María Gabriela Guevara
Gustavo Raúl Daleo *Editors*

Biotechnological Applications of Plant Proteolytic Enzymes

Springer

Biotechnological Applications of Plant Proteolytic Enzymes

María Gabriela Guevara • Gustavo Raúl Daleo
Editors

Biotechnological Applications of Plant Proteolytic Enzymes

Springer

Editors
María Gabriela Guevara
Biological Research Institute
University of Mar del Plata
Mar del Plata, Argentina

Gustavo Raúl Daleo
Biological Research Institute
University of Mar del Plata
Mar del Plata, Argentina

ISBN 978-3-319-97131-5 ISBN 978-3-319-97132-2 (eBook)
https://doi.org/10.1007/978-3-319-97132-2

Library of Congress Control Number: 2018956563

This Springer imprint is published by the registered company Springer Nature Switzerland AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

Preface

Proteases (also termed as proteolytic enzymes or proteinases) refer to a group of enzymes whose catalytic function is to hydrolyze peptide bonds of proteins. These enzymes are widely distributed in all plants, animals and microorganisms. Proteases fields of application are very diverse, including food science and technology, pharmaceutical industries and detergent manufacturing.

This book is a review about the results obtained during the last decade in biotechnological application of plant proteolytic enzymes. In the last years, plant proteolytic enzymes are the object of renewed attention from the pharmaceutical industry and biotechnology not only because of their proteolytic activity on a wide variety of proteins but also because often they are active over a range of temperatures and pHs. All these have stimulated the research and increased the number of works on plant proteolytic enzymes.

Mar del Plata, Argentina

María Gabriela Guevara
Gustavo Raúl Daleo

Acknowledgements

This work was supported in part by the National Scientific and Technical Research Council grant (CONICET) to M.G.G., Scientific Research Commission of the Province of Buenos Aires (CIC) grants to M.G.G. and G.D. and University of Mar del Plata grant to M.G.G. and G.D. M.G.G. is an established researcher of CONICET, and G.D. is an established researcher of CIC.

Contents

Chapter 1
An Overview of Plant Proteolytic Enzymes

D'Ipólito Sebastián, María Gabriela Guevara, Tito Florencia Rocío, and Tonón Claudia Virginia

Contents

1.1 Introduction

Enzymes are proteins that act as highly efficient catalysts in biochemical reactions. This catalytic capability is what makes enzymes unique and they work efficiently, rapidly, and are biodegradable. The use of enzymes frequently results in many benefits that cannot be obtained with traditional chemical treatments. These often include higher product quality and lower manufacturing cost, less waste, and reduced energy consumption. Industrial enzymes represent the heart of biotechnology processes and biotechnology (Whitehurst and van Oort 2009; Sabalza et al. 2014)

D. Sebastián · M. G. Guevara (✉) · T. F. Rocío
Biological Research Institute, National Council of Scientific and Technique Research (CONICET), University of Mar del Plata (UNMDP), Mar del Plata, Argentina
e-mail: gguevara@mdp.edu.ar

T. C. Virginia
Biological Research Institute, National Council of Scientific and Technique Research (CONICET), University of Mar del Plata (UNMDP), Mar del Plata, Argentina

Scientific Research Commission of the Province of Buenos Aires (CIC), Buenos Aires, Argentina

M. G. Guevara, G. R. Daleo (eds.), *Biotechnological Applications of Plant Proteolytic Enzymes*, https://doi.org/10.1007/978-3-319-97132-2_1

Proteases are enzymes that hydrolyze peptide bonds of proteins and, from the point of view of industry, are the most important type of enzymes because they represent ca. 60% of all commercialized enzymes in the world (Feijoo-Siota and Villa 2011). These enzymes polarize the carbonyl group of the substrate peptide bond by stabilizing the oxygen in an oxyanion hole, which makes the carbon atom more vulnerable for attack by an activated nucleophile. Proteases can do this in four major ways, which gives the names to four catalytic classes: cysteine proteases, serine proteases, metalloproteases, and aspartic proteases (Dunn 2002).

