis* PLA_1 namely DAD1 (*At*PLA_1-Iβ1) and DGL1 (*At*PLA1-Iα1), which both had experimentally been confirmed to have the activity of glycerolipid acyl hydrolase *sn*-1 specificity (Ishiguro et al., 2001; Hyun et al., 2008; Seo et al., 2009). It is worth noting that there is a difference in substrate preference for *sn*-1 specificity. Indeed, the substrate preference is broad toward different glycerolipids, in contrast to the strict *sn*-1 specificity (Wang et al., 2012).

DAD1 has been shown to display more important lipolytic activity toward phosphatidyl-choline (PC) than towards mono-galactosyldiacylglycerol (MGDG) or triacylglycerol (TAG) (Ishiguro et al., 2001). However, the enzyme DGL displayed more important activity toward di-galactosyldiacylglycerol (DGDG) than towards PC, TAG, and MGDG (Seo et al., 2009).

3.1.1.1.2. PLA_1 family: cellular function

According to the sequence analyses of PLA_1 family, the class I, II, and III of such enzyme were predicted to be localized in chloroplast, cytosol, and mitochondria, respectively (Ryu, 2004; Ishiguro et al., 2001).

DAD1 is a chloroplastic *At*PLA_1-Iβ1 that could release linolenic acid leading to the jasmonic acid (JA) biosynthesis, which is required for pollen maturation, anther dehiscence, and flower opening (Ishiguro et al., 2001).

Moreover, many reports have indicated the implication of *DAD1*, another PLA_1, *DGL* (*At*PLA_1-Iα1), and AtPLA_1-Ia1 in the biosynthesis of JA, which is required for wound response and/or defense responses to pathogens (Ishiguro et al., 2001; Hyun et al., 2008; Ellinger et al., 2010; Wang et al., 2012).

The reports of Hong et al. (2000) and Thompson et al. (2000) have indicated that the class II protein might have a function in the degradation of membrane lipids during senescence. The class III protein has no reported functions yet (Wang et al., 2012).

3.1.1.2. PA-PLA_1 family: some highlights in its enzymatic property and cellular function

Concerning plant PA-PLA_1 activity, a single gene homologous to PA-PLA_1 (AF045022) from bovine testis (Higgs et al., 1998) has been identified in the *Arabidopsis* genome database (Ryu, 2004; Kato et al., 2002). As well, homologous Expressed Sequence Tags (EST) clones for PA-PLA_1 were identified in other plants such as tomato and rice (Wang et al., 2012).

According to many reports, the mutant of a putative PA-PLA_1 gene (*sgr2*) of *Arabidopsis* (*At*PA-PLA_1) is involved in shoot gravitropism (Wang et al., 2012; Ryu, 2004; Kato et al., 2002; Morita et al., 2002).

The cells of *sgr2* embryos have shown abnormal structures in the vacuolar membranes, which are known to be involved in an early step in gravity sensing (Wang et al., 2012; Morita et al., 2002).

PA-PLA_1 could also hydrolyze PA generating lysophosphatidic acid (Wang et al., 2012). PA is involved in vesicular trafficking in eukaryotic cells (Wang et al., 2012; O'Luanaigh et al., 2002) and lysophosphatidic acid was found to be a bioactive molecule in animal systems (Yan et al., 2003).

3.1.2. Plant phospholipases A2: $sPLA_2$ and pPLA families

Phospholipases A2 (PLA_2s) catalyze the cleavage of the acyl groups of phospholipids and lysophospholipids from their *sn*-2 position.

The PLA_2 enzymes were originally defined in animals as a superfamily of PLA_2s (Six and Dennis, 2000), which comprise the cytosolic or Ca^{2+}-activated $cPLA_2$ (Dessen, 2000), the Ca^{2+}-independent $iPLA_2$ (Balsinde and Balboa, 2005) and the secreted $sPLA_2$ (Six and Dennis, 2000).

Several gene loci with high similarity of the sequence to PLA_2 from yeast and animals were identified in the genome of *Arabidopsis*.

The first plant PLA_2 has been isolated from elm (Ståhl et al., 1998, 1999).

The activity of plant PLA_2 was suggested to be involved in removing of unusual fatty acids from membrane lipids of different plant species such as those accumulating unusual fatty acids e.g., castor and elm, or those not accumulating such fatty acids e.g., rapeseed (Ståhl *et al.*, 1998, Dörmann, 2005).

3.1.2.1. $sPLA_2$ family

3.1.2.1.1. $sPLA_2$ family: enzymatic property

The phospholipase A isolated from elm is related to secretory PLA_2 ($sPLA_2$) from animals. The elm PLA_2 might be localized to the extracellular matrix; thus, it could not be related to microsomal PLA_2 activity that is involved in the lipid acylation/deacylation (Dörmann, 2005).

Four $sPLA_2$ members were identified in *Arabidopsis* namely, *Ats*PLA_2-α, *Ats*PLA_2-β, *Ats*PLA_2-γ and *Ats*PLA_2-δ (Ryu, 2004; Bahn et al., 2003; Lee et al., 2003).

There are several isolated $sPLA_2$ genes from rice and carnation (Kim et al., 1999; Stahl et al., 1999). Many reports were dealing with analyzing the biochemical and enzymatic properties of the $sPLA_2$ family (Bahn et al., 2003; Lee et al., 2003; Stahl et al., 1998, 1999; Kim et al., 1994; Mansfeld and Ulbrich-Hofmann, 2007).

$sPLA_2$s have a PA2c domain containing the highly conserved Ca^{2+}-binding loop and an active site motif with the conserved His/Asp dyad. The mature enzymes contain 12 cysteine residues, which can form six disulfide bonds (Wang et al., 2012).

Multiple forms of $sPLA_2$ exist in plants (Lee et al., 2005).

The substrate preference differs among the isoforms of plant $sPLA_2$, in particular, those from *Arabidopsis*. For example, the isoforms *Ats*PLA_2-γ and δ showed a very strong preference for PE over PC as substrate. As well, *Ats*PLA_2-α and -β preferred PE over PC (Lee et al., 2005). The report of Lee et al. (2005) highlighted some physicochemical properties of the four identified $sPLA_2$ members of Arabidopsis. Indeed, these authors reported that the ranges of optimal pH for the activities of *Ats*PLA_2-α, -β, -γ, and -δ were pH 6–11, 6–7, 7–9, and 8–9, respectively. The specific activity of *Ats*PLA_2-α continuously increased when the concentration of Ca^{2+} increased to 10 mM, while the activities of *Ats*PLA_2-β, -γ, and -δ reached a plateau at micromolar levels of Ca^{2+} (Lee et al., 2005).

Unlike AtsPLA_2-β, -γ, and -δ that show no distinct preferences for particular fatty acids, AtsPLA_2-α shows higher activity toward substrates containing acyl chain with a higher degree of unsaturation in *sn*-2 position (Mansfeld and Ulbrich-Hofmann, 2007; Lee et al., 2005).

The secretory low-molecular-weight PLA_2s ($sPLA_2$s) are calcium-dependent small enzymes of around 14 kDa and with strict *sn*-2 stereospecificity. Both the catalytic center and the enzymes are not related to patatin (Scherer et al., 2010).

3.1.2.1.2. $sPLA_2$ family: cellular function

The transgenic study in *Arabidopsis* was used to elucidate the function of plant secretory PLA_2 ($sPLA_2$) activity, and it was demonstrated that this enzyme is critical for the cell elongation and gravitropism (Lee et al., 2003).

Two genes with sequence similarity to mammalian cytosolic PLA_2 were identified in the *Arabidopsis* genome (Beisson et al., 2003). Nevertheless, as reported in Dörmann et al. (2005), it was not known whether these PLA_2s are involved in the plant lipid breakdown, turnover or signaling.

The hydrolysis of phosphatidyl-ethanolamine (PE) by PLA_2 generates lysophosphatidylethanolamine (LPE) that was reported to retard senescence in plants, for instance, tomato leaf and fruit senescence (Dörmann, 2005; Farag and Palta, 1993). Therefore, PLA_2 isoforms could be directly involved in cell signaling (Dörmann, 2005).

$sPLA_2$ genes from plants encode proteins with signal peptides in their N-terminal sequence and the proteins are, in general, of low molecular weight (13-18 kDa) after secretion (Wang et al., 2012).

*At*sPLA_2-α could found different localizations such as in apoplasts, Golgi bodies (GB) in mature leaves or GB in actively growing young leaves, and root tissues (Froidure et al., 2010; Lee et al., 2010; Wang et al., 2012).

The isoforms of *At*sPLA_2-β and -δ could be localized at the endoplasmic reticulum (ER) (Seo et al., 2008; Kim et al., 2011) and *At*sPLA_2-γ could be localized at ER and GB (Kim et al., 2011).

According to Lee et al. (2010), *At*sPLA_2-α could act in Golgi-related compartments and modulate the PIN proteins trafficking in root tissues.

The *At*sPLA_2-β has been shown to have a role in the stomatal opening, shoot gravitropism and shoot cell elongation (Lee et al., 2003; Seo et al., 2008).

The *At*sPLA_2-β, -γ, and -δ are expressed during the pollen development and have critical roles in the pollen development, in germination and in tube growth (Seo et al., 2008).

Two orthologs of $sPLA_2$s isolated from *Citrus sinensis* (CssPLA_2-α and CsPLA_2-β) were shown to exhibit diurnal rhythmicity in leaf tissue (Wang et al., 2012). The expressions of *CssPLA2-α* and *CsPLA2-β* have been redundantly mediated by different lights namely, blue, green, red, and red/far-red light. Nevertheless, blue light has been a major factor affecting the expression of *CssPLA2-α* and *CsPLA2-β* (Liao and Burns, 2010).

3.1.2.2. pPLA family

3.1.2.2.1. pPLA family: enzymatic property

Patatins or patatin-related enzymes are potato tuber proteins with acyl-hydrolyzing activity. In the review of Scherer et al. (2010), recent developments regarding the patatin-related enzymes have been highlighted and the authors have proposed the nomenclature pPLA for the patatin-related phospholipase A enzyme.

The patatin-like PLA (pPLA) gene family is represented by 10 genes in *Arabidopsis* (*Arabidopsis thaliana*) genome database (Ryu, 2004; Holk et al. 2002), which encode proteins related to the potato tuber storage protein patatin (Rosahl et al. 1986; Andrews et al. 1988). In *Arabidopsis*, these 10 identified pPLAs have been divided into 3 classes based on the deduced amino acid sequences and gene structures into pPLAI, pPLAII(α, β, γ, δ, ε), and pPLAIII(α, β, γ, δ) (Ryu, 2004; Holk et al., 2002; Scherer et al., 2010; Matos and Pham-Thi, 2009).

Unlike pPLAIIIs, pPLAI and pPLAII have the canonical lipase/esterase S-D dyad consensus catalytic sequences (Wang et al., 2012). In addition, pPLAI and pPLAIIs display a broad enzymatic specificity and could hydrolyze phospholipids as well as galactolipids (La Camera et al., 2005; Matos et al., 2001, 2008; Yang et al., 2007; Dhondt et al., 2002; Sahsah et al., 1998; Galliard, 1971; Reina-Pinto et al., 2009).

Moreover, pPLA-I has a preference of oxidized over nonoxidized galactolipids (Yang et al., 2007).

The plant pPLA-I has the ability to catalyze the hydrolysis of galactolipids as well as it has *in vitro* low specificity position for the *sn*-1 and *sn*-2 acyl chains and each of these features could be a plant-typical pPLA property (Wang et al., 2012). It is worth noting that there is no animal pPLA that has ever been tested with the substrate galactolipids or other glycolipids as a substrate (Wang et al., 2012).

According to many reports, none of the pPLAs examined so far possesses TAG lipase activity (Ishiguro et al., 2001; Scherer et al., 2010; Matos et al., 2001; Rietz et al., 2010).

The report of Li et al. (2011) showed that pPLAIIIβ could catalyze the hydrolysis of phospholipids and galactolipids, and additionally could have acyl-CoA thioesterase activity.

3.1.2.2.2. pPLA family: cellular function

It was reported that pPLAI could play a role in promoting basal, but not pathogen-induced, JA production (Yang et al., 2007).

Arabidopsis was more resistant to *Pseudomonas syringae* and *Botrytis cinerea* when pPLA-IIα was suppressed, while the overexpression of this enzyme enhanced the resistance against cucumber mosaic virus as well as plant cell death but decreased resistance against *Pseudomonas syringae* and *Botrytis cinerea* (Ackermann et al., 1994).

Analysis of gain- and loss-of-function suggested that pPLAIIα could be involved in the mobilization of fatty acid precursors for biosynthesis of specific oxylipins, but with limited impact on JA accumulation (La Camera et al., 2009). Phenotypic analysis of pPLA-IIγKO plants indicated that pPLA-IIγ could be necessary for root response to phosphate deficiency (Rietz et al., 2010).

KO plants did not display the classical auxin-related responses to the deficiency of phosphates, such as downregulation of the activity of the main root meristem and increase in lateral root formation (Perez-Torres et al., 2008).

Loss-of-function of pPLA-IIε could lead to a decrease of the numbers of the lateral root, thus, it seems that the expression of either one of these genes could be involved in auxin-related root responses (Rietz et al., 2010; Wang et al., 2012).

According to the report of Scherer et al. (2010), pPLAs enzymes could be involved in different cellular functions, including plant responses to auxin, elicitors or pathogens, abiotic stresses, and lipid mobilization during seed germination. Indeed, long-term responses to several biotic signals (salicylic acid, ABA, and pathogens) and abiotic stresses, in particular, drought and phosphate deficiency, involve transcriptional regulation of pPLA genes as part of the response after hours or days (Scherer et al., 2010; Wang et al., 2012).

The overexpression of pPLA-IIIδ could result in a STURDY mutant leading to thick leaves, large seeds, short siliques, delayed growth, stiff inflorescence stem, and round flowers (Huang et al., 2001).

The alteration of the expression of *pPLAIIIβ* as well as the resulting lipid changes could alter the cellulose content and the cell elongation in *Arabidopsis* (Li et al., 2011).

It was reported that the *pPLAIIIβ*-KO (knockout) plants had longer leaves, hypocotyls, petioles, primary roots, and root hairs than those observed in Wild Type (WT) plants, while the *pPLAIIIβ*-OE (overexpression) plants exhibited the opposite phenotype (Wang et al., 2012). Moreover, the *pPLAIIIβ*-OE plants had a significantly lower content of cellulose and mechanical strength than those observed in WT plants. All these data would suggest the involvement of PLA-mediated membrane lipid metabolism in the modulation of cellulose production in plants (Wang et al., 2012).

3.1.3. Potential applications

Phospholipases have many potential applications in particular for use in industries. PLAs are used in degumming of edible oils and in the synthesis of triacylglycerols enriched in polyunsaturated fatty acids (Gunstone, 1999). The plant phospholipases have gained much interest in various applications due to the recent progress on their biochemical properties and functions. The multiple forms of plant $sPLA_2$ are promising enzymes in diverse applications than that of animal sources due to their distinctive differences in substrate selectivity with respect to head groups and acyl chains of phospholipids (Lee et al., 2005; Mansfeld, 2009).

The unique properties found in plant phospholipases could lead to opening new horizons to deal with engineering such biocatalysts with improved performance, thus, raising the use of these enzymes in many relevant applications.

Phospholipases of plant origin could be particularly suitable for biodiesel production and enzymatic degumming in vegetable oil refinement because of the adaptation of substrate specificities to phospholipids, which is preferably occurring in plants (Wang et al., 2012).

The sPLAs from plants could have advantages over enzymes from animals concerning yield and performance in the incorporation of polyunsaturated fatty acids such as α-linolenic acid into egg-yolk PC for the production of a functional food with a high impact on human health. Therefore, the processes for producing phospholipids with uncommon fatty acids that match special functional requirements could be feasible using plant phospholipases (Wang et al., 2012). Plant phospholipases A2 could also

have another helpful aspect in applications consisting of their lower extent of interfacial activation that makes these biocatalysts further promising in relevant applications (Kim et al., 1999).

Moreover, plant phospholipases could be exploited in crop improvement by genetic manipulations. In fact, the overexpression or suppression of the gene encoded phospholipases was reported to enhance crop yield thanks to the involvement of such genetically modified plants to increase resistance to abiotic stresses and a variety of phytopathogenic infections (Yang, et al., 2007; Peng, et al., 2010; Vossen, et al., 2010; Yamaguchi, et al., 2009; Ryu, et al., 2009).

The report of Casado et al. (2012) reviewed some important uses of phospholipases in general such as their use in the food industry that includes edible oils, dairy, baking products, and emulsifying agents. Indeed, phospholipase is expected to be used as an emulsifier in aqueous food products, e.g., in the case of sauces, dressings and bakery products. Therefore, increasing in this end use application of phospholipases is expected to drive the market of such enzyme (Ben Halima, 2019).

3.2. Hydrolysis of head group of phospholipids: plant phospholipases C (PLCs) and D (PLDs)

Plant PLCs are used to catalyze the hydrolysis of phospholipids generating diacylglycerol (DAG) and a respective phosphorylated head group from the substrate (**Figure 1**). These biocatalysts could be classified into three distinguished families of enzymes according to substrate specificity and cellular function, namely, phosphoinositide-specific PLC (PI-PLC), non-specific PLC (NPC), and glycosylphosphatidylinositol (GPI)-PLC. In fact, the PI-PLCs, also known as phosphatidylinositol-4,5-bisphosphate ($PI(4,5)P_2$)-PLCs, are recognized to catalyze primarily phosphoinositides; while NPCs, also known as PC-PLCs, could catalyze common membrane PLs especially PC but other phospholipids such as PE; whereas (GPI)-PLCs hydrolyze GPI-anchors on proteins (Nakamura et al., 2005; Peters et al., 2010, Wang et al., 2012).

The phospholipase D (PLD) catalyzes the hydrolysis of glycerophospholipids, especially PC, to produce PA and a free head group, e.g., choline (**Figure 1**). PLDs can also catalyze the interconversion of polar head groups of phospholipids by transphosphatidylation reaction in the presence of primary alcohol within the substrate. The bio-catalysis of this reaction consists of transferring the phosphatidyl moiety to an alcohol

leading to the synthesis of naturally less abundant phospholipids such as phosphatidylserine (PS), phosphatidylglycerol (PG), or PE from highly abundant ones such as lecithin or PC (Danmjanovic and Iwasaki, 2013). The reaction of transphosphatidylation is unique to PLD and it would be used to identify PLD activity *in vivo*.

