UNIVERSITÀ DEGLI STUDI DI MILANO

PhD program in Biochemical Sciences

XXIX cycle

DeFENS - Department of Food, Environmental and Nutritional Sciences

BIOCHEMICAL FUNCTIONAL CHARACTERIZATION AND MOLECULAR BIOLOGY OF PLANT INHIBITOR PROTEINS ACTING AGAINST GLYCOSIDE HYDROLASE

Elisabetta GALANTI

Tutor: Prof. Alessio SCARAFONI

Director of the PhD program: Prof. Sandro SONNINO

A.A. 2015-2016

C’accht Crist n’de lupin?

Sommario

ABBREVIATIONS

ANXI: *Aspergillus niger* xylanase I

CESAs: cellulose synthase proteins

CSCs: cellulose synthase complexes

CSL: cellulose synthase-like (different family indicated with letter)

CWDEs: cell wall-degrading enzymes

DPA: days post-anthesis

EDGP: Extracellular dermal glycoprotein from carrot

EGL12s: endo-β-1,4-glucanase GH12 from *Phytophthora infestans*

G: unsubstituted glucosyl

GalUA: galacturonic acid units

GAX: heteroxylans containing more glucuronic acid and 4-O-methyl glucuronosyl residues

GH: Glycoside hydrolases (different family indicated with number)

GHIPs: glycoside hydrolase inhibitor proteins

GOX:

GT: Glycosyltransferases (different family indicated with number)

GX: Glucuronoxylan

HG: homogalacturonan

HVXI: *Hordeum vulgare* L. xylanase inhibitor

IL1 – IL2: inhibitor loop 1 and inhibitor loop 2

IL1; 1L2: in inhibition loop 1; in inhibition loop 2

MdXEGIP1: xyloglucan-specific endo-β-1,4-glucanase inhibitor proteins from *Malus domestica*

MLG: Mixed-linkage glucan

NbXEGIP2: xyloglucan-specific endo-β-1,4-glucanase inhibitor proteins from *Nicotiana benthamiana*

NEC4: nectarin 4 from *Nicotiana langsdorffii x Nicotiana sanderae*

NEC5: nectarin 5 from *Nicotiana langsdorffii x Nicotiana sanderae*

PAMP: pathogen-associated molecular pattern

PGIPs: polygalacturonase inhibiting proteins

RGI: rhamnogalacturonan

RGII: rhamnogalacturonan II

ROS: reactive oxygen species

SCXI: *Secale cerelae* L. xylanase inhibitor from isoform I to isoform IV

TAXIs: *Triticum aestivum* L. endoxylanase inhibitors

TDXI: *Triticum durum* L. endoxylanase inhibitors isoform I and isoform II.

TLXIs: thaumatin-like xylanase inhibitors

X: xylosyl residue

XEG: xyloglucan-specific endo-β-1,4-glucanase from *Aspergillus aculeatus*

XEGIPs-like: xyloglucan-specific endo-β-1,4-glucanase inhibitor proteins

XEH: xyloglucan endohydrolase

XET: xyloglucan endotransglucosylase

XGA: xylogalacturonan

XIPs: xylanase inhibitor proteins

XIs: xylanase inhibitors

XTH: xyloglucan endotransglucosylase/hydrolase

XyG: xyloglucan

INTRODUCTION

1. Plant cell wall:

Plant cells wall is a tough, rigid and made mostly of sugars (up to 75% of dry mass) structure that may consist of three layers: the primary cell wall, the secondary cell wall and the middle lamella. The middle lamella is the first layer formed during cell division, it is a pectin-rich intercellular material that glues the adjacent cells together. The primary wall is a thin flexible and extensible layer composed of cellulose, pectin and combinations of matrix molecules that can be composed of mixtures off different proportions of xylan, xyloglucan, mannan or mixed linkage glucan (collectively called hemicelluloses), besides structural proteins and lignin. The secondary cell wall is a thick layer rich in lignin that strengthens and waterproofs the wall; it is formed inside the primary cell wall that has stopped increasing in surface area when the cell is fully grown.

1.1 Structure

1.1.1 Cellulose

Cellulose was discovered in 1838 by Anselme Payen, who isolated it from plant matter and determined its chemical formula $(C_6H_{10}O_5)_n$. Cellulose molecules are long, unbranched chains of β1,4-linked glucose units. Each glucose is inverted with respect to its neighbours, resulting in a linear chain of at least 500 glucose residue that are covalently linked to one another to form a ribbon-like structure, which is stabilized by hydrogen bonds. In addition, intermolecular hydrogen bonds between adjacent cellulose molecules cause them to adhere strongly to one another in overlapping parallel arrays, forming a bundle of about 40 cellulose chain, all of which have the same polarity (Alberts et al. 2002). These highly ordered crystalline aggregates are called cellulose microfibrils.