These enzymes are widely distributed in all plants, animals, and microorganisms. Proteases account for approximately 2% of the human genome and 1–5% of genomes of infectious organisms (Puente et al. 2003). In plants, the *Arabidopsis* genome encodes over 800 proteases, which are distributed over almost 60 families, which belong to 30 different clans (van der Hoorn 2008). The distribution and the family size are well conserved within the plant kingdom because poplar and rice have similar distributions (García-Lorenzo et al. 2006).

1.2 Classification of Proteases

1.2.1 Catalyzed Reaction

This is the main property established by the Enzyme Nomenclature of the Committee of the International Union of Biochemistry and Molecular Biology (ENCIUBMB) that classifies these enzymes within group 3 (hydrolases) and subgroup 4 (hydrolases of peptide bonds). Subclass 3.4 may be in turn subdivided into endo- or exopeptidases (amino-terminal or carboxy-terminal) depending on their ability to hydrolyze internal peptide bonds or bonds located at the ends; endopeptidases are by far more important from the industrial point of view (Barrett 1994).

1.2.2 Nature of the Active Site

Hartley set a classification of proteases according to their catalytic site (Hartley 1960), in which proteases were divided into six mechanistic classes: serine endopeptidases (EC 3.4.21); cysteine endopeptidases, formerly denoted as thiol proteases (EC 3.4.22); aspartic endopeptidases, first known as acid proteases; glutamic endopeptidases (EC 3.4.23); metalloendopeptidases (EC 3.4.24); and threonine endopeptidases (EC 3.4.25), with a fifth group including peptidases with unidentified mode of action (EC 3.4.99).

Serine, threonine, and cysteine proteases are catalytically very different from aspartic and metalloproteases in that the nucleophile of catalytic site of the former group is part of an amino acid, whereas it is an activated water molecule for the last two types.

1.2.3 Structure-Based Evolutive Relationships

Rawlings and Barrett (1994) classified proteases according to their amino acidic sequence and relationships in families and clans. Proteases were placed within the same family if they shared sufficient sequence homology, and families believed to have a common ancestor placed within the same clan. The names of clans and families in the MEROPS database are built on the letters S, C, T, A, G, M, and U, which refer to the catalytic types. However, some of the clans are mixed type and contain families with two or more catalytic types and designated with the letter "P." This classification resulted in the creation of the MEROPS peptidase database, which is constantly revised: http://merops.sanger.ac.uk (Rawlings et al. 2010). The plant proteases most frequently used belong to the groups of cysteine, aspartic, and serine proteases (SPs).

1.3 Plant Proteases

Plant proteases are involved in many aspects of plant physiology and development (van der Hoorn 2008). They play a pivotal role in processes such as protein turnover, degradation of misfolded proteins, senescence, and the ubiquitin/proteasome pathway (Beers et al. 2000). Proteases are also responsible for the posttranslational modification of proteins by limited proteolysis at highly specific sites (Schaller 2004). They are involved in a great diversity of cellular processes, including photoinhibition in the chloroplast, defense mechanisms, programmed cell death, and photomorphogenesis in the developing seedling (Estelle 2001). Proteases are thus involved in all aspects of the plant life cycle ranging from mobilization of storage proteins during seed germination to the initiation of cell death and senescence programs (Schaller 2004).

1.3.1 Plant Cysteine Proteases

Fourthly, Cysteine proteases (CPs) family are recognized until today in which the nucleophile is the sulfhydryl group of a cysteine residue. The catalytic mechanism is similar to that of serine-type peptidases in that the nucleophile and a proton donor/general base are required, and the proton donor in all cysteine peptidases is a histidine residue as in the majority of the serine entered forms (Domsalla and Melzig 2008). Although there is evidence in some families that a third residue is required to orientate the imidazolium ring of the histidine, a role analogous to that of the essential aspartate seen in some serine peptidases. There are a number of families in which only a catalytic dyad is necessary (Barrett et al. 1998).

According to the MEROPS database, CPs are divided into ten clans: CA, CD, CE, CF, CH, CL, CM, CN, CO, and C-, and to date, plant CPs have been described as belonging to five of these clans (CA, CD, CF, CO, and CE). Most plant CPs belong to the C1 family, also known as the papain family (clan CA). A complete list of CPs may be found in the MEROPS database (Rawlings et al. 2010).