PLDs are among the main lipid hydrolyzing enzymes implicated in lipid-mediated signaling and are ubiquitous enzymes that are present in various living organisms from eukaryotes to prokaryotes and archaea.

PLDs and PLCs are important enzymes for the physiological processes and cellular functions of the plant. They could play critical roles in multiple plant growth and developmental processes, as well as in hormone and stress responses and possible multivalent functionality could be assigned to these enzymes (Hong et al., 2016; Novotna et al., 2000; Abdelkafi et al., 2012).

For more information on plant PLCs and PLDs, please see chapters 4 and 5 in the book of Ben Halima (2019).

3.3. A case study of oat PLAs: identification of oat (*Avena sativa*) PLAs based on bioinformatics analysis with blast search

Oat (*Avena sativa* L.) is a cultivated plant that belongs to the *Poaceae* family (*Gramineae*). This cereal is a nutrient-rich food and its lipids and nutrient components could be valorized in industrial applications (Ben Halima et al., 2015). Oats are unique among the common cereal grains since they have high lipid content and their lipolytic enzymes are 10–15 times more active than those of wheat (Matlashewski et al., 1982).

A. sativa is an allohexaploid (2n=6x=42) crop species with an estimated 1C genome size of 13.23pg, corresponding to about 13000 Mbp (Bennett and Smith, 1976). Therefore, oat has a large genome that has not yet been fully assessed. In fact, complex polyploidy genomes such as that of oat require substantial investigations and contributions of the bioinformatics pipeline. Until now, no cultivar of oat has been reported to have a completed genome sequence, thus, many oat protein sequences, for instance, PLAs, need further characterization.

Since no amino acid sequence of oat phospholipase, in particular, PLA, is available in the protein data bank, an attempt is made in this report to

identify and characterize the catalytic amino acid residues of oat PLA by bioinformatics analyses and *in silico* approach.

In fact, the primary aim of this work is to identify new insights into predicted sequences of oat PLA_1 and PLA_2 based on bioinformatics analysis and subsequently to get insight into the oat genome.

Here, we would report especially new insights from bioinformatics analysis into sequences of PLA_1 and $sPLA_2$ from oat that were predicted according to the Transcriptome Shotgun Assembly (TSA) database of *Avena sativa*; nevertheless, both such predicted sequences could be either full-length or truncated gene sequences. It could also be better to describe the identified oat proteins as hypothetical, predicted, uncharacterized, probable and putative proteins.

3.3.1 Retrieval of protein sequences

The amino acid sequences from plant phospholipases serving for comparison with the predicted oat proteins were retrieved from the protein database of the National Center for Biotechnology Information (NCBI, http://www.ncbi.nlm.nih.gov/protein/). The sequences were saved in FASTA format. An outline of the *in silico* approach steps that could be followed to define/predict oat PLAs in seedling extract has been portrayed in **Figure 2**.

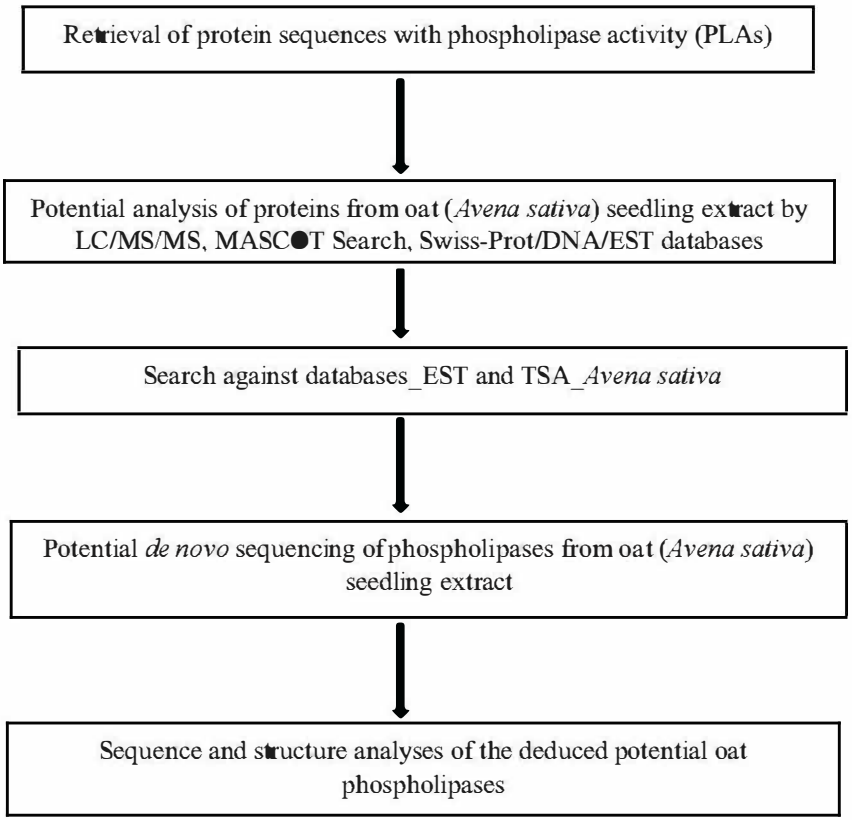


Figure 2. A flow chart depicting the *in silico* steps used in the study for the prediction and analysis of oat phospholipases.

3.3.2 Sequence analysis

Bioinformatic analysis of the *A. sativa* peptide sequences, ESTs, genomic sequences and deduced protein sequences could be performed using the following tools. Multiple sequence alignment could be performed using the ClustalW algorithm (Thompson et al., 1994). The peptide sequences could be compared with the NCBI (National Center for Biotechnology Information, USA) non-redundant sequence databases, the Transcriptome Shotgun Assembly (TSA) *A. sativa* database and the Expressed Sequence Tag (EST) *A. sativa* database using BLAST (Altschul et al., 2005).

Primary structure analysis using predicted amino acid sequences was performed using the ExPASy Proteomics tools. The Translate tool (web.expasy.org/translate/) was used to translate DNA sequences to protein sequences, whereas the Compute pI/Mw tool (web.expasy.org/compute_pi/) was used to compute the theoretical isoelectric point (pI) and molecular mass (Bairoch et al., 2005; Gasteiger et al., 2003). The BioEdit software package (Hall, 1999) was used to manipulate, edit and compare DNA and amino acid sequences.

The prediction of the signal peptide sequence could be performed using the signalP 4.1 application (Petersen et al., 2011).

Phylogenetic analyses could be performed using Molecular Evolutionary Genetics Analysis (MEGA) package version 7 (Kumar et al., 2016). The program MUSCLE (Edgar, 2004), implemented in MEGA7 package, was used to perform multiple alignments of amino acid sequences of oat phospholipases and their homologs for phylogenetic analysis. Evolutionary history could be inferred using the UPGMA method (Sneath and Sokal, 1973). The evolutionary distances were computed using the JTT matrix-based method (Jones et al., 1992) and were in the units of the number of amino acid substitutions per site. All positions containing gaps and missing data were eliminated. The robustness of the inferred tree was evaluated by bootstrap (1000 replications) (Felsenstein, 1985).

3.3.3 Conserved protein motifs analysis and subcellular location prediction

Conserved protein motifs of the protein sequences from oat were analyzed using Multiple Expectation Maximization for Motif Elicitation (MEME) v.5.0.3 (Bailey and Gribskov, 1998) (http://meme-suite.org) with the number of different motifs as 10, motif sites distribution as zero or one occurrence per sequence, and motifs width as 6 (minimum) and 50 (maximum).

The functional annotations of these motifs could be analyzed by InterProScan (http://www.ebi.ac.uk/Tools/pfa/iprscan/) (Finn et al., 2014). The mapping between Pfam (http://pfam.xfam.org) analysis and Gene Ontology (GO) could be provided by InterPro (Sangrador-Vegas et al., 2016).

The prediction on subcellular localization of oat protein was carried out using the CELLO v.2.5 server (http://cello.life.nctu.edu.tw/) (Yu et al., 2006).

3.3.4 Oat PLA_1

The following genomic sequences (TSA_*Avena sativa*: GGDX01164936.1 to GGDX01164941.1) were identified using BLAST (http://blast.ncbi.nlm.nih.gov) (Altschul et al., 1997). These genomic sequences are useful tools for the identification of at least two oat (*Avena sativa*) PLA_1 sequences. In this study, we characterized one protein sequence of *A. sativa* PLA_1 named AsPLA1.

The predicted AsPLA1 cDNA (978 bp) corresponds to a 326 amino acid residue protein (**Figure 3**). The predicted molecular weight of the AsPLA1 is 37.5 kDa and the theoretical isoelectric point (pI) is 6.49. The subcellular localization of **AsPLA1** is mainly plasma membrane with a reliability of 3.226.

As well, the deduced amino acid sequence of AsPLA1 was used to perform searches in the NCBI databases to choose homologous sequences, which were served to predict the conserved motifs (**Figure 4**) as well as to perform multiple sequence alignments and phylogenetic evolution (**Figure 5**).

GO term prediction of the conserved motifs found in AsPLA1 (**Figure 4**) showed the presence of a GO term for molecular function (GO:0008374) for O-acyltransferase activity. Distribution of GO term in the Biological Process category in AsPLA1 revealed a GO term for lipid metabolic process (GO:0006629).

From these analyses (**Figure 4 and Figure 5**), we could conclude that AsPLA1 belongs to the plant PLA_1 family.

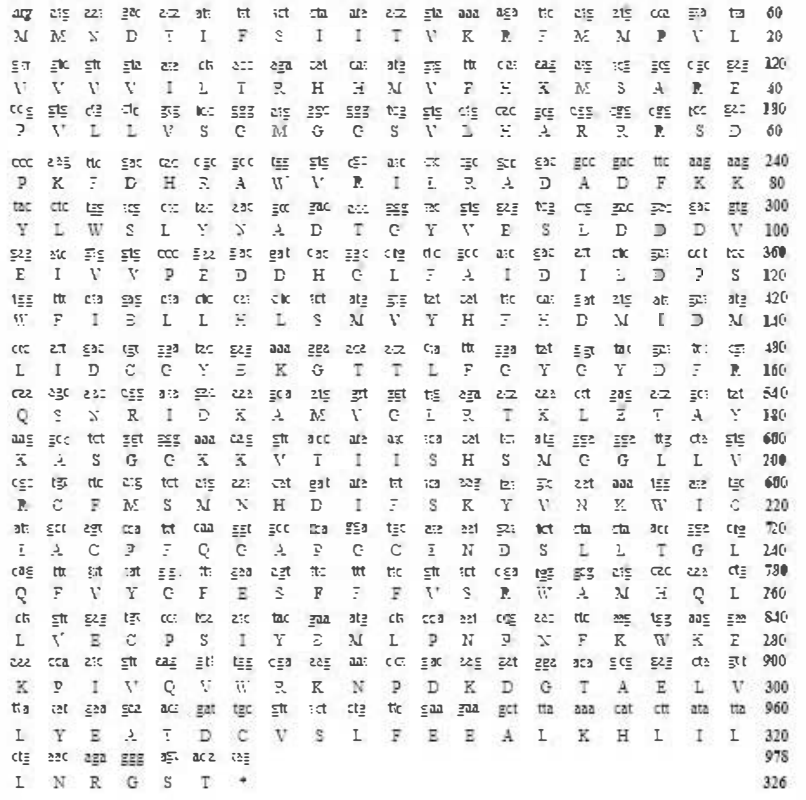


Figure 3. Open reading frame and amino acid sequences of AsPLA1. The amino acid sequence is shown in single code letters beneath the cDNA sequence. Bold italics mark the start codon (atg) and an asterisk indicates the translation stop codon (tag).

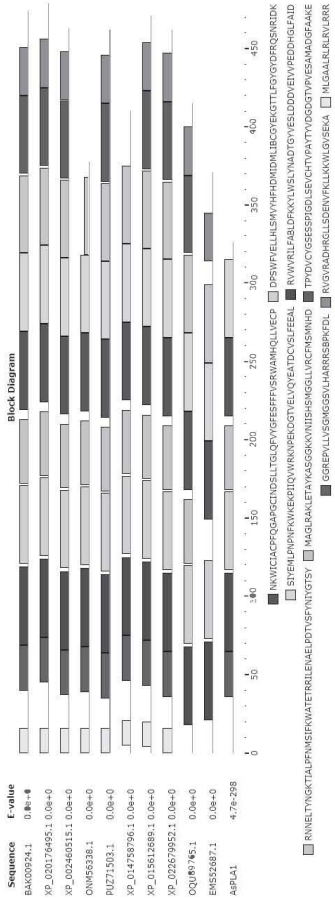


Figure 4. *Predicted structure of PLA1 proteins in A. sativa (AsPLA1), Hordeum vulgare (BAK00924.1), Aegilops tauschii (XP_020176495.1), Sorghum bicolor (XP_002460515.1), Zea mays (ONM56338.1), Panicum hallii (PUZ71503.1), Brachypodium distachyon (XP_014758796.1), Oryza sativa (XP_015612689.1), Setaria italica (XP_022679952.1), Sorghum bicolor (OQU89765.1) and Triticum urartu (EMS52687.1). The conserved motifs of the PLA1 proteins were obtained by the MAST/MEME Version 5.0.3 software. The highly conserved motifs ("DPSWFVELLHLSMVYHFHDMIDMLIBCGYEKGTTLFGYGYDFRQSNRIDK" and "MAGLRAKLETAYKASGGKKVNIISHSMGGLLVRCFMSMNHD") are functionally associated with O-acyltransferase activity (GO:0008374). Their Biological Process could be assigned as lipid metabolic process (GO:0006629).*

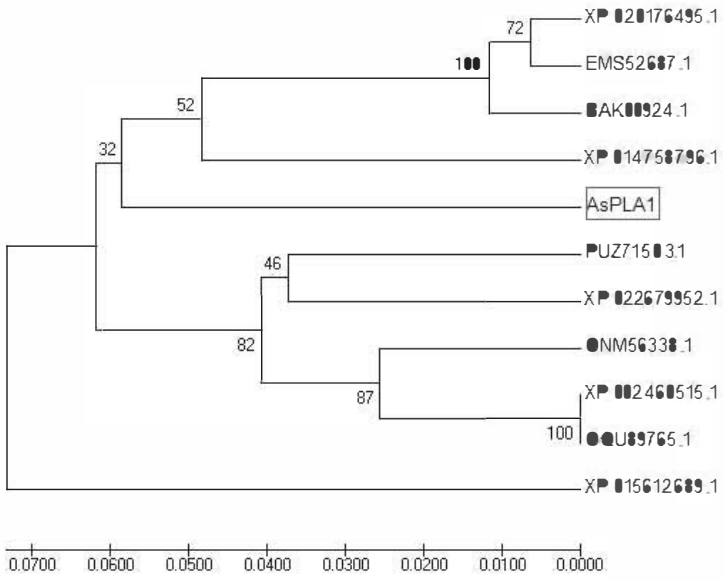


Figure 5. *Evolutionary relationships of taxa related to oat PLA1.*

The evolutionary history was inferred using the UPGMA method. The optimal tree with the sum of branch length = 0.43716168 is shown. The percentages of replicate trees in which the associated taxa clustered together in the bootstrap test (1000 replicates) are shown next to the branches. The tree is drawn to scale, with branch lengths in the same units as those of the evolutionary distances used to infer the phylogenetic tree. The evolutionary distances were computed using the JTT matrix-based method and are in the units of the number of amino acid substitutions per site. The analysis involved 11 amino acid sequences. All positions containing gaps and missing data were eliminated. There were a total of 251 positions in the final dataset. Evolutionary analyses were conducted inMEGA7.

3.3.5 Oat sPLA$_2$

Eight sequences from the TSA_*Avena sativa* database (TSA_*A. sativa*: GGDY01266282.1 to GGDY01266289.1) could be useful tools for the identification of at least two oat PLA2s sequences. In this study, we characterized three protein sequences of *A. sativa* PLA$_2$ named AsPLA2_1, AsPLA2_2, and AsPLA3.

The amino acid sequences of these predicted oat PLA_2s were used to perform searches in the NCBI databases to choose homologous sequences and to perform multiple sequence alignments (**Figure 6**) and phylogenetic evolution (**Figure 7**). From these analyses, we concluded that AsPLA2_2 could originate due to proteolysis of AsPLA2_3; nevertheless, all these three sequences (AsPLA2_1, AsPLA2_2, and AsPLA2_3) could either be truncated or full-length gene sequences as well as they could belong to the plant secretory low-molecular-weight PLA2 ($sPLA_2$) family based on their molecular weights. In fact, plant $sPLA_2$ genes encode proteins, which are generally of low molecular mass (13-18 kDa) after secretion (Wang et al., 2012).

In fact, the deduced amino acid sequences for AsPLA2_1, AsPLA2_2, and AsPLA2_3 were aligned with several plant PLAs and share the main domain and active sites (**Figure 6**). Indeed, similar to other PLA_2 proteins, the conserved catalytic His residues are well conserved (**Figure 6**). This, however, cannot be said about the second residue, which, in catalytically active PLA_2, is usually an Asp residue (Marchler-Bauer et al., 2017). The other four residues of the conserved domain of PLA_2 protein (highlighted in rectangles in **Figure 6**) are residues implicated in the putative metal binding site of PLA_2 enzymes.