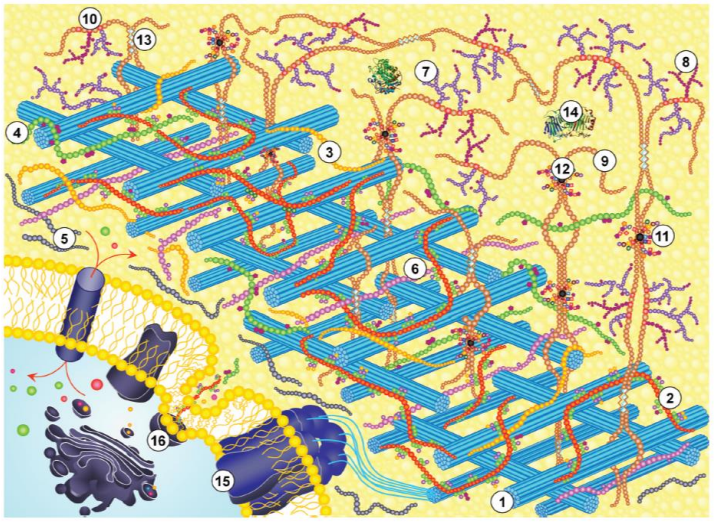


Fig. 1: Model of the polysaccharide framework in a plant cell wall, generalized for poalean and non-poalean walls.
1, Cellulose: cellulose microfibrils;
2–6, hemicelluloses: **2**, xyloglucan; **3**, mixed-linkage glucan; **4**, xylan and related heteroxylans; **5**, callose; **6**, mannan and related heteromannans;
7–11, Pectins: **7**, galactan; **8**, arabinan; **9**, homogalacturonan;**10**, rhamnogalacturonan I; **11**, rhamnogalacturonan II;
12, boron bridge; **13**, 'egg-box' with calcium bridges;
14–16, Non-polysaccharide components: **14**, enzymes and structural proteins; **15,** cellulose synthase complex; **16**, transport vesicles

1.1.2 Pectins

Pectins, first isolated and described in 1825 by Henri Braconnot, among the most complex and heterogeneous branched polysaccharides that contain many negatively charged galacturonic acid units. Because of their negative charge, pectins are highly hydrated and associated with a cloud of cations. The pectins include homogalacturonan (HG), xylogalacturonan (XGA), apiogalacturonan, rhamnogalacturonan I (RGI), and rhamnogalacturonan II (RGII). The ratio between HG, XGA, RGI, and RGII is also variable, but typically HG is the most abundant polysaccharide, constituting about 65% of the pectin, while RGI constitutes 20% to 35% (Mohnen 2008). XGA and RGII are minor components, each constituting less than 10% (Zandleven et al. 2007; Mohnen 2008). The backbone of galacturonic acid units (GalUA) residues can be substituted at various positions with other sugar moieties. In XGA, a single Xyl is attached to the O-3 position of some GalUA residues. Additional Xyl residues can be attached to the first Xyl with β-1,4 linkage (Zandleven et al. 2006). Besides the simple substitutions in XGA and apiogalacturonan, clusters of complex side chains are also attached onto the O-2 or O-3 position in the galacturonan backbone to form RGII (Harholt et al. 2010). Despite its complexity, the structure of RGII is highly conserved among vascular plants (Matsunaga et al. 2004; O'Neill et al. 2004). Frequently, RGII exists as a dimer mediated by borate ion attached to the A branch. Boron is an essential micronutrient for plants, most likely due to an important role of RGII dimerization in ensuring the integrity of the cell wall (O'Neill et al. 2001). RGI is the only type of pectin not built upon pure galacturonan backbones. Instead, it is a branched polymer with a backbone of disaccharide (α-1,4-D-GalA-α-1,2-L-Rha) repeats. The Rha residues in the backbone can be substituted with β-1,4-galactan, branched arabinan, and/or arabinogalactan side chains.

1.1.3 Hemicellulose

Hemicellulose are a group of complex polysaccharides, including xyloglucans, xylans and mannans, that are extracted from plant cell walls by use of strong alkali; characteristically they bind tightly to the surface of cellulose and have a long linear backbone made up of (1,4)-β-d-glycans (glucose, xylose or mannose) with an equatorial configuration, from which short side chains of other sugars protrude (Cosgrove 2005). The complexity of the cell wall organization increases with the cross-linking of hemicellulose leading to the aggregation of cellulose microfibrils into macrofibrils, structuring an entangles glucan assembly (Pauly et al. 2013).