Plants offer an attractive alternative for the production of CPs as they occur naturally in different tissues, in some cases in excessive amount (González-Rábade et al. 2011). CPs of the tropical plants *Carica papaya* (papain, chymopapain, caricain, and glycyl endopeptidase), *Ananas comosus* (fruit bromelain, stem bromelain, ananain, and comosain), and *Ficus glabrata* (ficin) are of considerable commercial importance, due to their strong proteolytic activity against a broad range of protein substrates and because they are active over a broad range of pH and temperature (Feijoo-Siota and Villa 2011). CPs have been isolated from: PsCYP15A from pea roots (Vincent and Brewin 2000); GP2 and GP3 from ginger rhizomes (Kim et al. 2007), from sweet potato (*Ipomoea batatas*) roots (Huang and McDonald 2009); FLCP-1 and FLCP-3 from *Phaseolus* leaves (Popovič et al. 2002); and bromelain (Rowan et al. 1990) and ananain (Lee et al. 1997) from stems. The most ubiquitous group are found in fruits, i.e., balansain I, macrodontain I in *Bromeliaceae* (Pardo et al. 2000; López et al. 2000), araujiain in *Asclepiadaceae* (Priolo et al. 2000); papain, chymopapain, papaya glycyl endopeptidase, and caricain from *C. papaya* latex (Azarkan et al. 2003). Papain-like cysteine proteases are usually lysosomal (vacuolar) or secreted proteins (Dubey et al. 2007).

According to the review by González-Rábade et al. (2011), proteases like papain, bromelain, and ficin are employed in different industrial processes and medicines (Uhlig and Linsmaier-Bednar 1998). Some of these proteases are used in the food industry for cheese, brewing and beverage industries for the preparation of highly soluble and flavored protein hydrolysates (papain-like proteases), as a food complement (Kleef et al. 1996; La Valle et al. 2000; Losada Cosmes 1999) to soften meats and dehydrated eggs (Bailey and Light 1989; Lawrie 1985; Miller 1982), and for the production of emulsifiers, among other uses (Pardo et al. 2000). Uses in other industries include culture medium formulation (Headon and Walsh 1994), isolation of genetic material (Genelhu et al. 1998), and the use of keratinases in the leather industry for dehairing and bating of hides to substitute toxic chemicals (Foroughi et al. 2006). Also, they are used in the production of essential amino acids such as lysine and for the prevention of clogging of wastewater systems (Rao et al. 1998). Proteases also have an important application in the pharmaceutical industry. Plant extracts with a high content of proteolytic enzymes have been used in traditional medicine for a long time. They have been used for the treatment of cancer (Batkin et al. 1988; Targoni et al. 1999), as antitumorals (Guimarães-Ferreira et al. 2007; Otsuki et al. 2010), for digestion disorders (Kelly 1996; Mello et al. 2008), and swelling and immune-modulation problems (Leipner et al. 2001; Lotti et al. 1993; Melis 1990; Otsuki et al. 2010). A good example is bromelain, derived from pineapple, which has been shown to be capable of preventing edema, platelet aggregation, and metastasis due to its capacity of modifying cell surface structures by

peptide cleavage. Salas et al. (2008) reviewed the pharmacological activity of plant cysteine proteases, emphasizing their role in mammalian wound healing, immunomodulation, digestive conditions, and neoplastic alterations.

1.3.2 Plant Aspartic Proteases

Aspartic proteases (APs, EC 3.4.23) are a family of proteolytic enzymes widely distributed among living organisms and are found in vertebrates, plants, yeast, nematodes, parasites, fungi, and viruses (Rawlings and Salvesen 2013). Aspartic proteases differ from the serine and cysteine peptidases in the way that the nucleophile that attacks the scissile peptide bond is an activated water molecule rather than the nucleophilic side chain of an amino acid (Domsalla and Melzig 2008).