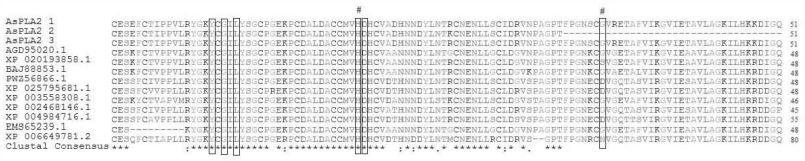


*Figure 6. Multiple amino acid sequences alignment of oat PLA2s (AsPLA2_1, AsPLA2_2, and AsPLA2_3) with homologous PLA2s identified from NCBI databases (http://www.ncbi.nlm.nih.gov/). We have chosen homologues of PLA2, which are as following: Triticum turgidum (AGD95020.1), Aegilops tauschii (XP_020193858.1), Hordeum vulgare (BAJ88853.1), Zea mays (PWZ56866.1), Panicum hallii (XP_025795681.1), Brachypodium distachyon (XP_003558308.1), Sorghum bicolor (XP_002468146.1), Setaria italica (XP_004984716.1), Triticum urartu (EMS65239.1) and Oryza brachyantha (XP_006649781.2). The predicted residues implicated in conserved plant PLA domain are highlighted by rectangles. **Hash marks (#)** above the aligned sequences show the location of the conserved feature residues for putative catalytic residues. The numbers of the residues relative to the first residue of each protein are shown on the right side of each sequence.*

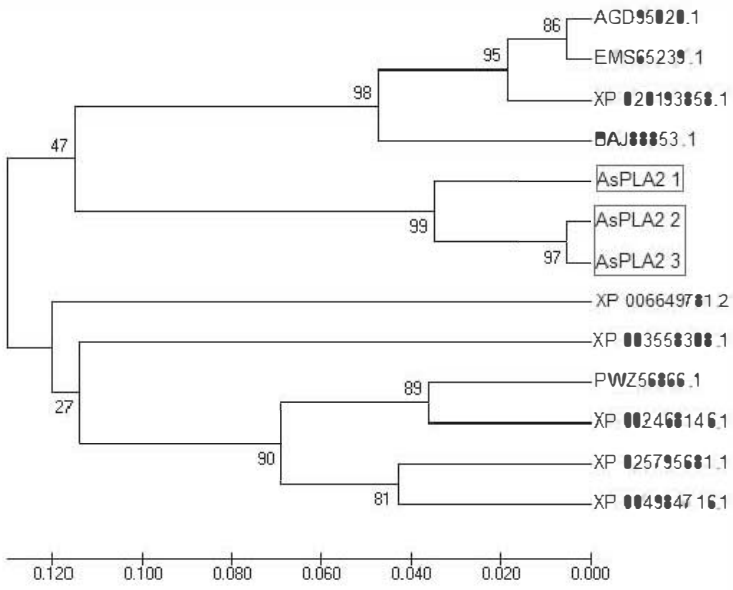


Figure 7. Evolutionary relationships of taxa related to cat PLA2. The evolutionary history was inferred using the UPGMA method. The optimal tree with the sum of branch length = 0.86845666 is shown. The percentages of replicate trees in which the associated taxa clustered together in the bootstrap test (1000 replicates) are shown next to the branches. The tree is drawn to scale, with branch lengths in the same units as those of the evolutionary distances used to infer the phylogenetic tree. The evolutionary distances were computed using the JTT matrix-based method and are in the units of the number of amino acid substitutions per site. The analysis involved 13 amino acid sequences. All positions containing gaps and missing data were eliminated. There were a total of 103 positions in the final dataset. Evolutionary analyses were conducted in MEGA7.

On the other hand, GO term prediction of the conserved motifs found in AsPLA2 (**Figure 8**) showed the presence of two GO terms for the molecular function: (GO:0004623) for phospholipase A2 activity and (GO:0005509) for calcium ion binding. Distribution of GO terms in the Biological Process category in AsPLA2 revealed three GO terms, which are as following: (GO:0006644) for phospholipid metabolic process, (GO:0016042) for lipid catabolic process and (GO:0050482) for arachidonic acid secretion.

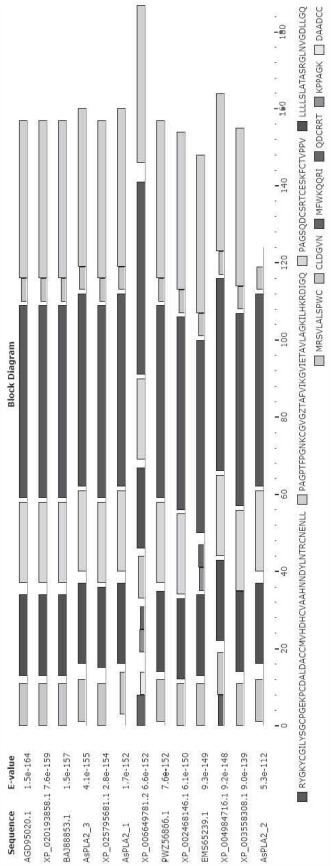


Figure 8. *Predicted structure of PLA2 proteins in A. sativa (AsPLA2 1, AsPLA2 2 and AsPLA2 3), Triticum turgidum (AGD95020.1), Aegilops tauschii (XP 020193858.1), Hordeum vulgare (BAJ88853.1), Zea mays (PWZ56866.1), Panicum hallii (XP 025795681.1), Brachypodium distachyon (XP 003558308.1), Sorghum bicolor (XP 002468146.1), Setaria italica (XP 004984716.1), Triticum urartu (EMS65239.1) and Oryza brachyantha (XP 006649781.2). The conserved motifs of the PLA₂ proteins were obtained by the MAST/MEME Version 5.0.3 software. The highly conserved motifs ("RYGKYCGILYSGCPGEKPCDALDACCMVHDHCVAAHNNDYLNTRCNENLL" and "PAGPTFPGNKCGVGZTAFVIKGVIETAVLAGKILHKRDIGQ") are functionally associated with phospholipase A2 activity (*GO:0004623*) and calcium ion binding (*GO:0005509*). Their biological process could be, essentially, as follows: phospholipid metabolic process (*GO:0006644*), lipid catabolic process (*GO:0016042*) and arachidonic acid secretion (*GO:0050482*).*

AsPLA2_1 was selected for further characterization. Indeed, the predicted AsPLA2_1 cDNA (480 bp) corresponds to a 160 amino acid residue protein (**Figure 9**). The predicted molecular weight of the AsPLA2_1 is 17.454 kDa and the theoretical isoelectric point (pI) is 7.46.

A 28-residue signal peptide was predicted using the Expasy SignalP V4.1 program and the N-terminal sequence of the mature AsPLA2_1 is expected to start at residue L29. The subcellular localization of **AsPLA2_1** is mainly extracellular with a reliability of 4.079.

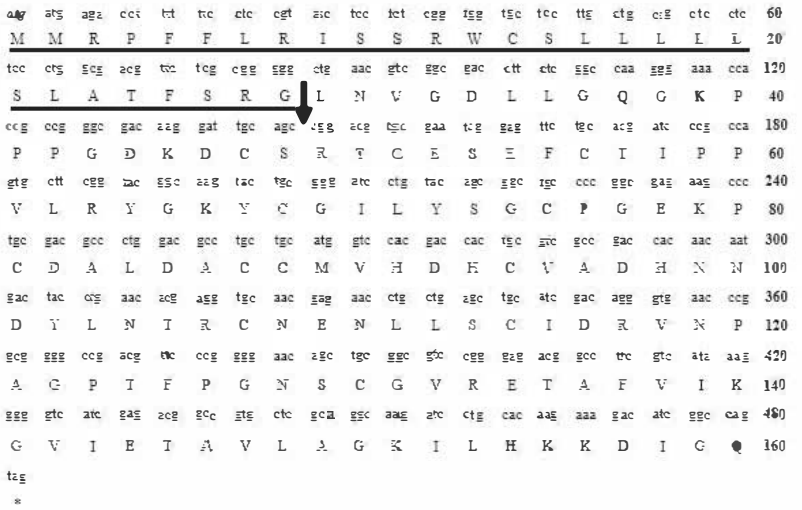


Figure 9. *Open reading frame and amino acid sequences of AsPLA2_1. The amino acid sequence is shown in single code letters beneath the cDNA sequence. Bold italics mark the start codon (**atg**) and an asterisk indicates the translation stop codon (tag). The putative signal peptide is underlined, with the cleavage site SRG/LN indicated by an arrow.*

4. Conclusions

The present chapter highlighted a global idea of plant phospholipases with focus on plant PLAs enzymes by underlining their most important relevant characteristics. A case study of oat (*Avena sativa*) PLAs was also reported. This case study was intended to identify oat PLAs by bioinformatics analyses using the TSA or the EST databases of *Avena sativa* and *in silico* approaches.

The results obtained in this case study on the identification of PLAs from *A. sativa* may be considered a further step in the characterization of such enzymes in plants. The complete sequencing of the *Avena sativa* genome (Gutierrez-Gonzalez et al., 2013) will certainly accelerate the identification of other oat lipolytic enzymes.

5. Acknowledgments

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

The author thanks the Tunisian Ministry of Higher Education and Scientific Research for facilities.

The author is very grateful to Igor Pokotylo, a researcher at the "Université Paris-Est, Institut d'Ecologie et des Sciences de l'Environnement de Paris, Créteil-France" and "Institute of Bioorganic Chemistry and Petrochemistry, National Academy of Sciences of Ukraine, Kiev-Ukraine", for his help reading the manuscript.

6. References

Abdelkafi, S., Abousalham, A., Fendri, I., Ogata, H., Barouh, N., Fouquet, B., Scheirlinckx, F., Villeneuve, P., and Carrière, F. (2012). Identification of a new phospholipase D in *Carica papaya* latex. *Gene* 499, 243–249.

Ackermann, E.J., Kempner, E.S., Dennis, E.A. (1994). Ca(2+)-independent cytosolic phospholipase A2 from macrophage-like P388D1 cells. Isolation and characterization. J. Biol. Chem. 269, 9227–9233.

Altschul S.F., Wootton J.C., Gertz E.M, Agarwala R., Morgulis A., Schaffer A.A., Yu Y.K. (2005). Protein database searches using compositionally adjusted substitution matrices, FEBS J. 272, 5101–5109.

Andrews, D.L. et al. (1988). Characterization of the lipid acyl hydrolase activity of the major potato (*Solanum tuberosum*) tuber protein, patatin, by cloning an abundant expression in a baculovirus vector. Biochem. J. 252, 199–206.

Assmann SM, Shimazaki K. (1999). The multisensory guard cell. Stomatal response to blue light and abscisic acid. *Plant Physiol.* 119, 809–15.

Bahn SC, Lee HY, Kim HJ et al. (2003). Characterization of *Arabidopsis* secretory phospholipase A2-gamma cDNA and its enzymatic properties. FEBS Lett. 553: 113–118.

Bailey, T.L., and Gribskov, M. (1998). Combining evidence using p-values: application to sequence homology searches. Bioinformatics. 14, 48-54.

Bairoch A., R. Apweiler, C.H. Wu, W.C. Barker, B. Boeckmann, S. Ferro, E. Gasteiger, H. Huang, R. Lopez, M. Magrane, M.J.Martin, D.A. Natale, C. O'Donovan, N. Redaschi, L.S. Yeh. (2005). The universal protein resource (UniProt), Nucleic Acids Res. 33, 154–159.

Balsinde, J. and Balboa, M.A. (2005). Cellular regulation and proposed biological functions of group VIA calcium-independent phospholipase A2 in activated cells. Cell Signal. 17, 1052–1062.

Barh, D. et al. (eds.). (2015). *PlantOmics: The Omics of Plant Science.* Springer India.

Barros, M., Fleuri, L.F., Macedo, G.A. (2010). Seed lipases: sources, applications and properties – a review. Braz. J. Chem. Eng. 27, 15–29.

Beisson, F., Koo, A.J.K., Ruuska, S. et al. (2003). *Arabidopsis* genes involved in acyl lipid metabolism. A 2003 census of the candidates, a study of the distribution of expressed sequence tags in organs, and a web-based database. Plant Physiology, 132, 681–697.

Ben Halima, N. (2019). *Unique Sequence Signatures in Plant Lipolytic Enzymes: Emerging Research and Opportunities.* Hershey, PA: IGI Global.

Ben Halima, N., Ben Saad, R., Khemakhem, B., Fendri, I. and Abdelkafi, S. (2015). Oat (*Avena sativa* L.): Oil and nutriment compounds valorization for potential use in industrial applications. *Journal of Oleo Science,* 64, 915-932.

Bennett, M.D. and Smith, J.B. (1976). Nuclear DNA amounts in angiosperms. Philos. Trans. R. Soc. Lond. B 274, 227–274.

Borrelli, G.M., and Trono, D. (2015). Recombinant lipases and phospholipases and their use as biocatalysts for industrial applications. Int. J. Mol. Sci. 16, 20774-20840.

Brockerhoff, H., & Jensen, R. G. (1974). *Lipolytic enzymes.* The United Kingdom/London: Academic Press.

Casado, V., Martín, D., Torres, C., and Reglero, G. (2012). Phospholipases in food industry: A review. Methods in Molecular Biology. 861, 495-523.

Chapman KD. (1998). Phospholipase activity during plant growth and development and in response to environmental stress. *Trends Plant Sci.* 11, 419–26.

Chen, G., Snyder, C.L., Greer M.S., and Weselake, R.J. (2011). Biology and biochemistry of plant phospholipases. *Critical Reviews in Plant Sciences,* 30:3, 239-258.

Danmjanovic, J., and Iwasaki, Y. (2013). Phospholipase D as a catalyst: Application in phospholipid synthesis, molecular structure and protein engineering. J. Biosci. Bioeng. xx:1-10.

De Maria, L., Vind, J., Oxenboll, K. M., Svendsen, A., and Patkar, S. (2007). Phospholipases and their industrial applications. *Appl. Microbiol. Biot.* 74, 290–300.

Dessen, A. (2000). Structure and mechanism of human cytosolic phospholipase A2. Biochim. Biophys. Acta 1488, 40–47.

Dhondt S, Gouzerh G, Muller A et al. (2002). Spatio-temporal expression of patatin-like lipid acyl hydrolases and accumulation of jasmonates in elicitor-treated tobacco leaves are not affected by endogenous levels of salicylic acid. Plant J. 32, 749–762.

Doolittle, M., & Reue, K. (Eds.). (1999). *Lipase and phospholipase protocols.* Humana Press.

Dörmann, P. (2005). Membrane lipids. In Murphy, D.J. (Ed.), *Plant Lipids. Biology, Utilisation and Manipulation* (pp. 123-161). Blackwell Publishing. CRC Press.

Drobak BK, Dewey RE, Boss WF. (1999). Phosphoinositide kinase and the synthesis of polyphosphoinositide in higher plant cells. *Int. Rev. Cytol.* 189, 95–130.

Edgar RC. (2004). MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res. 32, 1792–7.

Ellinger D, Stingl N, Kubigsteltig II et al. (2010). DONGLE and DEFECTIVE IN ANTHER DEHISCENCE1 lipases are not essential for wound- and pathogen-induced jasmonate biosynthesis: redundant lipases contribute to jasmonate formation. Plant Physiol. 153:114–127.

Farag, K.M. and Palta, J.P. (1993). Use of lysophosphatidylethanolamine, a natural lipid, to retard tomato leaf and fruit senescence. *Physiologia Plantarum,* 87, 515–521.

Felsenstein J. (1985). Confidence limits on phylogenies: An approach using the bootstrap. Evolution. 39, 783-791.

Finn RD, Bateman A, Clements J, Coggill P, Eberhardt RY, Eddy SR, Heger A, Hetherington K, Holm L, Mistry J, et al. (2014). Pfam: the protein families database. Nucleic Acids Res. 42, D222–230.

Froidure S, Canonne J, Daniel X et al. (2010). AtsPLA2-alpha nuclear relocalization by the *Arabidopsis* transcription factor AtMYB30 leads to repression of the plant defense response. Proc. Natl. Acad. Sci. USA 107, 15281–15286.

Galliard T. (1971). The enzymic deacylation of phospholipids and galactolipids in plants. Purification and properties of a lipolytic acylhydrolase from potato tubers. Biochem J. 121:379–390.

Gasteiger, E. Gattiker, A. Hoogland, C. Ivanyi, I. Appel, R.D. Bairoch, A. (2003). ExPASy: the proteomics server for in-depth protein knowledge and analysis. Nucl. Acids Res. 31, 3784-3788.

Gunstone, F.D. (1999). Enzymes as biocatalysts in the modification of natural lipids. J. Sci. Food Agric. 79, 1535–1549.

Gutierrez-Gonzalez, J.J., Tu, Z.J. and Garvin, D.F. (2013). Analysis and annotation of the hexaploid oat seed transcriptome. BMC Genomics 14,471.

Hall, T.A. (1999). BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT, Nucl. Acids Symp. Ser. 41, 95-98.

Higgs HN, Han MH, Johnson GE et al. (1998). Cloning of a phosphatidic acid-preferring phospholipase A1 from bovine testis. J. Biol. Chem. 273, 5468–5477.

Holk A, Rietz S, Zahn M et al. (2002). Molecular identification of cytosolic, patatin-related phospholipases A from *Arabidopsis* with potential functions in plant signal transduction. Plant Physiol. 130, 90–101.

Hong Y, Wang TW, Hudak KA et al. (2000). An ethylene-induced cDNA encoding a lipase expressed at the onset of senescence. Proc. Natl. Acad. Sci. USA 97, 8717–8722.

Hong, Y., Zhao, J., Guo, L., Kim, S.C., Deng, X., Wang, G., Zhang, G., Li, M., and Wang, X. (2016). Plant Phospholipases D and C and their diverse functions in stress responses. Progress in Lipid Research. 62, 55-74.

Huang S, Cerny RE, Bhat DS et al. (2001). Cloning of an *Arabidopsis* patatin-like gene, STURDY, by activation T-DNA tagging. Plant Physiol. 125, 573–584.

Hyun Y, Choi S, Hwang HJ et al. (2008). Cooperation and functional diversification of two closely related galactolipase genes for jasmonate biosynthesis. Dev. Cell 14, 183–192.

Ishiguro, S., Kawai-Oda, A., Ueda, J., Nishida, I., and Okada, K. (2001). The *DEFECTIVE IN ANTHER DEHISCIENCE1* gene encodes a novel phospholipase A1 catalyzing the initial step of jasmonic acid biosynthesis, which synchronizes pollen maturation, anther dehiscence, and flower opening in *Arabidopsis*. *Plant Cell 13*, 2191–2209.