1.1.3.1 Mannan

β-(1→4)-linked polysaccharides containing mannose are widely distributed and the main hemicellulose in Charophytes (Popper and Fry 2003; Popper 2008). The backbones may consist entirely of mannose, as in mannans and galactomannans, or with mannose and glucose in a nonrepeating pattern as in glucomannans and galactoglucomannans. Mannans and glucomannans are often acetylated (Scheller and Ulvskov 2010) (Fig. 1A). Mannans have been much studied in their role as seed storage compounds, in particular the seeds of many legumes are known to accumulate galactomannan in their endospermic cell walls. Moreover, in

gymnosperms, galactoglucomannans are major components of the secondary walls (Ebringerova et al. 2005) (Fig. 4). Mannans appear to have been very abundant in early land plants and are still abundant in mosses and lycophytes (Harholt et al. 2010, Moller et al. 2007)

1.1.3.2 Xyloglucan

XyG is a β-1,4 glucan that can be substituted with a diverse array of glycosyl and nonglycosyl residues (Pauly et al. 2013; Scheller and Ulvskov 2010). XyG is thought to form cross-links between cellulose microfibrils, forming a strong but extensible XyG-cellulose network that might function as the main load-bearing component of the primary cell wall. Because of the diversity of substituents and their linkages, a nomenclature based on one-letter codes has been. The letter **G** represents the unsubstituted glucosyl residue of the backbone. The most prominent substituent of XyG is an α-D-xylopyranose that is attached to the glucan backbone at O-6, and a backbone glucosyl residue that harbors such a xylosyl residue is denoted by the letter **X**. Both the backbone glucosyl and the xylosyl residue can be further substituted with D- and L-galactosyl, L-fucosyl, D-galacturonosyl, L-arabinopyranosyl, and/or L-arabinofuranosyl moieties at specific locations in specific linkages, resulting in the unique 24 structures identified to date (Pauly and Keegstra 2016). The structural features of XyG have been elucidated with the help of fungal XyG endoglucanases, enzymes that cleave the XyG polymer into oligosaccharides and then the released oligosaccharides have been structurally characterized in detail by mass spectrometry and NMR spectroscopy (Lerouxel et al. 2002; Tuomivaara et al. 2015; York et al. 1990). Based on the structure of the released XyG oligosaccharides, XyG polymers fall into one of two general types. In one type, three out of four backbone glucosyl residues are xylosylated, leading to an XXXG-type XyG (Vincken et al. 1997). This type is often galactosylated and fucosylated, resulting in the fucogalactoxyloglucan commonly present in most tissues of most dicots (Zablackis et al. 1995). Another type of XyG exhibits reduced xylosylation in that only two out of the four or more backbone glucosyl residues are xylosylated, resulting in the XXGGn-type XyG (Fig. 1B). This type often contains O-acetyl substituents on the glucan backbone, leading to the acetoxyloglucan common in many tissues of grasses. In many dicots xyloglucans constitute the major hemicellulose of growing cell walls, comprising ~20% of the dry mass of primary cell walls (Schultink et al. 2014). Grasses - but not monocots in general - have a reduced xyloglucan content; values of ~5% of primary walls are typical in grasses, but values as high as 10% occur (Carpita 1996; Gibeaut et al. 2005) (Fig 4). In *Arabidopsis thaliana* and many other dicots, O-acetylation of XyG is found principally on galactosyl residues at the O-3 or O-2 positions (Kiefer et al. 1989).

1.1.3.3 Xylan

Xylans are a diverse group of polysaccharides with the common feature of a backbone of β-(1→4)-linked xylose residues. Glucuronoxylans are substituted

with α-(1→2)-linked glucuronosyl and 4-O-methyl glucuronosyl residues and they are the dominating noncellulosic polysaccharide in the secondary walls of dicots. In commelinid monocots (which include grasses and some related species), xylans are the major noncellulosic polysaccharide in primary walls, constituting about 20% of the wall (Fig 4). Cereal endosperm arabinoxylan has very little glucuronic acid, but heteroxylans in vegetative parts of grasses are often called arabinoxylans, even though they tend to contain more glucuronic acid and 4-O-methyl glucuronosyl residues, making GAX a more appropriate name (Fig 1C). Gymnosperm walls also contain arabinoxylans in relatively high amounts (Scheller and Ulvskov 2010). Unlike XyGs, xylans do not have a repeated structure, and there are many variations in the structure that are not well known. An important feature of grass xylans is the presence of ferulic acid esters attached to O-5 of some of the arabinofuranosyl residues. Esters of p-coumaric acid are also abundant in grass cell walls, but it is not clear if they can be attached directly to the xylans, and they may be primarily associated with lignin (Hatfield et al. 2008). Ferulate can also be cross-linked with lignin (Grabber 2005) and we can therefore assume that GAX and lignin become covalently cross-linked through these linkages. Cross-linking through ferulate esters is widely assumed to render the cell wall recalcitrant to digestion, which would be an obvious benefit as a defence against microorganisms and herbivores.