Plant APs have been characterized and purified from a variety of tissues such as seeds, flowers, and leaves: (1) seeds of *Arabidopsis thaliana* (Mutlu et al. 1999), rice (Asakura et al. 1997; Doi et al. 1980), barley (Kervinen et al. 1999; Runeberg-Roos et al. 1991; Sarkkinen et al. 1992), hempseed (St. Angelo et al. 1969, 1970), cucumber, and squash (Polanowski et al. 1985); (2) leaves of the tomato plant (Rodrigo et al. 1989); (3) leaves and tubers of the potato plant (Guevara et al. 2001, 2004); (4) maize pollen (Radlowski et al. 1996); and (5) flowers of thistle (Heimgartner et al. 1990; Verissimo et al. 1996), among others. Some of these APs, like the ones found in barley, resemble mammalian cathepsin D. It has been suggested that plant APs are involved in the digestion of insects in carnivorous plants (Garcia-Martinez and Moreno 1986; Takahashi et al. 2009; Tökés et al. 1974), in the degradation of plant proteins in response to pathogens (Rodrigo et al. 1989, 1991), during development processes (Asakura et al. 1997; Runeberg-Roos et al. 1994), protein-storage processing mechanisms (Doi et al. 1980; Hiraiwa et al. 1997), stress responses (de Carvalho et al. 2001; Guevara et al. 1999, 2001), and senescence (Bhalerao et al. 2003; Buchanan-Wollaston 1997; Cordeiro et al. 1994; Lindholm et al. 2000; Panavas et al. 1999). These enzymes are distributed among families A1, A3, A11, and A12 of clan AA and family 22 of clan AD (Faro and Gal 2005; Mutlu et al. 1999; Rawlings et al. 2014; Simões and Faro 2004). The majority of plant APs have common characteristics as that of AP A1 family, are active at acidic pH, are specifically inhibited by pepstatin A, and have two aspartic acid residues responsible for the catalytic activity (Simões and Faro 2004).

Plant APs are classified into three categories: typical, nucellin-like, and atypical (Faro and Gal 2005). The swaposin domain is only present in typical plant APs inserted into the C-terminal domain as an extra region of approximately 100 amino acids known as "plant-specific insert" (PSI) (Simões and Faro 2004). The PSI domain has a high structural homology with saposin-like proteins (SAPLIPs), a large protein superfamily widely distributed from primitive eukaryotes to mammalians (Bruhn 2005; Michalek and Leippe 2015). Individual SAPLIPs generally share little amino acid sequence identity. However, SAPLIP protein sequences include highly conserved cysteine residues that form disulfide bonds and give

SAPLIPs a stable structure; secondary protein structure consists mainly of α-helices joined by loops (Andreu et al. 1999; Bruhn 2005; Munford et al. 1995). The SAPLIPs or Sap domains may exist for itself independently as a functional unit or as a part of a multidomain protein; they are autonomous domains with a variety of different cellular functions, all of them associated with lipid interaction. SAPLIP activities are classified into three major groups: (1) membrane targeting by the SAPLIP domain; (2) presentation of lipids as substrate for an independent enzyme, either by extraction from the membrane or by disturbance of the well-packed lipid order; and (3) membrane permeabilization by perturbation owing to single molecules or by pore formation of oligomeric proteins (Bruhn 2005). The SAPLIP family includes saposins, which are lysosomal sphingolipid-activator proteins (O'Brien and Kishimoto 1991), NK-lysin, granulysin, surfactant protein B, amoebapores, domains of acid sphingomyelinase and acyloxyacyl hydrolase, and the PSI domain of plant APs (Munford et al. 1995; Stenger et al. 1998; Vaccaro et al. 1999).

The PSI domains of plant APs are named swaposins since they arise from the exchange (swap) of the N- and C-terminal portions of the saposin-like domain, where the C-terminal portion of one saposin is linked to the N-terminal portion of the other saposins (Simões and Faro 2004). This segment is usually removed during the proteolytic maturation of the heterodimeric typical plant APs (Davies 1990; Domingos et al. 2000; Faro and Gal 2005; Glathe et al. 1998; Mutlu et al. 1999; Ramalho-Santos et al. 1997; Törmäkangas et al. 2001; White et al. 1999). However, in monomeric typical plant APs, the PSI domain is present in the mature protein (Guevara et al. 2005; Mendieta et al. 2006). *Solanum tuberosum* APs 1 and 3 (*St*APs 1 and 3) are included into the group of monomeric typical plant APs (Guevara et al. 1999, 2001, 2005).