Jones D.T., Taylor W.R., Thornton J.M. (1992). The rapid generation of mutation data matrices from protein sequences. Comput. Appl. Biosci. 8, 275–282.

Kato, T., Morita, M.T., Fukaki, H. et al. (2002). SGR2, a phospholipase-like protein, and ZIG/SGR4, a SNARE, are involved in the shoot gravitropism of *Arabidopsis*. Plant Cell, 14, 33–46.

Khatoon, H., Mansfeld, J., Schierhorn, A., and Ulbrich-Hofmann, R. (2015). Purification, sequencing and characterization of phospholipase D from Indian mustard seeds. Phytochemistry 117, 65–75.

Kim, D.K., Lee, H.J., Lee, Y. (1994). Detection of two phospholipase A2 (PLA2) activities in leaves of higher plant *Vicia faba* and comparison with mammalian PLA2's. FEBS Lett. 343:213–218.

Kim, J.Y., Chung, Y.S., Ok, S.H., et al. (1999). Characterization of the full-length sequences of phospholipase A2 induced during flower development. Biochim Biophys Acta. 1489, 389–392.

Kim, H.J., Ok, S.H., Bahn, S.C., et al. (2011). Endoplasmic reticulum- and Golgi-localized phospholipase A2 plays critical roles in *Arabidopsis* pollen development and germination. Plant Cell 23, 94–110.

Kumar, S., Stecher, G., and Tamura, K. (2016). MEGA7: Molecular Evolutionary Genetics Analysis version 7.0 for bigger datasets. Molecular Biology and Evolution. 33, 1870-1874.

La Camera, S., Balague, C., Gobel, C., et al. (2009). The Arabidopsis patatin-like protein 2 (PLP2) plays an essential role in cell death execution and differentially affects biosynthesis of oxylipins and resistance to pathogens. Mol. Plant Microbe Interact. 22, 469–481.

La Camera, S., Geoffroy, P., Samaha, H., et al. (2005). A pathogen-inducible patatin-like lipid acyl hydrolase facilitates fungal and bacterial host colonization in *Arabidopsis*. Plant J. 44, 810–825.

Lee, H.Y., Bahn, S.C., Kang, Y.-M., et al. (2003). Secretory low molecular weight phospholipase A2 plays important roles in cell elongation and shoot gravitropism in *Arabidopsis*. Plant Cell, 15, 1990–2002.

Lee, H.Y., Bahn, S.C., Shin, J.S. et al. (2005). Multiple forms of secretory phospholipase A2 in plants. Prog. Lipid Res. 44, 52–67.

Lee, O.R., Kim, S.J., Kim, H.J., et al. (2010). Phospholipase A(2) is required for PINFORMED protein trafficking to the plasma membrane in the *Arabidopsis* root. Plant Cell. 22, 1812–1825.

Li, M., Bahn, S.C., Guo, L., et al. (2011). Patatin-related phospholipase pPLAIII β–induced changes in lipid metabolism alter cellulose content and cell elongation in *Arabidopsis*. Plant Cell 23, 1107–1123.

Liao HL, Burns JK. (2010). Light controls phospholipase A2alpha and beta gene expression in *Citrus sinensis*. J. Exp. Bot. 61, 2469–2478.

Liscovitch, M., Czarny, M., Fiucci, G, and Tang, X. (2000). Phospholipase D: molecular and cell biology of a novel gene family. Biochem. J. 345:401–15.

Mansfeld, J. (2009). Plant phospholipases A2: perspectives on biotechnological applications. Biotechnol. Lett. 31, 1373–1380.

Mansfeld J, Ulbrich-Hofmann R. (2007). Secretory phospholipase A2-alpha from *Arabidopsis thaliana*: functional parameters and substrate preference. Chem. Phys. Lipids 150, 156–166.

Marchler-Bauer A. et al. (2017). CDD/SPARCLE: functional classification of proteins via subfamily domain architectures. Nucleic Acids Res. 45, 200-203.

Matlashewski, G.J., Urquhart, A.A., Sahasra-dudhe, M.R. and Altosaar, I. (1982). Lipase activity in oat flour suspensions and soluble extracts. Cereal Chem. 59, 418-422.

Matos AR, d'Arcy-Lameta A, Franca M et al. (2001). A novel patatin-like gene stimulated by drought stress encodes a galactolipid acyl hydrolase. FEBS Lett. 491, 188–192.

Matos AR, Gigon A, Laffray D et al. (2008). Effects of progressive drought stress on the expression of patatin-like lipid acyl hydrolase genes in *Arabidopsis* leaves. Physiol Plant 134, 110–120.

Matos, A.R., and Pham-Thi, A.T. (2009). Lipid deacylating enzymes in plants: old activities, new genes. Plant. Physiol. Biochem. 47, 491–503.

Morita MT, Kato T, Nagafusa K et al. (2002). Involvement of the vacuoles of the endodermis in the early process of shoot gravitropism in *Arabidopsis*. Plant Cell 14, 47–56.

Müller, G., & Petry, S. (Eds.). (2004). *Lipases and Phospholipases in Drug Development: From biochemistry to molecular pharmacology*. Weinheim, Germany: WILEY-VCH Verlag GmbH & Co. KGaA.

Munnik T, Irvine RF, Musgrave AP. (1998). Phospholipid signalling in plants. *Biochim. Biophys. Acta* 1389, 222–72.

Murphy, D.J. (Ed.). (2005). *Plant Lipids. Biology, Utilisation and Manipulation*. Blackwell Publishing. CRC Press.

Nakamura, Y., Awai, K., Masuda, T., et al. (2005). A novel phosphatidylcholine-hydrolyzing phospholipase C induced by phosphate starvation in *Arabidopsis*. Journal of Biological Chemistry. 280, 7469–7476.

Novotna, Z., Valentova, O., Martinec, J., Feltl, T., and Nokhrina, K. (2000). Study of phospholipases D and C in maturing and germinating seeds of *Brassica napus*. Biochem Society Transactions. 28, 817-8.

O'Luanaigh N, Pardo R, Fensome A et al. (2002). Continual production of phosphatidic acid by phospholipase D is essential for antigen-stimulated membrane ruffling in cultured mast cells. Mol. Biol. Cell 13, 3730–3746.

Peng, Y., Zhang, J., Cao, G. et al. (2010). Overexpression of a PLDalphal gene from *Setaria italica* enhances the sensitivity of *Arabidopsis* to abscisic acid and improves its drought tolerance. Plant Cell Rep. 29, 793–802.

Perez-Torres CA, Lopez-Bucio J, Cruz-Ramirez A et al. (2008). Phosphate availability alters lateral root development in *Arabidopsis* by modulating auxin sensitivity via a mechanism involving the TIR1 auxin receptor. Plant Cell 20, 3258–3272.

Peters C, Li M, Narasimhan R et al. (2010). Nonspecific phospholipase C NPC4 promotes responses to abscisic acid and tolerance to hyperosmotic stress in *Arabidopsis*. Plant Cell. 22, 2642–2659.

Petersen, T.N., Brunak, S., von Heijne, G., Nielsen, H. (2011). SignalP 4.0: discriminating signal peptides from transmembrane regions. Nature Methods. 8, 785-786.

Plaxton, W.C., and Lambers, H. (Eds.). (2015). *Phosphorus Metabolism in Plants*. Annual Plant Reviews, Volume 48. John Wiley & Sons, Ltd, The Atrium, Southern Gate, Chichester, West Sussex, PO19 8SQ, UK.

Pokotylo, I., Kolesnikov, Y., Kravets, V., Zachowski, A., and Ruelland, E. (2014). Plant phosphoinositide-dependent phospholipases C: Variations around a canonical theme. Biochimie 96, 144-157.

Pokotylo, I., Pejchar, P., Potocký, M., Kocourková, D., Krčková, Z., Ruelland, E., Kravets, V., and Martinec, J. (2013). The plant non-specific phospholipase C gene family. Novel competitors in lipid signaling. Prog. Lipid Res. 52, 62–79.

Ramnath, L., Sithole, B., and Govinden. R. (2017). Classification of lipolytic enzymes and their biotechnological applications in the pulping industry. Can. J. Microbiol. 63, 179–192.

Reina-Pinto JJ, Voisin D, Kurdyukov S et al. (2009). Misexpression of FATTY ACID ELONGATION1 in the *Arabidopsis* epidermis induces cell death and suggests a critical role for phospholipase A2 in this process. Plant Cell 21:1252–1272.

Rhee SG, Bae YS. (1997). Regulation of phosphoinositide-specific phospholipase C isozymes. *J. Biol. Chem.* 272, 15045-48.

Rietz S, Dermendjiev G, ●ppermann E et al. (2●1●). Roles of *Arabidopsis* patatin-related phospholipases a in root development are related to auxin responses and phosphate deficiency. Mol. Plant 3, 524–53●.

Rivera, I., Mateos-Díaz, JC. & Sandoval, G. (2●12). Plant lipases: partial purification of *Carica papaya* lipase. Methods in Molecular Biology, ●61, 115-22.

Rosahl, S. et al. (1●●6). Isolation and characterization of a gene from *Solanum tuberosum* encoding patatin, the major storage protein of potato tubers. Mol. Gen. Genet. 2●3, 214–22●.

Ryu, S.B. (2●●4). Phospholipid-derived signaling mediated by phospholipase A in plants. Trends Plant Sci. ●, 22●–235.

Ryu, S.B., Lee, H.J., Hwang, I.W., et al. (2●●●). Transgenic plants exhibiting increased resistance to biotic and abiotic stresses or accelerated flowering time by overexpression of a secretory phospholipase A_2. US 12/527,323 United State Patent and Trademark ●ffice.

Sahsah, Y., Campos, P., Gareil, M., et al. (1●●8). Enzymatic degradation of polar lipids in *Vigna unguiculata* leaves and influence of drought stress. Physiol. Plant 1●4, 577–58●.

Sandoval, G. (Ed.). (2●12). *Lipases and Phospholipases: Methods and Protocols*. New York, USA: Humana Press.

Sangrador-Vegas, A., Mitchell, A.L., Chang, H.Y., Yong, S.Y., and Finn, R.D. (2●16). G● annotation in InterPro: why stability does not indicate accuracy in a sea of changing annotation. Database, 1–8.

Scherer, G.F., Ryu, S.B., Wang, X., Matos, A.R., and Heitz, T. (2●1●). Patatin-related phospholipase A: Nomenclature, subfamilies and functions in plants. Trends Plant Sci. 5, 6●3–7●●.

Selvy, P.E., Lavieri, R.R., Lindsley, C.W., and Brown, H.A. (2●11). Phospholipase D: Enzymology, functionality, and chemical modulation. Chem. Rev. 111, 6●64–611●.

Seo J, Lee HY, Choi H et al. (2●●8). Phospholipase A2beta mediates light-induced stomatal opening in *Arabidopsis*. J. Exp. Bot. 5●, 3587–359●4.

Seo YS, Kim EY, Kim JH et al. (2●●●). Enzymatic characterization of class I DAD1-like acylhydrolase members targeted to chloroplast in *Arabidopsis*. FEBS Lett 583, 23●1–23●7.

Servi, S. (1●●●). Phospholipases as synthetic catalysts. In: *Biocatalysis – from Discovery to Application*. pp. 127–158. Fessner, W.-D. Ed., Springer, New York.

Six, D.A. and Dennis, E.A. (2●●●). The expanding superfamily of phospholipase A2 enzymes: classification and characterization. Biochim. Biophys. Acta 1488, 1–1●.

Sneath P.H.A and Sokal R.R. (1973). Numerical Taxonomy. Freeman, San Francisco.

Ståhl, U., Ek, B., and Stymne, S. (1998). Purification and characterization of a low-molecular-weight phospholipase A2 from developing seeds of elm. Plant Physiology, 117, 197–205.

Ståhl, U., Lee, M., Sjödahl, S., et al. (1999). Plant low-molecular-weight phospholipase A2s (PLA2s) are structurally related to the animal secretory PLA2s and are present as a family of isoforms in rice (*Oryza sativa*). Plant Molecular Biology, 41, 481–490.

Stevenson, J.M., Perera, I.Y., Heilmann, I., Persson, S., Boss, W.F. (2000). Inositol signaling and plant growth. *Trends Plant Sci.* 5, 252–58.

Thompson, J.D., Higgins, D.G., Gibson, T.J. (1994). ClustalW, improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position specific gap penalties and weight matrix choice. Nucleic Acids Res. 22, 4673–4680.

Thompson, J., Taylor, C., Wang, T.W. (2000). Altered membrane lipase expression delays leaf senescence. Biochem. Soc. Trans. 28:775–777.

Verlotta, A., Liberatore, M.T., Cattivelli, L., and Trono, D. (2013). Secretory phospholipases A2 in durum wheat (*Triticum. durum* Desf.): Gene expression, enzymatic activity, and relation to drought stress adaptation. Int. J. Mol. Sci. 14, 5146–5169.

Vossen, J.H., Abd-El-Haliem, A., Fradin, E.F., et al. (2010). Identification of tomato phosphatidylinositol-specific phospholipase-C (PI-PLC) family members and the role of PLC4 and PLC6 in HR and disease resistance. Plant J. 62, 224–239.

Wang, X. (1999). The role of phospholipase D in signaling cascade. *Plant Physiol.* 120, 645–51.

Wang, X. (2000). Multiple forms of phospholipase D in plants: the gene family, catalytic and regulatory properties, and cellular functions. *Prog. Lipid Res.* 39, 109–49.

Wang, X. (2001). Plant phospholipases. Annu. Rev. Plant Physiol. Plant Mol. Biol. 52, 211-231.

Wang, X. (2004). Lipid signaling. Curr. Opin. Plant Biol. 7, 329–336.

Wang, G., Ryu, S. & Wang, X. (2012). Plant phospholipases: an overview. Methods in Molecular Biology, 861, 123-37.

Williams, R.L. (1999). Mammalian phosphoinositide-specific phospholipase C. Biochim. Biophys. Acta 1441, 255–67.

Yamaguchi, T., Kuroda, M., Yamakawa, H., et al. (2009). Suppression of a phospholipase D gene, OsPLDbeta1, activates defense responses and increases disease resistance in rice. Plant Physiol. 150, 308–319.

Yan, H., Lu, D., Rivkees, S.A. (2003). Lysophosphatidic acid regulates the proliferation and migration of olfactory ensheathing cells in vitro. Glia 44, 26–36.

Yang, W., Devaiah, S.P., Pan, X., et al. (2007). AtPLAI is an acyl hydrolase involved in basal jasmonic acid production and *Arabidopsis* resistance to *Botrytis cinerea*. J. Biol. Chem. 282, 18116–18128.

Yu, C.S., Chen, Y.C., Lu, C.H., Hwang, J.K. (2006). Prediction of protein subcellular localization. Proteins: Structure, Function and Bioinformatics. 64, 643-651.

Zhang, T. T., Song, Y. Z., Liu, Y. D., Guo, X. Q., Zhu, C. X., and Wen, F. J. (2008). Overexpression of phospholipase Dalpha gene enhances drought and salt tolerance of *Populus tomentosa*. *Chinese Sci. Bull.* 53, 3656–3665.

7. Additional Reading

Brown, H.A. (Ed.). (2007). *Lipidomics and bioactive lipids: Lipids and cell signaling*. Academic Press, Elsevier.

Burke, J.E., and Dennis, E.A. (2009). Phospholipase A_2 structure/function, mechanism, and signaling. Journal of Lipid Research. S237-S242.

Casas-Godoy, L., Duquesne, S., Bordes, F., Sandoval, G. & Marty, A. (2012). Lipases: an overview. Methods in Molecular Biology. 861, 3-30.

Helmy, M., Lombard, S., and Piéroni, G. (1999). Ricin RCA60: Evidence of its phospholipase activity. Biochem. Biophys. Res. Commun. 258, 252–255.

Inoue A., and Aoki, J. (2006). Phospholipase A_1: structure, distribution and function. Future Lipidology. 1:6, 687-700.

Kim, D.K., Lee, H.J., and Lee, Y. (1994). Detection of two phospholipase A2 (PLA_2) activities in leaves of higher plant *Vicia faba* and comparison with mammalian PLA2s. FEBS Lett. 343, 213–218.

Murphy, D.J. (Ed.). (2005). *Plant Lipids. Biology, Utilisation and Manipulation*. Blackwell Publishing. CRC Press.

Richmond, G.S., and Smith, T.K. (2011). Phospholipases A_1. Int. J. Mol. Sci. 12, 588-612.

Senda, K., Yoshioka, H., Doke, N., and Kawakita, K. (1996). A cytosolic phospholipase A2 from potato tissues appears to be patatin. *Plant Cell Physiol.* 37, 347–353.

Williams, J.P., Khan, M.U., and Lem, N.W. (Eds.). (1997). *Physiology, biochemistry and molecular biology of plant lipids*. Kluwer Academic Publishers.

CHAPTER FOUR

IMPORTANCE OF LIPIDS FROM PLANTS: PROMISING INSIGHTS INTO ALGAE

Abstract

Plant lipids are interesting biomolecules due to their multiple uses such as in nutrition and energy storage. Lipids originating from algae, especially microalgae are very promising into eventual biotechnological applications. Indeed, algae, i.e., macro-algae and microalgae, are a diverse group of organisms that inhabit a vast range of ecosystems from the Antarctic to the deserts. Algae are unique compared to other living organisms thanks to their valuable biochemical compounds, for instance, lipids, to be used in many biotechnological applications such as in aquaculture, biofuels, human health food, and other applications. In this chapter, updated knowledge of lipids and recent research studies for algae's valorization performed by several authors were reviewed.

Keywords: ●leaginous algae, Biotechnology, Applications, Lipids, Valorization.

1. Introduction

In its large concept, biotechnology is understood as "technologies supported by living beings". By such definition, since over thousands of years ago, biotechnological processes were used by mankind to, for example, bread making, brewing, and animal breeding. These processes were being used for a long time without knowing the scientific basis (e.g., genetic and fermentation theories, etc.). The ●rganization for Economic Co-operation and Development (●ECD) described biotechnology as "the application of science and technology to living organisms as well as parts, products, and models thereof, to alter living or non-living materials for the production of knowledge, goods and services" (●ECD, 2●●5). Therefore, biotechnology would be used in multiple productive activities, which

generate a significant impact on many fields (new and old areas) of exchange at technological, productive, scientific, and social levels. To facilitate the broadest conception of biotechnology, each of their fields was associated with a specialty coding by colors according to a methodology proposed by DaSilva (2004). This approach which categorized biotechnology by colors was referred to as "Rainbow" summarized in **Table 1**.