1.1.3.4 Mixed-linkage glucan

MLG (1,3; 1,4-β-glucan) is a glucose-based unsubstituted, non-branched homopolymer, whereby randomly distributed β-1,4-linked cellotriosyl and cellotetraosyl units are connected by β-1,3 linkages. β-(1→4)-linked glucans with interspersed single β-(1→3)-linkages are well known in grasses (Fig. 1D). The occurrence of β-(1→3,1→4)-glucans in many primitive taxa could indicate that they represent an ancient trait. However, if β-(1→3,1→4)-glucan in grasses were a conserved ancient trait, then we would have to postulate the independent disappearance in a large number of Spermatophyte taxa that have been investigated (Fig 4). It seems more likely that β-(1→3,1→4)-glucan has evolved independently in grasses (Smith and Harris. 1999).

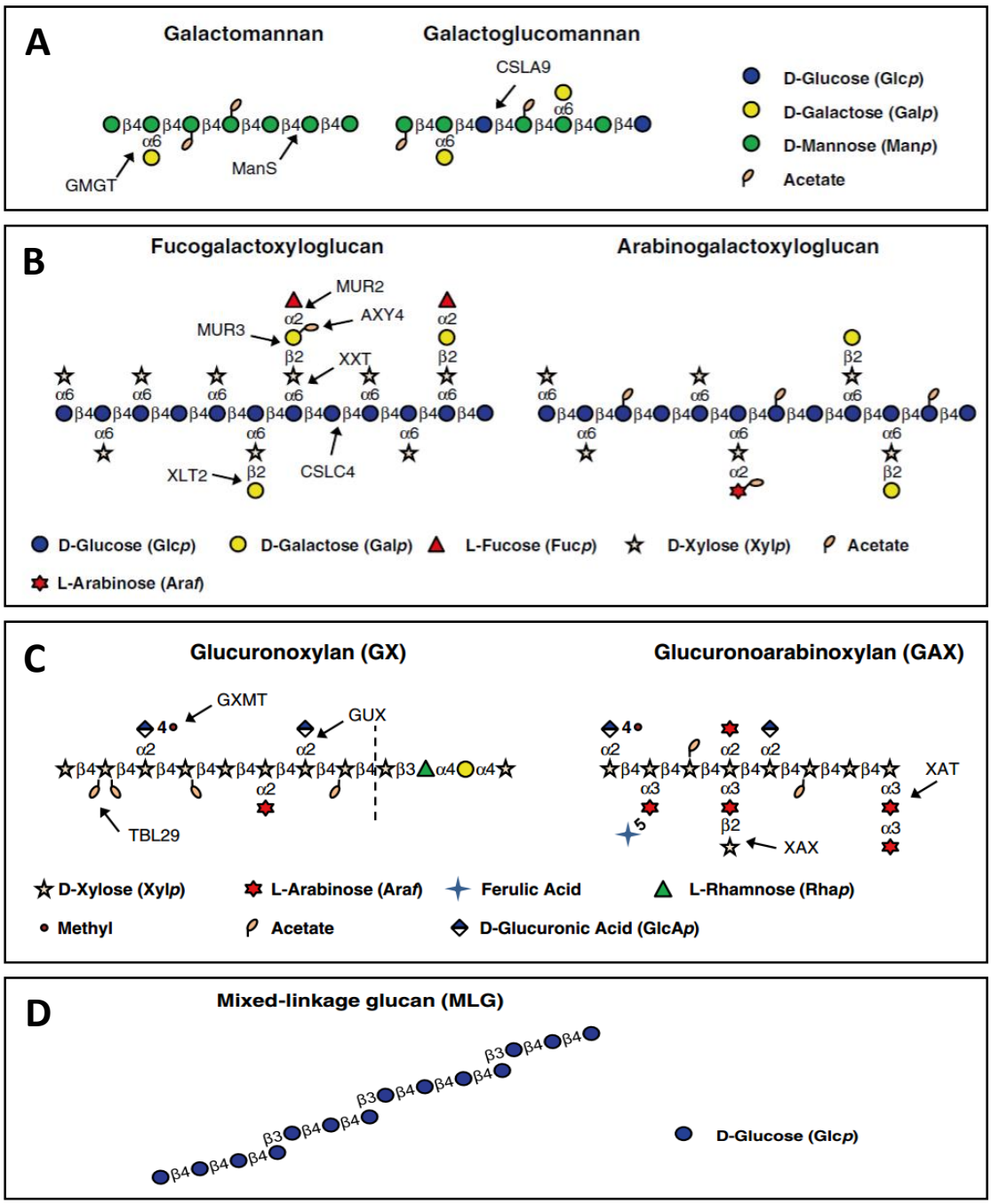


Fig. 2: Schematic representation of hemicellulose structure and known proteins involved in its synthesis. Symbols representing the various monosaccharides were adopted from the Nomenclature Committee Consortium for Functional Glycomics (Varki et al. 2009). Glycosidic linkages between monosaccharides are represented in their anomeric configurations (α or β) and their position. If glycosyltransferases are known to add a sugar in a certain linkage, they are indicated by their protein abbreviation and an arrow.
A: Mannan: **Galactomannan and Galactoglucomannan**; ManS Mannan synthase, CSLA9 Glucomannan synthase, Galactosyl transferase GMGT.
B: Xyloglucan: **Fucogalactoxyloglucan in dicots and Arabinogalactoxyloglucan in monocots/solanales**; CSLC4 Glucan synthase, XXT Xylosyl transferase, MUR3 Galactosyl transferase, XLT2 Galactosyl transferase, MUR2 Fucosyl transferase, AXY4 Acetyl transferase.
C: Xylan: **Glucuronoxylan GX and Glucuronoarabinoxylan GAX**; Xylan synthase, GUX Glucouronsyl transferase, XAX Xylosyl transferase, XAT Arabinosyl transferase, GXMT Methyltransferase, TBL29 Acetyl transferase
D: Mixed-linkage glucan MLG: Glucan synthase

1.2 Biosynthesis

1.2.1 Cellulose

Multiple glucan chain are simultaneously synthesized by a membrane-localized cellulose synthase complexes (CSCs), a rose shaped array of proteins in the plasma membrane that synthesize the cellulose microfibrils. Rosettes CSC is composed by six subunits each of that contains 4-6 enzymatically active cellulose synthase proteins (CESAs) (Li et al. 2014). CESAs are integral plasma membrane proteins with multiple transmembrane domains and a central catalytic domain (Richmond et Somerville 2000) (Fig. 2). Based on genetic analysis, CESAs are roughly classified into two groups that correspond to CESAs that are involved in cellulose synthesis in primary cell walls (primary CESAs) and CESAs that are involved in cellulose synthesis in secondary cell walls (secondary CESAs).