All enzymes employed commercially in milk coagulation are APs, with acidic optima pH, and high levels of homology between their primary structures and similarity between their catalytic mechanisms (Silva and Malcata 2005).

The most widely used AP is rennet, which has chymosin as its active component (Vioque et al. 2000). Rennet is obtained from the stomach of calves but it is costly and scarce. Most companies produce recombinant rennet of calf origin in different microbial hosts (Seker et al. 1999). Some plant APs have shown to possess similar characteristics to calf-derived rennet and hence have attracted attention in the food industry. In Portugal and some regions of Spain, the use of extracts from dried flowers of *C. cardunculus* L. has been successfully maintained since ancient times for the production of many traditional varieties of sheep and goat cheeses, further strengthening the suitability of this rennet for the production of high-quality cheeses (Reis and Malcata 2011; Roseiro et al. 2003a, b; Sousa and Malcata 2002). Therefore, over the last decades, much effort has been made in understanding the properties of this unique plant coagulant. Several authors have dedicated their research efforts to characterize the milk-clotting enzymes present in cardoon flowers, their role in the hydrolysis of caseins in vitro, and their effect in the proteolysis process during ripening (Agboola et al. 2004; Brodelius et al. 1995; Esteves et al. 2001; Esteves et al. 2003; Faro et al. 1992; Ramalho-Santos et al. 1996; Roseiro et al. 2003a, b; Silva et al. 2003; Silva and Xavier Malcata 1998; Silva and

Malcata 1999, 2000, 2005; Sousa and Malcata 1997, 1998, 2002). Due to the potential of these cardoon enzymes to serve as alternative rennets in larger-scale production processes, several strategies have also been undertaken in more recent years to either develop more standardized formulations of the native enzymes or explore their production in heterologous systems to generate synthetic versions of these proteases (Almeida and Simões 2018).

Several APs from *Cynara* sp. have been cloned and expressed successfully in bacterial and eukaryotic expression systems. The production by yeast of the cyprosin B (cynarase 3) has been patented (Planta et al. 2000). This was the first clotting protease of plant origin produced by fermentation technology (Sampaio et al. 2008). Recombinant cyprosin produces a proteolysis similar to that obtained with natural enzymes present in the crude extract from *C. cardunculus* (Fernández-Salguero et al. 2003).

In this regard, studies with APs from *Solanum tuberosum* have revealed their antimicrobial activity (Guevara et al. 2002) including antifungal activity (Mendieta et al. 2006). The authors cloned, expressed, and purified PSI (swaposin domain) from *St*APs (*St*Asp-PSI) and found that the recombinant protein still maintained its cytotoxic activity (Muñoz et al. 2010). Both *St*APs and *St*Asp-PSI were able to kill human pathogenic bacteria in a dose-dependent manner but were not toxic to human red blood cells under the experimental conditions tested. Therefore, *St*APs and *St*Asp-PSI could contribute to the generation of new tools to solve the growing problem of resistance to conventional antibiotics (Zasloff 2002).

1.3.3 Plant Serine Proteases

Serine proteases (SPs) use the active site Ser as a nucleophile. The catalytic mechanism is very similar to that of cysteine proteases, and some serine proteases are even evolutionarily related to cysteine proteases. With more than 200 members, serine proteases are the largest class of proteolytic enzymes in plants. Plant serine proteases are divided into 14 families. These families belong to nine clans that are evolutionarily unrelated to each other. Families S8, S9, S10, and S33 are the largest serine protease families in plants, with each containing approximately 60 members. Biological functions for serine proteases have been described for some of the subtilases (SDD1 and ALE1; family S8, clan SB), carboxypeptidases (BRS1 and SNG1/2; family S10, clan SC), and plastid-localized members of the S1, S26, and S14 families (DegPs, Plsp1, and ClpPs) (van der Hoorn 2008).