Table 1. Colors describing the different fields of biotechnology (DaSilva 2004)

Color type	Area of Biotech Activities
Red	Health, Medical, Diagnostics
Yellow	Food Biotechnology, Nutrition Science
Blue	Aquaculture, Coastal, and Marine Biotech
Green	Agriculture, Environmental Biotechnology-Biofuels, Biofertilizers, Bioremediation, Geomicrobiology
Brown	Arid Zone and Desert Biotechnology
Dark	Bioterrorism, Biowarfare, Biocrimes, Anticrop warfare
Purple	Patents, Publications, Inventions, IPRs
white	Gene-based Bioindustries
Gold	Bioinformatics, Nanobiotechnology
Grey	Classical Fermentation and Bioprocess Technology

Hence, living organisms are key factors for further investigations in such biotechnological domains. Worthy of note is the fact that evolutionary relationships among living organisms remain uncertain because of a lack of evidence from fossil for some taxa, which can lead to many controversies. The proposed phylogenetic relationships among Bacteria, Animalia, Plantae, Algae and Fungi kingdoms depend on the method used and the molecular regions. For example, the phylogenetic analysis in **Figure 1** had shown that the fungal kingdom is part of the great eukaryotic groups in the terminal radiation, which has occurred some one billion years ago (Sogin, 1989; Taylor et al., 1994; Guarro et al., 1999).

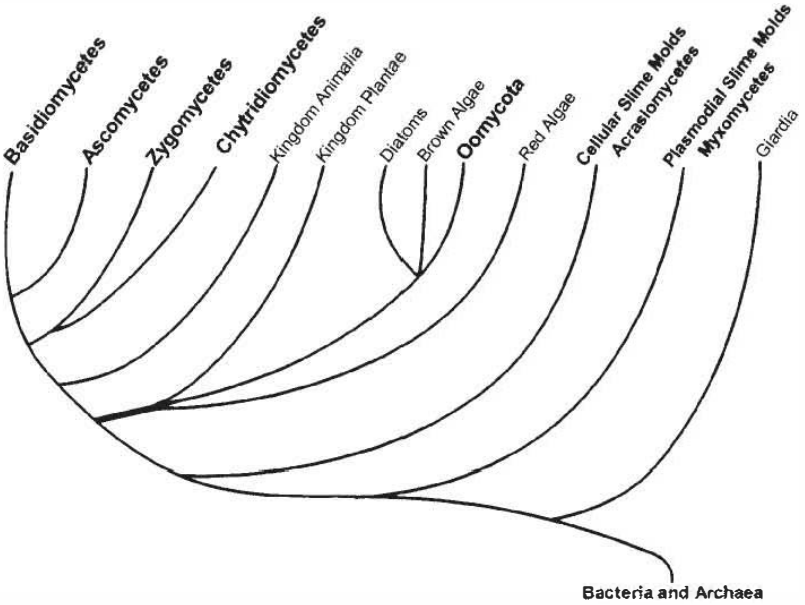


Figure 1. *Phylogenetic tree showing relationships of eukaryotes, based on the nucleotide sequence of 18S DNA. Reprinted from Taylor et al. (1994).*

Moreover, plants and algae are important constituents of diverse ecosystems extending from freshwater and marine environments to dry deserts and from hot to cold climates. Special attention will be highlighted through this review to algae, which account for more than half of the total primary production regarding the food chain worldwide (Van den Hoek et al., 1995). Obviously, algae and microalgae are a promising and very diverse group of organisms, which can be unicellular or pluricellular, closely related to Plantae and in some cases to bacteria when talking about Cyanobacteria (blue-green algae). Furthermore, living cells need metabolites or nutrients involved in their metabolism. The often important nutrients are lipids and their derivatives which are organic compounds are part of the primary metabolic processes in the living organisms due to their direct implication in general maintenance of the cell systems, reproduction, growth, etc.

This review intends to highlight the different structures of lipids from various origins and in particular the ability to select oleaginous algae as potential sources of industrial applications.

2. Fatty acids and lipids: generalities

In general, fatty acids (FAs) are aliphatic, monocarboxylic acids with straight chains. The most natural FAs have chain lengths between C_4 and C_{22}. The structure of these FAs is reflected by their common biosynthesis. In fact, the chain is built in two units; carbon and double bonds (cis) can be inserted at their specific positions relative to carboxyl carbons. Hence, many types of FAs can be found with different chain lengths, configurations, positions and types of unsaturation, as well as additional substituents along the aliphatic chain. Moreover, saturated FAs would be obtained by hydrogenation of the more readily available unsaturated acids. However, the most widely spread FAs in nature are palmitic (C16:0), oleic (C18:1) and linoleic (C18:2) acids which can make up 80 % of commodity fats and oils (http://sofa.mri.bund.de/). Many other kinds of FAs will occur as conjugated acids, trans acids, branched chain acids, or cyclic FAs. Unusual FAs structures could be found from some extremophile bacteria, which were able to produce FAs with a terminal cyclohexyl group (Oshima and Agriga, 1975). Also, a terminal concatenated cyclobutane occurred as glycerol and methyl esters were found in unusually dense membranes of anammox bacteria (Damste et al., 2002). Interestingly, the production of polyunsaturated FAs (PUFAs) is intensively researched and increased in commercial attention because of their essential role in maintaining steady health. In addition, it is important to know whether these PUFAs are present as part of triacyglycerols in the cytosol, or if they are present within membrane lipids (phospho- or glyco- lipids) (Sijtsma and de Swaaf, 2004). The important PUFAs could be the very long-chain PUFAs of the ω-3 family, e.g., eicosapentaenoic acid (C20:5n-3 (EPA)) and docosahexaenoic acid (C22:6n-3 (DHA)). Thus, it is noteworthy that identification and characterization of enzymes involved in the biosynthesis of such PUFAs in organisms such as mammals or algae has a biotechnological significance due to their promising effects for the production of the very long-chain PUFAs in transgenic plants i.e., transgenic oilseed crops (Meyer et al., 2004; Domergue et al., 2005; Drexler et al., 2003). The pathway for DHA biosynthesis in mammals and algae is represented in **Figure 2**.

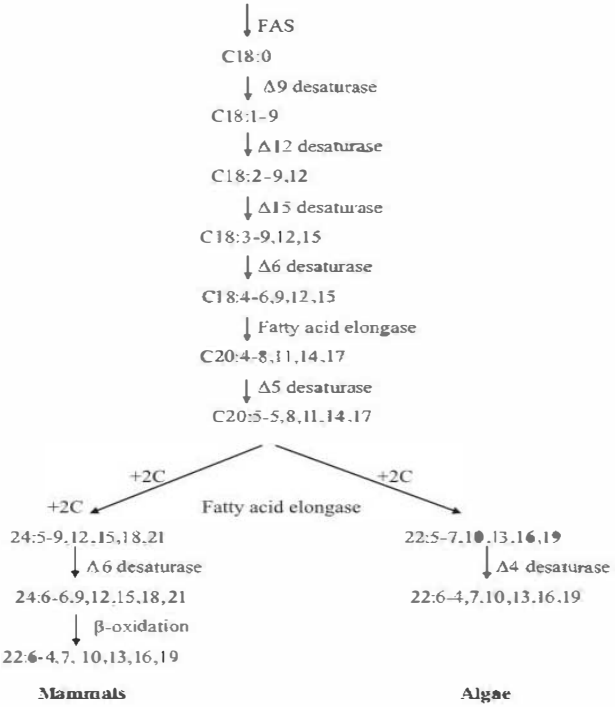


Figure 2. Biosynthesis of DHA by the aerobic pathway in mammals and algae (Guschina and Harwood, 2006).

Various natural lipids exist, a wide range of which are FAs as the major constituents. Lipids are apolar compounds, i.e., are insoluble in water and can be extracted with organic solvents such as hexane or chloroform. Lipids are biomolecules with most probably the highest structural variability and the extraction of lipids from a given sample is, therefore, the first step of lipid analysis (Carrapiso and Garcia, 2000). These biomolecules are involved in the energy storage, especially for fat tissue, in nutrition and are also of relevance for signal transduction processes, for instance, apoptosis, cell differentiation and phagocytosis (Yeung et al., 2006). Lipids structure could add information about the occurrence of lipids that cover important materials such as:

-*Glycerides* (Acylglycerols): which are composed of esters of glycerol and FAs, and triacylglycerols (TAGs) are major constituents of the most natural oils and fats. However, partial glycerides, i.e., monoacylglycerols

(monoglycerides) and diacylglycerols (diglycerides), are important intermediates in metabolism.

-*Glycerol ethers*: which can be divided into alkyl-ether and alk-1-enyl ethers (or plasmalogens). This class of lipids includes Betaine lipids that have been reported from plants or algae (Dembitsky, 1996). In addition, Ether lipids are found also in three classes of Archaebacteria (the extreme halophiles, the thermoacidophiles, and the methanogens) (Kates, 1990, 1993).

-*Phospholipids*: which are generally divided into two major classes: sphingophospholipids and glycerophospholipids depending on whether they contain a sphingosyl or a glycerol backbone, respectively.

-*Glycosylglycerides*: which are important in the photosynthetic membranes of higher plants, algae, and *Cyanobacteria*. They contain sugars linked glycosidically to diacylglycerol.

Hence, selected lipid species determination is very interesting for better analysis, processes metabolic characterization and for further lipid investigation in many relevance applications such as in medicinal approaches like diagnostics of lipoproteins or in nutrition and food application.

3. Characterization of oils and fats in some organisms

3.1. Cereal lipids

Cereals constitute the main source of food. Cereal lipids are of increasing interest because of their multiple roles, e.g., in food technology and as potential sources of dietary PUFAs in significant amounts (Price and Parsons, 1974). Yet, oilseeds represent many versatile commodities due to their use essentially in food, feed, and medicine. Nevertheless, the content of the lipids of total cereals varies considerably. Furthermore, different parts of cereal seed have different lipid contents. In wheat, for example, the germ contains about 25-30 % lipid while its endosperm has only 1 % lipid. However, considerable amounts of germ lipid of wheat are lost during milling (Galliard and Barnes, 1980). Oat (*Avena sativa* L.) belongs to the *Poaceae* family (Kellogg, 1998), is a cool-season annual cereal crop that thrives in cool, moist climates, but adapts to various soil types (Islam et al., 2011). This cereal is one of the crops cultivated by mankind for the longest time (Lásztity, 1998). Oats are cultivated for grain, fodder, straw,

and feed (Särkijärvi and Saastamoinen, 2006) over more than 9 million hectares globally (FAO, 2014). Compared with other cereals, the specific feature of *Avena sativa* L. grains is the high oil content, which can range from 3 % to 18 % of the grain weight in different cultivars (Banas et al., 2007; Frey and Holland, 1999; Peterson and Wood, 1997), and deposited mainly in the grain endosperm tissues (Price and Parsons, 1979; Youngs et al., 1977). However, most oat cultivars comprise about 5–6 % of oil and 55–60 % of the starch in the grain (Banas et al., 2007; Welsh, 1995; Doehlert et al., 2001), and high-lipid oat remains a potential oil crop (Heneen et al., 2009). Oats contain a wide range of active compounds, including avenanthramides, starches, hydrocolloid β-D-glucan, vitamins, saponins and other antioxidants (mostly phenolic esters) and a relatively high content of total lipids with a high unsaturated fatty acids percentage (Kurtz and Wallo, 2007; Sur et al., 2008). Total lipids can reach 18 % (Frey and Holland, 1999; Peterson and Wood, 1997), and about 41 % of the groat lipids are triglycerides, while 5 % are free fatty acids (Youngs, 1978). Oat lipids have been reviewed by Ben Halima et al. (2015).

3.2. Leaf lipids

FA and lipid compositions of plant leaves denote a remarkable consistency among species; and are usually slightly altered by developmental period and external factors (Hitchcock and Nichols, 1971). Lipid can account for about 7 % of dry weight leaves, with galactosylglycerides being the major constituents (Harwood, 1980). However, essential oils from different plant leaves have shown potent antioxidant properties, for instance, essential oil from *Pelargonium graveolens L'Her* in reproductive damage induced by deltamethrin in mice (Ben Slima et al., 2013); or in the suppression of crown gall disease thanks to the healthy composition of essential oils from *Ruta montana* L. leaf (Hammami et al., 2015). Overall, distribution of lipids in leaves reflects a high chloroplast content Their thylakoid membranes have rather simple acyl lipids (Hardwood, 1980), which resemble those found in cyanobacterial membranes. When plants are grown in the dark, they have pale yellow leaves consisting of etioplasts rather than chloroplasts. Thus, glycosylglycerides are less important in etioplasts than in normal green leaves and there are changes in the composition of FAs (Harwood, 1980).

3.3. Fungal lipids

The lipid content of many fungi has proved to be variable, depending on the individual species as well as the growth conditions. Besides, lipids of yeast cells, mycelium, reproductive structures and that of different stages of development, all vary. The major components of fungi lipids are, in general, phospholipids and sterols (and of course their esters, TAGs, and FAs) (Weete, 1974, 1980; Rattray, 1988; Losel, 1980). The use of fungi and yeasts as a source of oils is covered by Cohen and Ratledge (2005).

3.4. Algal lipids

There is an increasing interest in the exploitation of microalgae and algae from different origins (marine and freshwater) as sources of important natural products as well as for many potential industrial applications. Because of the extremely diverse groups of algae, it is unsurprising that their lipids are diverse, and enormous attention is paid to FAs content. The particular importance of many of these organisms is that of richness in PUFAs (Cohen and Ratledge, 2005). Cyanobacteria, often called blue-green algae, have a much simpler lipid and FA composition than eukaryotic algae. Certain diatoms, unicellular brown algae, named the *Bacillariophyceae*, have been found to have other novel sulfur-containing lipids (Kates, 1987). Guschina and Harwood (2006) reported unusual lipids, including oxylipins in their review on lipid metabolism in eukaryotic algae. Oxylipins may have an important physiological function. Some other algae (microalgae or macroalgae) such as *Dunaliella salina* and *Botrycoccus braunii* can produce hydrocarbons at levels of 15-75 %. *Porphyridium cruentum*, a red alga, has been proposed as a possible source of arachidonic acid which is an essential FA and is advantageous in aquaculture (Volkman, 1989). Nowadays, there are considerable interests for exploiting macroalgae and microalgae in the production of PUFAs especially those of very long-chains like DHA and EPA (Guschina and Harwood, 2006).

For biosynthesis of quality lipids in algae, acetyl-Co-A is required as the starting point. As such, acetyl-CoA carboxylase and synthetase and the other enzymes involved in the different types of lipid biosynthesis pathway have been targeted for enhancing lipid production (Ratledge, 2004; Richmond, 2004; Avidan and Pick, 2015). Unusual hydrocarbons from algae have also been reported such as trienes, triterpenoid botryococcenes, n-alkadienes, tetraterpenoids lycopadiene and methylated squalenes (Achitouv et al., 2004; Metzger and Largeau, 2005; Riekhof et

al., 2005). A proposed pathway of lipid biosynthesis (FAs and TAG biosynthesis) was highlighted by Scott et al. (2010) which involved diacylglycerol-acyl-transferase (DGAT), a key enzyme for lipid biosynthesis (**Figure 3**).

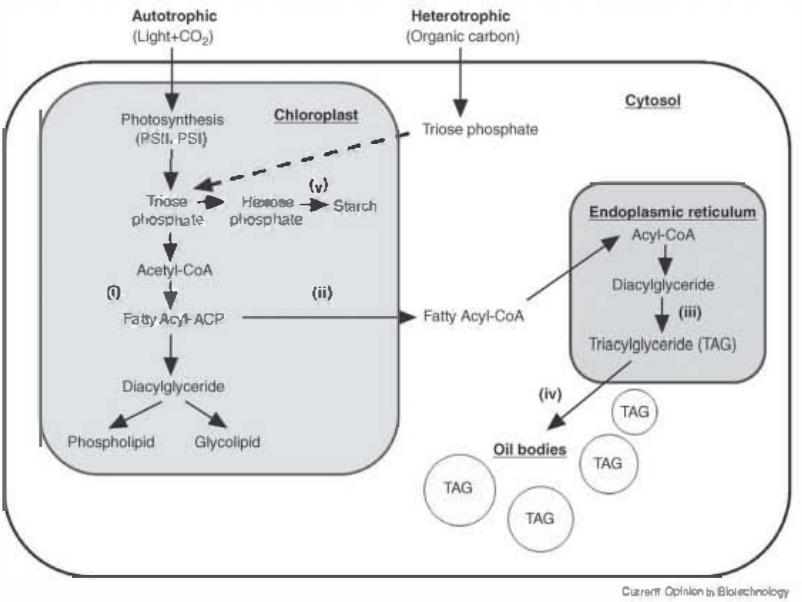


Figure 3. Basic overview of the pathway of carbon capture and lipid biosynthesis. Only the major steps are indicated for clarity. Precursor fatty acids are synthesized de novo in the chloroplast, using either carbon fixed during photosynthesis or from an exogenous supply of organic carbon; the exact nature of what enters the chloroplast is unknown in algae (dashed line). Free fatty acids are exported from the chloroplast and then converted to TAGs in the endoplasmic reticulum (ER), where they bud off into oil bodies in the cytosol. Key: (i) = acetyl-CoA carboxylase (ACCase) and fatty acid synthase (FAS); (ii) = Fatty acid thioesterases and acyl-CoA synthetases; (iii) = TAG biosynthesis enzymes, including acyl-CoA:diacylglycerol acyltransferase (DGAT); (iv) = oil body formation; and (v) = ADP-glucose pyrophosphorylase and starch synthase (Scott et al., 2010).

Genetic engineering of key enzymes involved in specific pathways of FA production within lipid biosynthesis is, remarkably, a promising goal to improve both quality and quantity of lipids, like saturation grade and chain length (Singh and Dhar, 2011).