1.2.2 Pectin

Pectin is synthesized in the Golgi lumen by Golgi-localized glycosyltransferases (GTs) (Fig. 2) that transfer glycosyl residues from nucleotide-sugars onto oligosaccharide or polysaccharide acceptors. During synthesis some pectic glycosyl residues are modified by methyltransferase-catalyzed esterification or O-methylation, by acetyltransferase-catalyzed acetylation or feruloylation driven by feruloyltransferases. It has been estimated that as many as 67 different transferases are required for the biosynthesis of pectin (Mohnen 2008; Harholt et al. 2010)

1.2.3 Hemicellulose

1.2.3.1 Mannan

Heteromannans are synthesized from activated nucleotide sugars. For mannans these nucleotide sugars are GDP-mannose, GDP-glucose, and UDP-galactose (Liepman et al. 2005). The activated nucleotide sugars are then utilized by highly specific, Golgi-localized glycosyltransferases (GTs), which facilitate the formation of the specific linkage between the monomers and thus synthesize the polymer (Breton et al. 2006). β-mannan synthase (ManS) is a member of the cellulose synthase-like family A (CSLA) (Fig. 2) from GT family 2. (Pauly et al. 2013) The identification of an apoplastic mannan transglycosylase activity in plants suggests that once the polymer is deposited in the wall it can undergo further modification adding complexity to the biosynthesis process (Schröder et al. 2004).

1.2.3.2 X*yloglucan*

The β-(1,4)-D-glucan backbone of xyloglucans is synthesized by Golgi-localized glycan synthases, a member of the cellulose synthase-like family A (CSLA) (Fig. 2) (Cocuron et al. 2007). Sidechain substitution is tuned by different glycosyltransferases that add the same

glycosyl-moiety to different positions (Schultink et al. 2014). Complementary to these biosynthetic enzymes, plants also have a suite of glycosidases that can trim xyloglucan side chains after deposition to the cell wall (Pauly et al. 2001; Sampedro et al. 2010). Such enzymes may be involved in turnover or recycling of sugars from non-structural xyloglucans in the cell wall. A flexible O-acetylation mechanism transfers acetyl groups to a number of polysaccharides and that is conserved across kingdoms (Gille and Pauly 2007). The targeting of acetylation is apparently well regulated but not understood in detail.