Feijoo-Siota and Villa (2011) have reviewed several origins to SPs. These enzymes have been found and extracted from the seeds of barley (*Hordeum vulgare*), soybean (*Glycine max*), and rice (*Oryza sativa*), from the latex of *Euphorbia supina*, *Wrightia tinctoria*, dandelion (*Taraxacum officinale*), African milkbush (*Synadenium grantii*), and jackfruit (*Artocarpus heterophyllus*); from the flowers, stems, leaves, and roots of *Arabidopsis thaliana*; from the storage roots of sweet potato (*I. batatas*) and corn (*Zea mays*); from the sprouts of bamboo (*Pleioblastus hindsii*); from the leaves of

tobacco (*Nicotiana tabacum*), lettuce (*Lactuca sativa*), common bean (*Phaseolus vulgaris*), and tomato (*Lycopersicon esculentum*); and from the fruits of melon (*Cucumis melo*), *Cucurbita ficifolia*, osage orange (*Maclura pomifera*), suzumeuri (*Melothria japonica*), "Ryukyu white gourd" (*Benincasa hispida*), Japanese large snake gourd (*Trichosanthes bracteata*), and yellow snake gourd (*Trichosanthes kirilowii*) (Antão and Malcata 2005).

SPs are usually considered to act principally as degradative enzymes. Plant subtilases (subtilisin-like SPs), however, have been documented to be involved in several physiological processes including symbiosis (Takeda et al. 2007), hypersensitive response, the infection of plant cells (Laplaze et al. 2000), pathogenesis in virus infected plants (Tornero et al. 1997), germination (Sutoh et al. 1999), signaling (Déry et al. 1998), tissue differentiation (Groover and Jones 1999), xylogenesis (Ye and Varner 1996), senescence (Distefano et al. 1999; Huffaker 1990), programmed cell death (Beers et al. 2000), and protein degradation/processing (Antão and Malcata 2005).

Cucumisin, an enzyme derived from melon, remains the best plant SP characterized to date (Antão and Malcata 2005), purified from *Cucumis melo* (Kaneda and Tominaga 1975). Plant subtilisins, also referred to as cucumisin-like proteases (in recognition of the first subtilisin isolated from a plant (Yamagata et al. 1994)), have been isolated from *Cucumis melo* (Yamagata et al. 1994), *Solanum lycopersicum* (Meichtry et al. 1999), *Alnus glutinosa* (Ribeiro et al. 1995), and *Arabidopsis* (Zhao et al. 2000). Cucumisin-like SPs have also been isolated from other sources, like the latex of *E. supina* (Arima et al. 2000a; Taylor et al. 1997), the sprouts of bamboo (*Pleioblastus hindsii* Nakai) (Arima et al. 2000b), and the fruits of *Melothria japonica* (Uchikoba et al. 2001). A cucumisin-like protease from kachri fruit (*Cucumis trigonus Roxburghi*) is used as a meat tenderizer in the Indian subcontinent (Asif-Ullah et al. 2006).

Macluralisin, from the fruits of *M. pomifera* (Raf.) Schneid (Rudenskaya et al. 1995), taraxilisin, from the latex of dandelion (*T. officinale Webb s. I.*) roots (Rudenskaya et al. 1998), SP A and B, from the sarcocarp of yellow snake gourd (*Trichosanthes kirilowii*) (Uchikoba et al. 1990), an SP from the seeds of tropical squash (*C. ficifolia*) (Dryjanski et al. 1990), and several other SPs from barley (Fontanini and Jones 2002), oat (Coffeen and Wolpert 2004), soybean (Tan-Wilson et al. 1996), and common bean (Popovič et al. 2002) are other SPs isolated from plants.

Some SPs have been studied for their medicinal properties (Andallu and Varadacharyulu 2003; Andallu et al. 2001; Doi et al. 2000; Jang et al. 2002; Andallu and Varadacharyulu 2003; Andallu et al. 2001; Doi et al. 2000; Jang et al. 2002), examples of these are the a subtilisin-like SP, named as indicain, isolated from the latex of *Morus indica* by Singh et al. (2008) and milin, an SP purified from the latex of *Euphorbia milii* (Yadav et al. 2006). Many of the medicinal applications have been proved by clinical studies (Asano et al. 2001; Cheon et al. 2000; Doi et al. 2001; Nomura 1999). Milin, another SP of plant origin, isolated from *E. milii*, is a good candidate for applications in the food industry (Souza et al. 1997; Schall et al. 2001).