The algal ability for adaptation to different environmental conditions is particularly reflected in the aptitude of an exceptional diversity of lipid patterns with the capacity to synthesize unusual compounds (Thompson, 1996). Furthermore, modern analytical techniques (high-performance chromatographic techniques and mass spectrometry) and the different molecular biology tools would lead, certainly, to positive development in the identification of new classes of lipids (usual and unusual classes of lipids) and their derivatives in algae or in other living organisms. For more information about lipid metabolism in algae refer to Guschina and Harwood (2006).

The selection of lipid-rich strains of algae can be done using high-throughput staining, for instance, the fluorescent dyes; Nile red and BODIPY 505/515 which are commonly used to quantify algal lipids. Indeed, lipids fluorescent staining is a rapid analysis tool and also inexpensive to measure lipid content (Rumin et al., 2015).

It is clear that microscopic morphology without fluorescent dyes is less intensive to visualize lipid bodies (Shrivastav et al., 2015; Balduyck et al., 2015), as shown in **Figure 4** with the green fluorescent dye, BODIPY 505/515, which was served to stain oil-containing lipid bodies of *Monoraphidium sp.* or with the Nile red staining to visualize oil in *T-Isochrysis* (**Figure 5**).

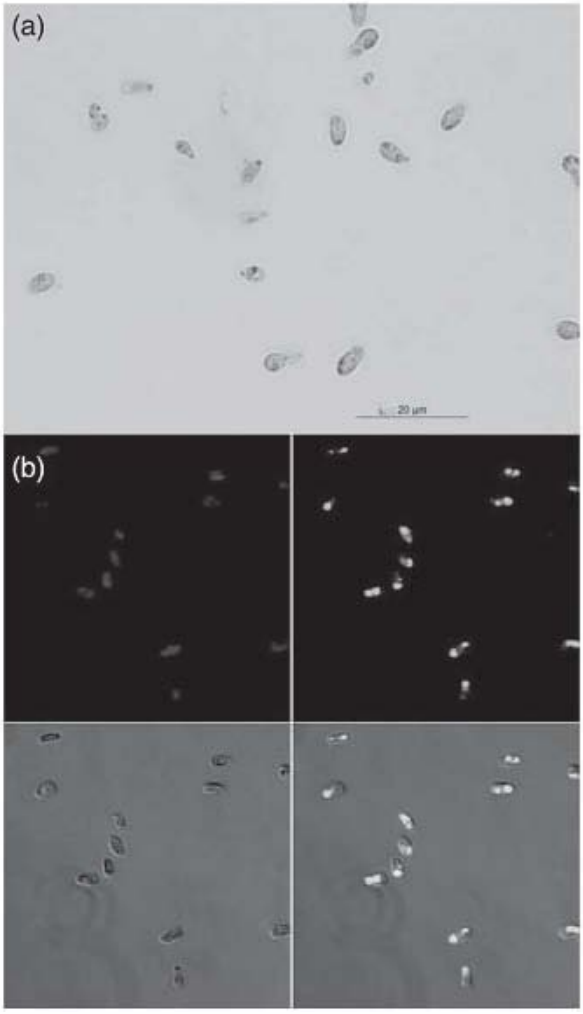


Figure 4. *Microscopic morphological observation of Monoraphidium sp. (a) Light micrograph; (b) fluorescent confocal micrograph of algae stained with* ***BODIPY*** *505/515* ***Oil-containing lipid bodies can be vitally stained and visualized in live oleaginous (oil-containing) algal cells using the green fluorescent dye,*** ***BODIPY*** *505/515 (Shrivastav et al., 2015).*

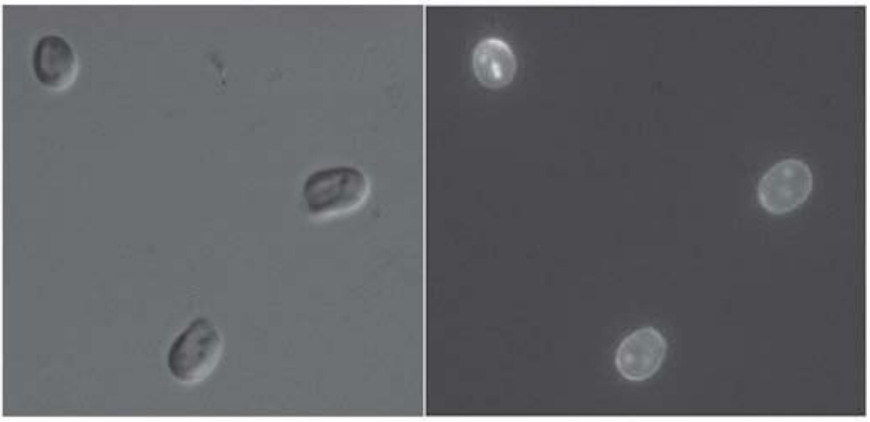


Figure 5. Visualization of NR staining in T-Isochrysis with light microscopy (left) and fluorescence microscopy (right) (Balduyck et al., 2015).

4. Taxonomic characteristics of some oleaginous algae

The systematic classification and the definition of various algal groups are based on the types and combinations of photosynthetic pigments, as well as the algal cell walls and the chemical nature of storage products present in diverse algal species. Thus, thousands of eukaryotic algae are grouped into nine divisions and some of their classes include *Rhodophyceae* (red algae), *Phaeophyceae* (brown algae), *Chlorophyceae* (green algae), *Chryso-phyceae* (yellow-green algae) and *Pyrrophyceae* (dinoflagellates) (Van den Hoek et al., 1995).

Sakthivel et al. (2011) described some microalgae for biodiesel production; such kinds of microalgae are qualified as oleaginous organisms due to their high lipid content (Ren et al., 2015; Shrivastav et al., 2015; Liu et al., 2015). The most important ones are detailed in Table 2.

Table 2. Taxonomic information of some oleaginous microalgae (Sakthivel et al., 2011)

	Dunaliella	*Nannochloropsis*	*Botryococcus braunii*	*Dunaliella salina*	*Isochrysis*	*Phaeodactylum*
Domain	Eukaryota	Eukaryota	-	-	-	-
Kingdom	Viridiplantae	Chromalyeolata	Plantae	Plantae	Chromalyeolata	Chromalyeolata
Phylum (Division)	Chlorophyta	Heterokontophyta	Chlorophyta	Chlorophyta	Haptophyta	Heterokontophyta
Class	Chlorophyceae	Eustigmatophyceae	Chlorophyceae	Chlorophyceae	Prymnesioophyceae	Bacillariophyceae
Order	Chlamydomonadales	-	Chlorococcales	Volvocales	Isochrysidales	Naviculales
Family	Dunaliellaceae	-	Dictyosphaeriaceae	Dunaliellaceae	Isochrysidaceae	Phaeodactylaceae
Genus	Dunaliella	Nannochloropsis	Botryococcus	Dunaliella	Isochrysis	Phaeodactylum
Species	*D. acidophila;* *D. lateralis;* *D. bardawil;* *D. bioculata;* *D. maritima;* *D. tertiolecta*	*N. gaditana;* *N. granulate;* *N. limnetica;* *N. oceanic;* *N. oculata;* *N. salina*	*B. braunii*	*D. salina*	*I. galbana*	*P. tricornutum*

Dunaliella tertiolecta, a marine green flagellate microalga, is reported to be a fast-growing strain and have oil yield around 37 % (in organic basis). *Nannochloropsis* differs from the other related microalgae in that it lacks chlorophyll b and c. This alga is able to accumulate high levels of PUFAs. As such, it is a promising alga for industrial applications, mainly as an energy-rich food for rotifers and fish larvae. *Botryococcus braunii* (Bb) is a planktonic green microalga of potential importance in biotechnology. Its colonies are held together by a lipid biofilm matrix. This species is able to produce high amounts of hydrocarbons, particularly oils in the form of triterpenes (30-40 % of dry weight). *Dunaliella salina* is a halophilic micro-alga with pink color, exclusively found in salty fields. This microalga is known to produce high concentrations of antioxidants especially carotenoids (β-carotene) and is also able to produce large amounts of glycerol for protection against osmotic pressure. This would offer many opportunities for exploiting such substances for eventual industrial applications (cosmetics, nutrition, medicine, etc.). *Isochrysis galbana* is widely used in the bivalve aquaculture industry as an outstanding food for bivalve larvae. Unlike *Isochrysis*, *Phaeodactylum tricornutum* is a diatom which can exist in different morphotypes (oval, triradiate and fusiform) and can grow without silicon.

5. Harmful algae

Many health problems are associated with significant increase and accumulation of toxic phototrophic organisms such as cyanobacteria and algae, notably with the harmful algal blooms (HABs) which are commonly called red tides (Anderson et al., 2012; Sun et al., 2016). These problems are becoming more frequent in aquatic systems. Therefore, accurate methods for rapid detection of harmful organisms as well as their toxins are required and very important to monitor the occurrence of such particular HABs. Orellana et al. (2015) developed a useful method for the quantification and profiling of lipophilic marine toxins in microalgae-based on ultra-high-performance liquid chromatography (UHPLC) coupled to high-resolution Orbitrap mass spectrometry (HR-Orbitrap MS). These authors studied three key algal species: *Protoceratium reticulatum*, *Prorocentrum lima*, and *Alexandrium ostenfeldii* which were used to test their efficiency to extract selected toxins, i.e., yessotoxin (YTX), okadaic acid (OA) and desmethyl spirolide (SPX), respectively. In addition, through a homemade database, screened algae culture was deemed for analogs and metabolites; and by using the ToxID software program, many toxin derivatives would be detected in the harmful microalgae species.

On the other hand, the identification of microalgae in HABs was achieved traditionally by light and electron microscopy examination. However, these methods may present some difficulties and are in general time-consuming. With the development of the real-time polymerase chain reaction (PCR) assays within the last decade, rapid detection of microorganisms in HABs has been more efficient (Galluzzi et al., 2004). Thus, molecular techniques such as real-time PCR are more and more useful for detecting desirable microorganisms. Nevertheless, exact identification of separate strains is practically impossible when using PCR for reasons of the limited availability of the specific primer-probe sets. Therefore, more specific techniques are needed for the exact identification and characterization of microorganisms. Recently, the well-known techniques are the mass spectrometry-based methods, which become more sensitive and specific techniques for such identification (Lee et al., 2015; Orellana et al., 2015). Lee et al. (2015) have mentioned that Matrix-assisted laser desorption/ionization with time-of-flight-Mass spectrometry (MALDI-TOF-MS) is suitable for the rapid identification of fungi and bacteria in clinical microbiology thanks to the accuracy of the technique as well as to a distinctive molecular fingerprint of highly abundant microbial proteins which would be measured successfully using MALDI-TOF-MS. Besides, their work was achieved to analyze the phylogenetic of some harmful microalgal strains of the genus *Chaetoceros*, *Heterocapsa*, and *Alexandrium*, as well as microalgal strains belonging to the genus *Chlorella*, *Nannochloropsis,* and *Dunaliella*; using MALDI-TOF-MS. In conclusion, their research for the rapid identification of microalgae highlighted that MALDI-TOF-MS was efficient and comparable to conventional methods such as classical morphology and molecular genetic methods based on ribosomal RNA genes. Thus, these authors suggested the application of mass spectrometry-based methods in the rapid identification of dangerous aquatic microorganisms (Lee et al., 2015).

6. Biotechnological applications of algae

The high-quality composition of algae including, for instance, lipids, starch, and bioactive and antioxidant components is globally recognized for a wide range of commercial applications, namely in aquaculture and energy industries. The prominent steps to satisfy the final target application is to select the appropriate strains of algae/microalgae as well as the optimal conditions to overproduce the targeted substances.

On the one hand, optimization for a high quantity of energy reserves in algae, especially lipids and starch is very promising to the biofuel industries as lipids can be transformed to biodiesel and starch to bioethanol (Vitova et al., 2015). The need for renewable fuels (biofuels) that are non-pollutants and friendly to the environment, as well as more disposable and which agree with the concept of sustainable development (as compared to fossil fuels), has, nowadays, become more and more exigent. Furthermore, microalgae as feedstocks for biofuels production are much more competitive than those from traditional plant feedstocks (rapeseed, soybean, sunflower or palm oil) which would cater for food starvation in poor countries instead of being feedstocks for biofuels. Many other types of research dealt with the efficiency of microalgae as bioenergy sources (biodiesel, bioethanol, biomethane, biohydrogen, bioelectricity) when satisfying specific parameters. These parameters could be culture techniques (photobioreactors, fermenters, etc.), environmental conditions (light, temperature, gas exchange, nutrients, etc.), and harvesting (flocculation, centrifugation, filtration) of the microalgal biomass as well as efficient strain selection of microalgae in response to the different optimal conditions for potential applications (Singh and Dhar, 2011; Scott et al., 2010; Moody et al., 2014; Tan et al., 2015).

On the other hand, microalgae are very useful to animal feed, especially in the aquaculture industry. Microalgae are consumed either as individual diets or mixed components diets to supply basic nutrients for marine organisms particularly fish and their preys such as rotifers. Among the valuable nutrients provided by different strains of microalgae are the PUFAs which can be in insufficient amounts at various stages of fish growth and that lead to lesser meat fish quality. Thus, incorporating adequate microalgae for this purpose would resolve the quality of fish for healthier and commercial uses. PUFAs like EPA and DHA are essential for the survival of fish larvae (Skjånes et al., 2013). In addition, several other authors tended to determine important species of microalgae in aquaculture (Santos-Ballardo et al., 2015). These authors confirmed the high importance of four different species of microalgae in aquaculture i.e., *Chaetoceros calcitrans*, *Isochrysis affinis galbana* (T-Iso), *Nannochloropsis gaditana* and *Phaeodactylum tricornutum* via developing a specific spectrophotometric methodology for cell counting (biomass measurements) of these species of microalgae used in aquaculture.

For more information on algal lipids and their uses in aquaculture or in other biotechnological applications, refer to Ben Halima (2017, 2018).

7. Concluding remarks

The starting point for this review is the importance of lipids in biotechnology as well as the multidisciplinary process found with the combination of algal technology (Skjånes et al., 2007). The research on microalgal biotechnologies has increased steadily over the last decade (Plaza et al., 2009; Lu and Xu, 2015). Microalgae are able to accumulate over half cell dry weight of lipids, mostly in the form of TAGs as storage lipids. Though these lipids are produced in small amounts under optimized growth conditions, the large amounts of such algal TAGs are accumulated only under growth stress conditions such as in the stress case of pH (Razzak et al., 2015), temperature (Wei et al., 2015), light stress (Gris et al., 2014; He et al., 2015) or nutrient limitation (phosphate and nitrate). This latter is a well-known phenomenon to lipid accumulation, while under nitrogen deficiency, lipid accumulation is the most widely studied (Yang et al., 2011). Indeed, stress conditions on the growth of microalgae are essential to induce high quantities of desirable products (**Figure 6**), e.g., lipids, hydrogen, glycerol, antioxidants and other secondary metabolites (Skjånes et al., 2013; Shrivastav et al., 2015).

During the last decades, microalgae have raised great interests as renewable feedstocks thanks to their numerous advantages including, for instance, higher biomass production, higher photosynthetic efficiency, important sequestration level of atmospheric carbon dioxide (Hu et al., 2008) and important rates of growth as compared to the other energy crops (Huang et al., 2010). Moreover, the oil yield from algae can be higher than that of plant oil commonly used by mankind, mainly because algae do not require harvest cycles or annual/biannual growth thanks to their unicellular nature (Wijffels and Barbosa, 2010). Thus, microalgae are unique compared to other food crops and their cultivation would be done on non-arable land leading to reduced competition with plant/vegetable crops. Furthemore, the selection of microalgal species with desirable attributes, such as suitable fatty acid profiles, high lipid content and high nutraceutical constituents is a promising challenge associated with the relevant applications of algal biomass. Moreover, many health problems were found recently because of the increase in industrial activities, which release various types of contaminations worldwide. The concept of microalgae's valorization is nowadays promising and it is at the forefront of sustainable development in the field of industrial applications.

Finally, future investigations on different kinds of microalgae are essential for deeper understanding mechanisms for their high quality of output

products. Besides, biotechnological engineering like genomic approaches (Hlavova et al., 2015) would provide additional information on how to boost microalgae breeding.

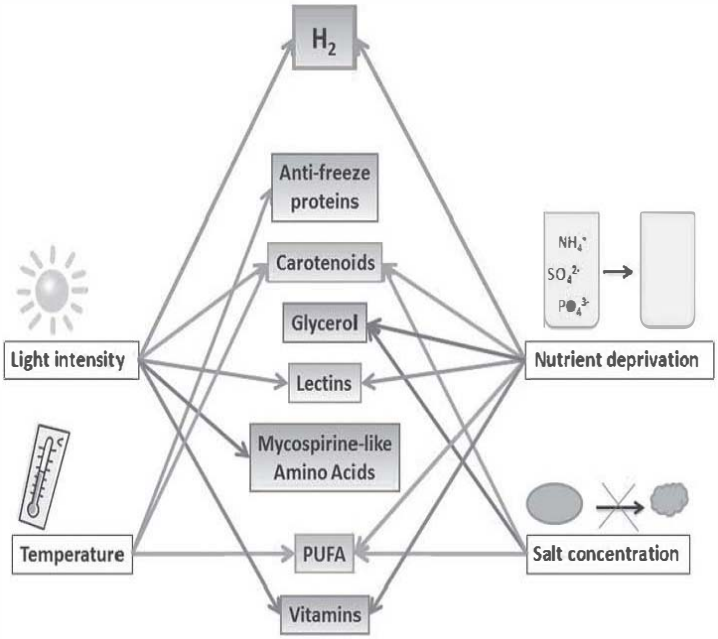


Figure 6. Schematic summary of the most common representative stress reactions, which influence the synthesis of some important products in green algae (Skjånes et al., 2013).

8. Acknowledgments

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

The author thanks the Tunisian Ministry of Higher Education and Scientific Research for facilities.

The author is very grateful to Erasmus Nchuaji Tang, a researcher at the Department of Biochemistry, University of Yaounde I-Cameroon, for his help reading the manuscript.