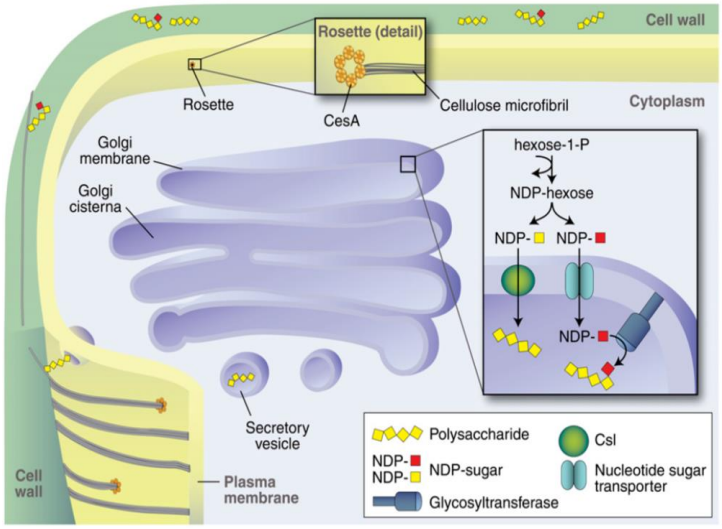


Fig. 3: Schematic representation of the key events in cell wall biosynthesis. Cellulose biosynthesis occurs at the plasma membrane in large complexes visualized as rosettes. The synthesis of matrix polysaccharides and glycoproteins occurs in the Golgi where the products accumulate in the lumen before transport to the cell wall via vesicles. The regulation of these biosynthetic events is an important issue that needs more study. Abbreviations used in the figure: CesA, cellulose synthase proteins that form the rosette; NDP-sugar, nucleotide sugars that act as donors for the sugars that go into polysaccharides; Csl, cellulose synthase-like proteins that are known to be involved in hemicellulose biosynthesis. (Harholt et al. 2010)

1.2.3.3 Xylan

Unlike other hemicelluloses there is no evidence that a CSL plays a role in the backbone formation of xylan. Instead, several glycosyltransferases from other GT families have been identified that are thought to be involved in xylan backbone elongation: GT43, GT47 and GT8, are thought to play a role in forming this oligosaccharide, but hitherto no defined enzyme activity has been demonstrated (Brown et al. 2009; Keppler and Showalter 2010; Wu et al. 2009, 2010; Chiniquy et al. 2013; Hörnblad et al. 2013; Peña et al. 2007; Persson et al. 2007;). GT8 has a unique localization in both ER and Golgi while the other GTs seem to be only present in the Golgi. For this reason, it was proposed that GT8 catalyses the initiation of the reducing end sequence by transferring xylose to an acceptor in an earlier stage of xylan synthesis. There is still a large gap in our knowledge of the mechanism of xylan backbone biosynthesis. In contrast, more progress has been made in recent years on the substitution of xylan: various glucuronosyl transferase enzymes lead to distinct differences in glucuronic acid (GlcA) substitution patterns of xylan (Bromley et al. 2013); xylosyltransferase is involved in adding xylosyl units (Chiniquy et al. 2012); a methyltransferase transfers specifically a methyl group from S-adenosyl-methionine to the O-4 position of glucuronosyl residues linked to xylan (Urbanowicz et al. 2012).

1.2.3.4 Mixed-linkage glucan

Two gene classes have been identified in mediating MLG synthesis, CSLF and CSLH both representing grass-specific branches of the CSL cellulose synthase-like gene family (Scheller and Ulvskov 2010). These experiments (Doblin et al. 2009; Vega-Sánchez et al. 2012; Burton et al. 2011) demonstrate that CSLF and CSLH are each independently sufficient for MLG biosynthesis, but as an in vitro mono-component assay has not been established it is not known what precise role those proteins have in MLG biosynthesis.

Amount of polysaccharide in wall (% w/w)[a]						
	Dicot walls		Grass walls		Conifer walls	
Polysaccharide	Primary	Secondary	Primary	Secondary	Primary	Secondary
Xyloglucan	20–25	Minor	2–5	Minor	10	–[b]
Glucuronoxylan	–	20–30	–	–	–	–
Glucuronoarabinoxylan	5	–	20–40	40–50	2	5–15
(Gluco)mannan	3–5	2–5	2	0–5	–	–
Galactoglucomannan	–	0–3	–	–	+[b]	10–30
β-(1→3,1→4)-glucan	Absent	Absent	2–15	Minor	Absent	Absent

Fig. 4: **Occurrence of hemicelluloses in primary and secondary walls of plants (**Scheller and Ulvskov 2010)