9. References

Achitouv, E., Metzger, P., Rager, M.N. (2004). C31–C34 methylated squalene from a Bolivian strain of *Botryococcus braunii*. Phytochem 65, 3159–3165.

Anderson, D.M., Cembella, A.D., Hallegraeff, G.M. (2012). Progress in Understanding Harmful Algal Blooms: Paradigm Shifts and New Technologies for Research, Monitoring, and Management. Annu. Rev. Mar. Sci. 4, 143–176.

Avidan, O., Pick, U. (2015). Acetyl-CoA synthetase is activated as part of the PDH-bypass in the oleaginous green alga *Chlorella desiccata*. J. Exp. Bot. 66, 7287-98.

Balduyck, L., Veryser, C., Goiris, K., Bruneel, C., Muylaert, K., Foubert, I. (2015). Optimization of a Nile Red method for rapid lipid determination in autotrophic, marine microalgae is species dependent. J. Microbiol Methods 118, 152–158.

Banas, A., Dexbski, H., Banas, W., Heneen, W.K., Dahlqvist, A., Bafor, M., Gummeson, P.O., Marttila, S., Ekman, A., Carlsson, A.S., Stynme, S. (2007). Lipids in grain tissues of oat (*Avena sativa*): differences in content, time of deposition, and fatty acid composition. J. Exp. Botany 58, 2463–2470.

Ben Halima, N., Ben Saad, R., Khemakhem, B., Fendri, I., Abdelkafi, S. (2015). Oat (*Avena sativa* L.): Oil and Nutriment Compounds Valorization for Potential Use in Industrial Applications. J. Oleo Sci. 64, 915-932.

Ben Halima, N. (2017). Occurrence and Importance of Plant Lipids: A Promising Insight into Algae. Recent Pat. Food Nutr. Agric. 8, 166-174.

Ben Halima, N. (2018). Why is it important to use algae in aquaculture? Journal of Biochemistry and Biotechnology Volume 1, Issue 1.

Ben Slima, A., Ali, M.B., Barkallah, M., Traore, A.I., Boudawara, T., Allouche, N., Gdoura, R. (2013). Antioxidant properties of *Pelargonium graveolens L'Her* essential oil on the reproductive damage induced by deltamethrin in mice as compared to alpha-tocopherol. Lipids Health Dis. 12:30.

Carrapiso, A.I., Garcia, C. (2000). Development in lipid analysis: some new extraction techniques and in situ transesterification. Lipids 35, 1167-1177.

Cohen, Z., Ratledge, C., Eds. (2005). Single Cell Oils, AOCS, Champaign, IL.

Damste, J.S.S., Strous, M., Rijpstra, W.I.C., Hopmans, E.C., Geenevasen, J.A.J., van Duin, A.C.T., van Niftrik, L.A., Jetten, M.S.M. (2002). Linearly concatenated cyclobutane lipids form a dense bacterial membrane. Nature 419, 708-712.

DaSilva, E.J. (2004). The colours of biotechnology: science, development and humankind. Electron. J. Biotechnol. 7 (3).

Dembitsky, V.M. (1996). Betaine ether-linked glycerolipids: chemistry and biology. Prog. Lipid Res. 35, 1-51.

Doehlert, D.C., McMullen, M.S., Hammond, J.J. (2001). Genotypic and environmental effects on grain yield and quality of oat grown in North Dakota. Crop Sci. 41, 1066–1072.

Domergue, F., Abbadi, A., Zahringer, U., Moreau, H., Heinz, E. (2005). *In vivo*-characterization of the first acyl-CoA D6-desaturase from the plant kingdom. Biochem. J. 389, 483–90.

Drexler, H., Spiekermann, P., Meyer, A., Domergue, F., Zank, T., Sperling, P., Abbadi, A., Heinz, E. (2003). Metabolic engineering of fatty acids for breeding of new oilseed crops: strategies, problems and first results. J. Plant Physiol. 160, 779–802.

FAO. (2014). Food and Agriculture Organization of the United Nations: FAOSTAT. http://faostat.fao.org/ (accessed April 2014).

Frey, K.J., Holland, J.B. (1999). Nine cycles of recurrent selection for increased grain-oil content in oat. Crop Sci. 39, 1636–1641.

Galliard, T., Barnes, P.J. (1980). The biochemistry of lipids in cereal crops. In Biogenesis and Function of Plant Lipids, Eds, Mazliak P, Benveniste P, Costes C, Douce R. Elsevier, Amsterdam, pp. 191-198.

Galluzzi, L., Penna, A., Bertozzini, E., Vila, M., Garces, E., Magnani, M. (2004). Development of a real-time PCR assay for rapid detection and quantification of *Alexandrium minutum* (a Dinoflagellate). Appl. Environ. Microbiol. 70, 1199–1206.

Gris, B., Morosinotto, T., Giacometti, G.M., Bertucco, A., Sforza, E. (2014). Cultivation of Scenedesmus obliquus in photobioreactors: effects of light intensities and light–dark cycles on growth, productivity, and biochemical composition. Appl. Biochem. Biotechnol. 172, 2377–2389.

Guarro, J., Gené, J., Stchigel, A.M. (1999). Developments in fungal taxonomy. Clin. Microbiol. Rev. 12, 454-500.

Guschina, I.A., Harwood, J.L. (2006). Lipids and lipid metabolism in eukaryotic algae. Prog. Lipid Res. 45, 160–186.

Hammami, I., Smaoui, S., Hsouna, A.B., Hamdi, N., Triki, M.A. (2015). *Ruta montana* L. leaf essential oil and extracts: characterization of

bioactive compounds and suppression of crown gall disease. EXCLI J. 14, 83-94.

Harwood, J.L. (1980). Plant acyl lipids: structure, distribution and analysis. In The Biochemistry of Plants, Eds, Stumpf PK, Conn EE. Academic Press, New York, pp. 1-55.

Hlavova, M., Turoczy, Z., Bisova, K. (2015). Improving microalgae for biotechnology — From genetics to synthetic biology. Biotechnol. Advances 33, 1194–1203.

He, Q., Yang, H., Xu, L., Xia, L., Hu, C. (2015). Sufficient utilization of natural fluctuating light intensity is an effective approach of promoting lipid productivity in oleaginous microalgal cultivation out-doors. Bioresour. Technol. 180, 79–87.

Heneen, W.K., Banas, A., Leonova, S., Carlsson, A.S., Marttila, S., Debski, H., Stynme, S. (2009). The distribution of oil in the oat grain. Plant Signal Behav. 4, 55-56.

Hitchcock, C., Nichols, B.W. (1971). Plant Lipid Biochemistry. Academic Press, London.

Hu, Q., Sommerfeld, M., Jarvis, E., Ghirardi, M., Posewitz, M., Seibert, M., Darzins, A. (2008). Microalgal triacylglycerols as feedstocks for biofuel production: perspectives and advances. Plant J. 54, 621–639.

Huang, G., Chen, F., Wei, D., Zhang, X., Chen, G. (2010). Biodiesel production by microalgal biotechnology. Appl. Energy 87, 38–46.

Islam, M.R., Xue, X., Mao, S., Ren, C., Eneji, A.E., Hu, Y. (2011). Effects of water-saving superabsorbent polymer on antioxidant enzyme activities and lipid peroxidation in oat (*Avena sativa* L.) under drought stress. J. Sci. Food Agric. 91, 680-686.

Kates, M. (1987). Lipids of diatoms and of halophilic *Dunaliella* species. In The Metabolism, Structure and Function of Plant Lipids. Eds, Stumpf PK, Mudd JB, Nes WD. Plenum, New York, pp. 613-621.

Kates, M. (1990). Glyco-, phosphoglyco- and sulfoglycoglycerolipids of bacteria. In Handbook of Lipid Research, 6, Kates M., Ed., Plenum: New York, pp. 1-122.

Kates, M. (1993). Membrane lipids of Archaea. In: The biochemistry of Archaea (Archaeobacteria), Kates M., Kushner DJ., Matheson AT. Eds., Amsterdam: Elsevier, pp. 261-295.

Kellogg, E.A. (1998). Relationships of cereal crops and other grasses. Proc. Natl. Acad. Sci. USA 95, 2005-2010.

Kurtz, E.S., Wallo, W. (2007). Colloidal oatmeal: history, chemistry and clinical properties. J. Drugs Dermatol. 6, 167–170.

Lásztity, R. (1998). Oat grain – a wonderful reservoir of natural nutrients and biologically active substances. Food Rev. Int. 14, 99-119.

Lee, H.W., Roh, S.W., Cho, K., Kim, K.N., Cha, I.T., Yim, K.J., Song, H.S., Nam, Y.D., ●da, T., Chung, Y.H., Kim, S.J., Choi, J.S., Kim, D. (2●15). Phylogenetic analysis of microalgae based on highly abundant proteins using mass spectrometry. Talanta 132, 63●-634.

Liu, B., Liu, J., Chen, T., Yang, B., Jiang, Y., Wei, D., Chen, F. (2●15). Rapid characterization of fatty acids in oleaginous microalgae by near-infrared spectroscopy. Int. J. Mol. Sci. 16, 7●45-7●56.

Losel, D.M. (1●8●). Fungal Lipids. In Microbial Lipids. Eds, Ratledge C, Wilkinson SG. Academic Press. London, pp. 6●9-8●6.

Lu, Y., Xu, J. (2●15). Phytohormones in microalgae: a new opportunity for microalgal biotechnology? Trends Plant Sci. 2●, 273-2●2.

Metzger, P., Largeau, C. (2●●5). *Botryococcus braunii*: a rich source for hydrocarbons and related ether lipids. Appl. Microbiol. Biotechnol. 66, 4●6-4●6.

Meyer, A., Kirsch, H., Domergue, F., Abbadi, A., Sperling, P., Bauer, J., Cirpus, P., Zank, T.K., Moreau, H., Roscoe, T.J., Zähringer, U., Heinz, E. (2●●4). Novel fatty acid elongases and their use for the reconstitution of docosahexaenoic acid biosynthesis. J. Lipid Res. 45, 1●9●–1●●9.

Moody, J.W., McGinty, C.M., ●uinn, J.C. (2●14). Global evaluation of biofuel potential from microalgae. PNAS 111, 8●9●-8●9●.

●ECD. (2●●5). The ●rganisation for Economic Co-operation and Development. A framework for biotechnology statistics. Paris.

●rellana, G., Van Meulebroek, L., Van Vooren, S., De Rijcke, M., Vandegehuchte, M., Janssen, C.R., Vanhaecke, L. (2●15). ●uantification and profiling of lipophilic marine toxins in microalgae by UH■PLC coupled to high-resolution orbitrap mass spectrometry. Anal. Bioanal. Chem. 4●7, 6345-6356.

●shima, M., Agriga, T. (1●75). ω-Cyclohexyl fatty acids in acidophilic thermophilic bacteria. J. Biol. Chem. 25●, 6●63-6●6●.

Peterson, D.M., Wood, D.F. (1●●7). Composition and structure of high-oil oat. J. Cereal Sci. 26, 121–12●.

Plaza, M., Herrero, M., Cifuentes, A., Ibáñez, E. (2●●9). Innovative natural functional ingredients from microalgae. J. Agric. Food Chem. 57, 7159–717●.

Price, P.B., Parsons, J.G. (1●74). Lipids of six cultivated barley (*Hordeum vulgare*) varieties. Lipids 9, 56●-566.

Price, P.B., Parsons, J. (1●79). Distribution of lipids in embryonic axis, bran-endosperm, and hull fractions of hulless barley and hulless oat grain. J. Agric. Food Chem. 27, 813–815.

Rattray, J.B.M. (1988). Yeasts. In Microbial Lipids. Eds, Ratledge C, Wilkinson SG. Academic Press. London, pp. 555-697.

Ratledge, C. (2004). Fatty acid biosynthesis in microorganisms being used for single cell oil production. Biochimie 86, 807–815.

Razzak, S.A., Ilyas, M., Ali, S.A., Hossain, M.M. (2015). Effects of CO_2 concentration and pH on mixotrophic growth of *Nannochloropsis oculata*. Appl. Biochem. Biotechnol. 176, 1290-1302.

Ren, H.Y., Liu, B.F., Kong, F., Zhao, L., Ren, N. (2015). Hydrogen and lipid production from starch wastewater by co-culture of anaerobic sludge and oleaginous microalgae with simultaneous COD, nitrogen and phosphorus removal. Water Res. 85, 404-412.

Richmond, A. (2004). Principles for attaining maximal microalgal productivity in photobioreactors: an overview. Hydrobiologia 512, 33–37.

Riekhof, W.R., Sears, B.B., Benning, C. (2005). Annotation of genes involved in glycerolipid biosynthesis in *Chlamydomonas reinhardtii*: discovery of the betaine lipid synthase BTA1Cr. Eukaryot Cell 4, 242–252.

Rumin, J., Bonnefond, H., Saint-Jean, B., Rouxel, C., Sciandra, A., Bernard, O., Cadoret, J.P., Bougaran, G. (2015). The use of fluorescent Nile red and BODIPY for lipid measurement in microalgae. Biotechnol. Biofuels 8:42.

Sakthivel, R., Elumalai, S., Mohammad Arif, M. (2011). Microalgae lipid research, past, present: A critical review for biodiesel production, in the future. J. Exp. Sci. 2, 29-49.

Santos-Ballardo, D.U., Rossi, S., Hemández, V., Gómez, R.V., Rendón-Unceta, M.D.C., Caro-Corrales, J., Valdez-Ortiz, A. (2015). A simple spectrophotometric method for biomass measurement of important microalgae species in aquaculture. Aquaculture 448, 87–92.

Särkijärvi, S., Saastamoinen, M. (2006). Feeding value of various processed oat grains in equine diets. Livestock Sci. 100, 3–9.

Scott, S.A., Davey, M.P., Dennis, J.S., Horst, I., Howe, C.J., Lea-Smith, D.J., Smith, A.G. (2010). Biodiesel from algae: challenges and prospects. Curr. Opin. Biotechnol. 21, 277-286.

Shrivastav, A., Mishra, S.K., Suh, W.I., Farooq, W., Moon, M., Kim, T.H., Kumar, K., Choi, G.G., Park, M.S., Yang, J.W. (2015). Characterization of newly isolated oleaginous microalga *Monoraphidium* sp. for lipid production under different conditions. Algal Res. 12, 289–294.

Sijtsma, L., de Swaaf, M.E. (2004). Biotechnological production and application of the ω-3 polyunsaturated fatty acid docosahexaenoic acid. App. Microbiol. Biotechnol. 64, 146–153.

Singh, N.K., Dhar, D.W. (2011). Microalgae as second-generation biofuel. A review. Agron. Sustain. Dev. 31, 605–629.

Skjånes, K., Lindblad, P., Muller, J. (2007). BioCO_2 – a multidisciplinary, biological approach using solar energy to capture CO_2 while producing H_2 and high-value products. Biomol. Eng. 24, 405–413.

Skjånes, K., Rebours, C., Lindblad, P. (2013). Potential for green microalgae to produce hydrogen, pharmaceuticals and other high-value products in a combined process. Crit. Rev. Biotechnol. 33, 172–215.

Sogin, M.L. (1989). Evolution of eukaryotic microorganisms and their small subunit ribosomal RNAs. Am. Zool. 29, 487-499.

Sun, Y.Y., Wang, H., Guo, G.L., Pu, Y.F., Yan, B.L., Wang, C.H. (2016). Isolation, purification, and identification of antialgal substances in green alga *Ulva prolifera* for antialgal activity against the common harmful red tide microalgae. Environ. Sci. Pollut. Res. Int. 23, 1449-59.

Sur, R., Nigam, A., Grote, D., Liebel, F., Southall, M.D. (2008). Avenanthramides, polyphenols from oats, exhibit anti-inflammatory and anti-itch activity. Arch. Dermatol. Res. 300, 569–574.

Tan, C.H., Show, P.L., Chang, J.S., Ling, T.C., Lan, J.C.W. (2015). Novel approaches of producing bioenergies from microalgae: A recent review. Biotechnol. Advances 33, 1219–1227.

Taylor, J.W., Swann, F.C., Berbee, M.L. (1994). Molecular evolution of ascomycete fungi: phylogeny and conflict, p. 201-212. In Hawksworth DL (ed.). Ascomycete systematics: problems and perspectives in the nineties. Plenum Press, New York, N.Y.

Thompson, G.A. (1996) Lipids and membrane function in green algae. Biochim. Biophys. Acta 1302, 17–45.

Van den Hoek, C., Mann, D.G., Jahns, H.M. (1995). Algae. An introduction to phycology. Cambridge: University Press.

Vitova, M., Bisova, K., Kawano, S., Zachleder, V. (2015). Accumulation of energy reserves in algae: From cell cycles to biotechnological applications. Biotechnol. Advances 33, 1204–1218.

Volkman, J.K. (1989); Fatty acids of microalgae used as feedstocks in aquaculture. In Fats for the Future, Ed, Cambie RC, Horwood E, Chichester, UK, pp. 263-283.

Weete, J.D. (1974). Fungal Lipid Biochemistry. Plenum, New York.

Weete, J.D. (1980). Lipid Biochemistry of Fungi and Other Organisms. Plenum, New York.

Wei, L., Huang, X., Huang, Z. (2015). Temperature effects on lipid properties of microalgae *Tetraselmis subcordiformis* and *Nannochloropsis oculata* as biofuel resources. Chin. J. Oceanol. Limnol. 33, 99–106.
Welsh, R.W. (1995). The chemical composition of oats. In: Welsh RW (ed.). The oat crop, London: Chapman & Hall, 279–320.
Wijffels, R.H., Barbosa, M.J. (2010). An outlook on microalgal biofuels. Science 329, 796–799.
Yang, J., Xu, M., Zhang, X., Hu, Q., Sommerfeld, M., Chen, Y. (2011). Life-cycle analysis on biodiesel production from microalgae: water footprint and nutrients balance. Bioresour. Technol. 102, 159–165.
Yeung, T., Ozdamar, B., Paroutis, P., Grinstein, S. (2006). Lipid metabolism and dynamics during phagocytosis. Curr. Opin. Cell Biol. 18, 429-437.
Youngs, V.L. (1978). Oat lipids. Cereal Chem. 55, 591-597.
Youngs, V.L., Püskülcü, M., Smith, R.R. (1977). Oat lipids I. Composition and distribution of lipid components in two oat cultivars. Cereal Chem 54, 803–812.

CHAPTER FIVE

PLANT PROTEOMICS AND LIPIDOMICS: GENERALITIES, PERSPECTIVES, AND PROSPECTS

Abstract

Plants can synthesize a wide range of hydrophilic compounds, namely proteins, as well as hydrophobic compounds, which are generally known as lipids. Plant proteomics and lipidomics are comprehensive systems of all proteins and lipids in plants, respectively. The study of these systems includes cell signaling, transcriptional and translational modulation, and cellular interactions in response to environmental conditions. In order to ameliorate protein and lipid research from plants, various analytical methods, as well as databases, were developed. New insights into relevant proteomic and lipidomic techniques would be of great importance to improve knowledge of plant science.

Keywords: Plants, Proteomes, Lipidomes, Analytical techniques, Databases, Online resources, Proteins, Lipids.

1. Introduction

Plants, including higher plants and algae, are good sources of lipids and proteins as well as their relevant enzymes that play critical roles in the biological turnover of their proper materials and are used in many industrial applications, e.g., food, nutraceutical, and medicine, etc.

Plant proteomics and lipidomics could be considered as emerging research fields in plant science, which play important roles in understanding the action of proteins and lipids, respectively, on signaling pathways in plants against various abiotic and biotic stresses.

Bioinformatics for plant proteomics and lipidomics are emerging research fields that give further information about proteins and lipids from plants. Many interesting resources for proteins and lipids can be found online. Table 1 lists some online resources and databases about proteomes and lipidomes.

The following databases namely, LIPIDAT (Caffrey and Hogan, 1992), LIPIDBANK (Yasugi and Watanabe, 2002) and LMSD (Sud et al., 2007) are important databases for the study of lipids. LED (Fischer and Pleiss, 2003), ESTHER (Hotelier et al., 2004; Lenfant et al., 2013) and BIOPEP (Minkiewicz et al., 2008; Iwaniak et al., 2016) are important databases for the study of lipase engineering, α/β-hydrolase fold superfamily of proteins and bioactive peptide sequences, respectively.

Unlike genes and proteins databases, lipids databases are created based on their different scope and organization owing to the lack of a universal classification scheme of lipids (Namasivayam et al., 2015).

Table 1. Some online bioinformatics resources with regards to proteomics and lipidomics

LIPIDAT	http://www.lipidat.ul.ie/
LipidBank	http://lipidbank.jp
LMSD	http://www.lipidmaps.org/data/structure/index.html
LIPID MAPS	http://www.lipidmaps.org/
BIOPEP	http://www.biopeps.com/
LMPD	http://www.lipidmaps.org/data/proteome/index.html
KEGG	http://www.genome.jp/kegg/pathway.html
ESTHER	http://bioweb.ensam.inra.fr/ESTHER/general?what=index
LED	http://www.led.uni-stuttgart.de/

The current chapter will take just a retrospective look at the importance of proteomics and lipidomics in plant science. It would give an outline description of such field, rather than a comprehensive history of this topic. Maybe the latter task is something that one of the readers of this chapter might be inspired to undertake at some time in the future.

In fact, the goal of this chapter is to provide some idea relative to the importance of plant proteomics and lipidomics to plant science.

2. Diverse origins of plant lipids and lipolytic enzymes

The principal sources of plant lipids and lipolytic enzymes are found in laticifers and in oleaginous plants including cereals and oleaginous algae.

For example, olive (*Olea europaea* L.) is one of the oldest known plant characterized by its edible oil, which is derived essentially from the mesocarp of the olive tree fruits. In the report of Panzanaro et al. (2010), an acid lipase activity in olive oil mesocarp was detected, albeit with low catalytic activity. The detected acid lipase in the oil body fraction of mesocarp olive fruit could suggest that such enzyme could be closely associated with oil droplets (Panzanaro et al., 2010). Some other reports (Murphy, 1993; Huang, 1983, 1996) describing plant lipases and relevant themes such as the characterization of oleosins and oil bodies in seeds could be helpful to elucidate the important physiological roles of plant lipids and lipolytic enzymes.

Furthermore, the reports of Ben Halima (2017, 2018 and 2019) highlighted some characteristics relative to lipids and lipolytic enzymes of higher plants and algae.

3. Plant proteomics

Plant proteomics is a research field that focuses on large-scale functional analyses of proteins extracted from plants. There are increasing numbers of publications on plant proteomics, which deal with the importance of this research field for the advancement of plant science. Indeed, plant proteomics has many goals, in particular, the study of the assembly and functional interactions of proteins from plants. The rapidly increasing number of reports on this research field demonstrates its immense potential impacts for current and future plant study. In general, proteins often function as molecular machines that could be organized into multi-protein complexes localized in specialized subcellular compartments. Plant proteomics has gained much interest by studying post-translational modifications of plant proteins. Many developmental methods and techniques would be required to satisfy the large-scale study of plant post-translational modifications of proteins, which are involved in protein turnover, cellular signaling and membrane association.

The book of Šamaj and Thelen (2007) highlighted the rapid progress in plant proteomics and emphasized on model species, subcellular organelles and specific aspects such as plant reproduction, signaling, stress biotic and/or pathogen/symbiotic interactions between microorganisms and plants.

Brief historical overviews concerning plant proteomics, two-dimensional gel electrophoresis and an introduction to bioinformatics were provided in the book of Šamaj and Thelen (2007).

Indeed, the book of Šamaj and Thelen (2007) is a monograph that represented a synthesis of data on plant proteomics including the biological aspects and some important methodological approaches. These latter approaches could be the high-resolution two-dimensional electrophoresis, the protein microchips, the MudPIT (multidimensional protein identification technology), the fluorescent DIGE (difference gel electrophoresis) alone and/or in combination with stable isotope reagents such as ICAT (isotope-coded affinity tag) and iTRAQ (isobaric tag for relative and absolute quantitation), which allow relative protein quantification.

Many chapters on proteomes of selected plants such as *Arabidopsis*, tobacco, oilseed plants, cereals, and legumes, have been provided in the book of Šamaj and Thelen (2007). Several of its chapters have been devoted to plant proteomics of organelles and compartments such as plant plasma membrane, cell wall, mitochondria, plastids, and nucleolus. Additionally, proteomic approaches related to plant reproduction as well as proteomes in regards to plants and microorganisms interactions were also reported in the book of Šamaj and Thelen (2007).

Therefore, the book of Šamaj and Thelen (2007) is an interesting book that covered some progress in the field of plant proteomics and can be recommended to a broad spectrum of readers, in particular, researchers, teachers, and students.

Recently, the review of Bansal et al. (2017) has discussed the applications of cereal proteomics approaches in nutritional sciences. Indeed, these authors reported a brief review of some efforts made by different researchers on the comparative description and functional analyses of cereal proteomics.

Without a doubt, there will be significant challenges to plant proteomics advancements using relevant new techniques and instruments for more quantification, characterization, and analyses of all proteins in plants.

The complete genomic sequence of rice is already available. The proteome of root, leaf and seed tissues from rice using the application of two-dimensional gel electrophoresis and mass spectrometry analyses have already been obtained. The rice proteomics study also covered the growth and developmental stages as well as abiotic and biotic stresses responses (Bansal et al., 2017).

There is continuous progress in genomics studies for wheat as well as subsequent development in wheat proteomics. The review of Bansal et al. (2017) has also evaluated some increased efforts in wheat proteomics that were based on qualitative and quantitative characterizations of proteins, organelles and regulatory networks in wheat.

The availability of the genome sequence of maize has enabled the researchers to better explore the properties of maize. The maize proteomes, and the proteome differences between mutants and wild type genotypes of maize, in particular, have also been identified (Bansal et al., 2017).

The recently completed genome sequence of barley, which is the fourth most important cereal crop worldwide, led to its detailed proteomics and functional studies (Bansal et al., 2017).

Several bioinformatics resources have recently become available, leading to provide more catalogs of genome sequence-predicted proteins in plants and cereals in particular (Bansal et al., 2017).

4. Plant lipidomics

Lipids have many essential roles in cellular functions in plants. Cellular signaling and lipid metabolism are integrated processes, which regulate cell growth, proliferation, and survival.

Recent developments in plant lipid profiling, as well as quantitative analyses, would define new roles of plant lipids in complex biological functions.

The development in lipidomics as biological approach systems could lead to a better understanding of changes in plant lipid composition within

cells, organelles or tissues with respect to stresses and metabolism. Lipidomics would provide an approach to determine relationships between a precursor-product as well as order the temporal-spatial events, which constitute vital processes (Brown, 2007).

The chapter of Namasivayam et al. (2015) is an interesting report of plant lipidomics with emphasis on the role of plant lipids in signaling pathways during stress conditions. Plant lipids could have interesting biological roles (Namasivayam et al., 2015). Various analytical methods were developed to quantify and characterize lipids. Namely, this was done for plant lipids with data available in the same report of Namasivayam et al. (2015). An application of liquid chromatography-ion trap time-of-flight mass spectrometry (LC-IT-TOF-MS) for plant lipidomics was reported in the publication of Okazaki et al. (2013).

In addition, various lipid databases are available online and could be useful information sources about plant lipids and lipidomics (Namasivayam et al., 2015).

5. Current and future developments

Different plant origins, i.e., higher plants and algae, could be in favor to provide many types of proteins and lipids with interesting properties that could be used in many applications, e.g., food, medicine, and biotechnology. The selection of the prominent varieties of plants to extract proteins and lipids could be of great importance to satisfy the demand for application of these components.

Proteins and peptides could have important biological activities such as anti-oxidant and antimicrobial activities. The importance of proteomics and bioinformatics for these bioactive compounds resides in elucidating the main amino acids that bring out such biological activities. Moreover, lipids, in particular, include the lipophilic compounds with interesting properties such as hormones, vitamins, and fatty acids. They could have outstanding biological activities, e.g., anti-oxidant and anti-inflammatory activities. Hence, the search in the lipids databases and the lipidomics, in general, could highlight the mechanism of action of such lipids.

Several resources that assist the study of proteomics and lipidomics are available via the internet, for instance, databases that could provide access to gene mapping data, phenotype information, and amino acid sequences.

A list of selected online resources to boost knowledge on plant proteomics and lipidomics is as following (Ben Halima, 2019).

- **GenBank** - https://www.ncbi.nlm.nih.gov/genbank/. GenBank is the NIH (National Institutes of Health) genetic sequence database.

- **GrainGenes** - https://wheat.pw.usda.gov/GG3/. An online database for *Triticeae* and *Avena*.

- **Integbio Database Catalog** - https://integbio.jp/dbcatalog/?lang=en. An online catalog of life science databases.

- **STRING** - https://string-db.org/. Database of known and predicted protein-protein interactions.

- **TriFLDB** - https://integbio.jp/dbcatalog/en/record/nbdc01047. A database for *Triticeae* Full-Length CDS.

- **Lipidomics-related web sites:** http://www.ksu.edu/lipid/lipidomics;

http://hcc.musc.edu/research/shared_resources/lipidomics.cfm and

http://medschool.mc.vanderbilt.edu/brownlab/comlip.html.

Different lipids and proteins with interesting biological activities are found in nature. Their origin is an important factor. Those originating from the plant should gain more importance due to many advantages over the other sources in terms of availability, safety, and potential applications.

Plant lipidomics in association with plant genomics, plant proteomics, and plant metabolomics will contribute to defining molecular mechanisms of plant lipid and protein actions. Mapping plant lipid and protein contents would help in revealing the exact plant defense mechanisms in response to various stress conditions. Like glycomics, lipidomics and proteomics involve system-level identification and have vital roles in revealing metabolic pathways. Novel methodologies and analytical techniques have been greatly developed to separate the proteins and lipids and to analyze their compositions.

Further improvements in analytical methods, lipidomics and proteomics databases, and universal bioinformatic databases have to be further developed to allow effective utilization of these data of plant lipids and proteins.

Therefore, it is very important that diverse integrated approaches including advanced proteomic and lipidomic techniques combined with functional genomics, bioinformatics, metabolomics and/or with advanced molecular cell biology would be developed in the future as the coupling of these technologies will serve as a powerful tool in the development of plant science.

Regarding future perspectives, it is very interesting to study the evolutionary history of lipids and proteins in comparison with the different kingdoms of life origins.

6. Acknowledgments

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

The author thanks the Tunisian Ministry of Higher Education and Scientific Research for facilities.

The author is very grateful to Igor Pokotylo, a researcher at the "Université Paris-Est, Institut d'Ecologie et des Sciences de l'Environnement de Paris, Créteil-France" and "Institute of Bioorganic Chemistry and Petrochemistry, National Academy of Sciences of Ukraine, Kiev-Ukraine", for his help reading the manuscript.

7. References

Bansal, M., Sharma, M., Kanwar P., and Goyal, A. (2017). Recent advances in proteomics of cereals. Biotechnology and Genetic Engineering Reviews. DOI: 10.1080/02648725.2016.1272357.

Ben Halima, N. (2017). Occurrence and Importance of Plant Lipids: A Promising Insight into Algae. Recent Pat. Food Nutr. Agric. 8, 166-174.

Ben Halima, N. (2018). Why is it important to use algae in aquaculture? Journal of Biochemistry and Biotechnology Volume 1, Issue 1.

Ben Halima, N. (2019). *Unique Sequence Signatures in Plant Lipolytic Enzymes: Emerging Research and Opportunities.* Hershey, PA: IGI Global.

Brown, H.A. (Ed.). (2007). *Lipidomics and bioactive lipids: Lipids and cell signaling.* Academic Press, Elsevier.

Caffrey, M., and Hogan, J. (1992). LIPIDAT: a database of lipid phase transition temperatures and enthalpy changes. DMPC data subset analysis. *Chem. Phys. Lipids. 61,* 1-109.

Fischer, M., and Pleiss, J. (2003). The Lipase Engineering Database: a navigation and analysis tool for protein families. *Nucleic Acids Research, 31,* 319–321.

Hotelier, T., Renault, L., Cousin, X., Negre, V., Marchot, P., and Chatonnet, A. (2004). ESTHER, the database of the α/β-hydrolase fold superfamily of proteins. *Nucleic Acids Research, 32,* D145-D147.

Huang, A.H.C. (1996). Oleosins and oil bodies in seeds and other organs. Plant Physiol. 110, 1055-1061.

Huang, A.H.C. (1983). Plant lipases. in: H.L. Brockman, D. Borgstrom (Eds.), Lipases. Elsevier Publishing, Amsterdam, pp. 419-442.

Iwaniak, A., Minkiewicz, P., Darewicz, M., Sieniawski, K., and Starowicz, P. (2016). BIOPEP database of sensory peptides and amino acids. Food Research International. 85, 155-161.

Lenfant, N., Hotelier, T., Velluet, E., Bourne, Y., Marchot, P., and Chatonnet, A. (2013). ESTHER, the database of the α/β-hydrolase fold superfamily of proteins: tools to explore diversity of functions. *Nucleic Acids Research, 41,* D423–D429.

Minkiewicz, P., Dziuba, J., Iwaniak, A., Dziuba, M., and Darewicz, M. (2008). BIOPEP database and other programs for processing bioactive peptide sequences. Journal of AOAC International. 91, 965-980.

Murphy, D.J. (1993). Structure, function and biogenesis of storage lipid bodies and oleosin in plant. Prog. Lipid Res. 32, 247-280.

Namasivayam, E., Kowsalya, R., Padarthi, P.K., Manigandan, K., Jayaraj, R.L., Johnravindar, D., and Jagatheesh, K. (2015). Plant lipidomics: Signalling and analytical strategies. In D. Barh et al. (eds.), *PlantOmics: The Omics of Plant Science.* (pp. 331-356). Springer India.

Okazaki, Y., Kamide, Y., Hirai, M.Y., and Saito, K. (2013). Plant lipidomics based on hydrophilic interaction chromatography coupled to ion trap time-of-flight mass spectrometry. Metabolomics 9, S121–S131.

Panzanaro, S., Nutricati, E., Miceli, A., De Bellis, L. (2010). Biochemical characterization of a lipase from olive fruit (*Olea europaea* L.). Plant Physiology and Biochemistry. 48, 741-745.

Šamaj, J., and Thelen, J.J. (Eds). (2007). *Plant Proteomics.* Springer-Verlag Berlin Heidelberg.

Sud, M., Fahy, E., Cotter, D., Brown, A., Dennis, E.A., Glass, C.K., Merrill, AH Jr., Murphy, R.C., Raetz, C.R., Russell, D.W., and

Subramaniam, S. (2007). LMSD: LIPID MAPS structure database. *Nucleic Acids Res. 35,* D527-532.

Yasugi, E., and Watanabe, K. (2002). LIPIDBANK for Web, the newly developed lipid database. *Tanpakushitsu Kakusan Koso. Protein, nucleic acid, enzyme. 47,* 837-841.

8. Additional Reading

Ben Halima, N., Khemakhem, B., Fendri, I., Ogata, H., Baril, P., Pichon, C. and Abdelkafi, S. (2016). Identification of a new oat β-amylase by functional proteomics. *Biochimica et Biophysica Acta, Proteins and Proteomics,* 1864: 52–61.

Morikawa, R., Tsujimoto, M., Arai, H., and Aoki, J. (2007). Phospholipase A_1 Assays Using a Radiolabeled Substrate and Mass Spectrometry. In Brown, H.A. (Ed.), *Lipidomics and Bioactive Lipids: Lipids and Cell Signaling* (pp. 1-14). Academic Press, Elsevier.

Müller, G., & Petry, S. (Eds.). (2004). *Lipases and Phospholipases in Drug Development: From biochemistry to molecular pharmacology.* Weinheim, Germany: WILEY-VCH Verlag GmbH & Co. KGaA.

Rosahl, S., and Feussner, I. (2005). Oxylipins. In Murphy, D.J. (Ed.), *Plant Lipids. Biology, Utilisation and Manipulation* (pp. 329-354). Blackwell Publishing. CRC Press.

Sandoval, G. (Ed.). (2012). *Lipases and Phospholipases: Methods and Protocols.* New York, USA: Humana Press.