1.3 Functions

Cell walls perform a number of essential functions: providing shape to the many different cell types needed to form the tissues and organs of a plant; forming the interface between adjacent cells; playing important roles in intercellular communication (Keegstra 2010). The cell wall is a dynamic structure that often determines the outcome of the interactions between plants and pathogens. It is a barrier that pathogens need to breach to colonize the plant tissue (Bellincampi et al. 2015). Cell walls gives the cell a definite shape and structure, it provides structural support and it separates interior of the cell from the outer environment. But the cell walls aren't a rigid structure rather it is flexible: the original hypothesis of a macromolecular matrix made of covalently-linked domains of xyloglucan, pectins and structural proteins (Keegstra et al. 1973) was replaced by the simpler tethered network model (Hayashi 1989; Carpita and Gibeaut 1993; Cosgrove 2001) which highlighted direct coating and tethering of cellulose by xyloglucan as the key structural determinant of wall extensibility. Also it helps in osmotic-regulation, preventing water loss and rupture due to turgor pressure. It enables transport of substances and information from the cell insides to the exterior and vice versa. Aids in diffusion of gases in and out of the cell. Also provides mechanical protection from insects and pathogens. Especially HG and RGII, because of them mechanical properties, are well known to be involved in strengthening the wall. Plant pathogens cause degradation of pectin, and oligogalacturonides (i.e. α-1,4-linked oligomers of GalUA) are well established to be part of a signalling cascade that senses wall degradation upon pathogen attack (Ridley et al. 2001; D'Ovidio et al. 2004; Kohorn et al. 2009). The hypothesis is that RGI side chains plasticize cell walls that undergo large physical remodelling: the arabinans work as spatial regulators of the proximity of HG domains and this regulation may prevent the formation of Ca^{2+}-mediated interactions between HG regions. XyG metabolism and turnover is thought to play an important role in cell elongation

(Takeda et al. 2002; Pauly et al. 2001). XyG may have a role as a spacer-molecule by preventing the formation of microfibrilar cellulosic aggregates (Thompson 2005; Anderson et al. 2010) or an adapter molecule, which enables cellulose to interface with other cell wall matrix components (Keegstra et al. 1973; Talbott and Ray 1992; Ha et al. 1997; Cavalier et al. 2008).

2. Enzyme

2.1 Cell wall modified enzyme

Growing cell wall must be “loosened” in order to expand its surface arises from various biophysical, biochemical, and physiological considerations. Plant primary cell walls constitute a flexible and metabolically active extra protoplasmic compartment; they control cell expansion by varying their extensibility (Franková and Fry 2013). The primary wall confers the cell’s ability to define its own shape and size. Cell expansion is an irreversible increase in cell volume often exceeding 1000-fold. Controlled plant cell expansion demands the reversible ‘loosening’ of the cellulose–hemicellulose–pectin primary wall. Thus, plants will require a battery of wall-manipulating enzymes not found in other organisms (Labavitch 1981; Fry 1995, 2004; de la Torre et al. 2002; Minic 2008).

2.1.1 GH

Glycoside hydrolases are enzymes that catalyse the hydrolysis of the glycosidic linkage of glycosides, leading to the formation of a sugar hemiacetal or hemiketal and the corresponding free aglycon (www.cazypedia.org).

2.1.1.1 Definition

The glycoside hydrolases have been classified into more than 100 families (Henrissat 1991). Each family (GH family) contains proteins that are related by sequence, and by corollary, fold. This allows a number of useful predictions to be made since it has long been noted that the catalytic machinery and molecular mechanism is conserved for the vast majority of the glycosidase families as well as the geometry around the glycosidic bond (irrespective of naming conventions) (Henrissat et al. 1995). *exo*- and *endo*- refers to the ability of a glycoside hydrolase to cleave a substrate at the end (most frequently, but not always the non-reducing end) or within the middle of a chain (Davies and Henrissat 1995). Classification of GH families into larger groups, termed "clans", has been proposed (Henrissat and Bairoch 1996; Davies and Sinnott 2008). A clan is a group of families that possess significant similarity in their tertiary structure, catalytic residues and mechanism (Fig 5). Thus knowledge of three-dimensional structure and the functional assignment of catalytic residues is required for classification into clans. Families within clans are thought to have a common evolutionary ancestry.

GH-A (β/α)8
GH-B β-jelly roll
GH-C β-jelly roll
GH-D (β/α)8
GH-E 6-fold β-propeller
GH-F 5-fold β-propeller
GH-G (α/α)6
GH-H (β/α)8
GH-I α+β
GH-J 5-fold β-propeller
GH-K (β/α)8
GH-L (α/α)6
GH-M (α/α)6
GH-N β-helix

Fig. 5: Clan classification of GH

2.1.1.1.1 Reaction mechanisms

It can be defined "Inverting glycoside hydrolases" the hydrolysis of a glycoside with net inversion of anomeric configuration. It is generally achieved via a one step, single-displacement mechanism involving oxocarbenium ion-like transition states. The reaction typically occurs with general acid and general base assistance from two amino acid side chains, normally glutamic or aspartic acids, that are typically located 6-11 A apart (McCarter and Withers 1994). It can be defined "Retaining glycoside hydrolases" (Fig 6) the hydrolysis with net retention of configuration. It is most commonly achieved via a two-step, double-displacement mechanism involving a covalent glycosyl-enzyme intermediate. Each step passes through an oxocarbenium ion-like transition state. Reaction occurs with acid/base and nucleophilic assistance provided by two amino acid side chains, typically glutamate or aspartate, located 5.5 A apart. In the first step (often called the glycosylation step), one residue plays the role of a nucleophile, attacking the anomeric centre to displace the aglycon and form a glycosyl enzyme intermediate. At the same time the other residue functions as an acid catalyst and protonates the glycosidic oxygen as the bond cleaves. In the second step (known as the deglycosylation step), the glycosyl enzyme is hydrolysed by water, with the other residue now acting as a base catalyst deprotonating the water molecule as it attacks (McIntosh et al. 1996).

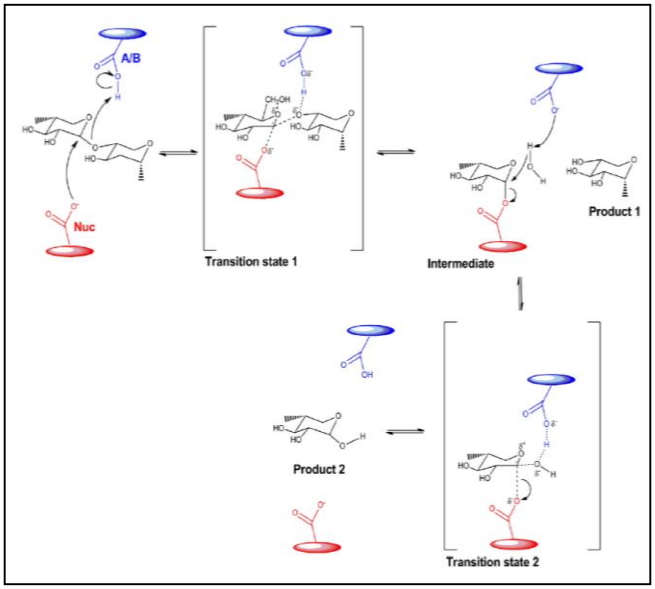


Fig. 6: Catalytic mechanism of GH11. Nuc is the nucleophile catalytic residue, A/B is the acid/base catalytic. residue.

2.1.1.2 Function

Cell wall polysaccharide biogenesis includes polymer synthesis, secretion, assembly, and rearrangement during development. The regulation of all of these processes requires precise genetical, hormonal, temporal and environmental control for proper completion of plant development. Synthesized polysaccharides are integrated into the cell wall immediately after they arrive in the apoplast. In

addition, during plant development and growth, the size and morphology of cells change, resulting in the degradation of cell wall components. GH enzymes located in the wall or in the plasma membrane play a crucial role in the degradation of different cell wall polysaccharides. These modifications allow changes in the structure and composition of polysaccharides.

2.1.2 XHT

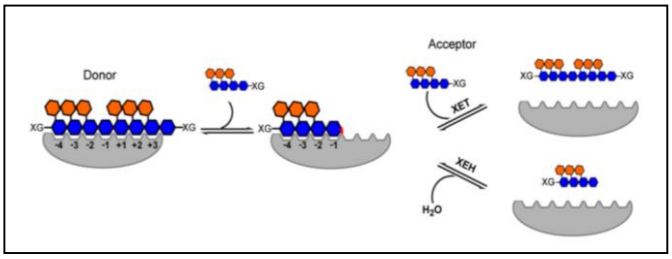


Fig. 7: A schematic representation of the mechanism used by XETs and XEHs. Xyloglucan binds to XETs and XEHs in both negative and positive subsites (glucosyl units of xyloglucan in blue and xylosyl units in orange. After binding, the substrate is cleaved, resulting in a covalent glycosyl-enzyme intermediate. In the last step, the glycosyl-enzyme is broken down by an incoming acceptor, either water (XEH) or the nonreducing end of a xyloglucan molecule (XET).

Polysaccharide-remodelling enzymes create new polysaccharide–polysaccharide linkages and thus play a role in recruiting newly secreted polysaccharides into the wall fabric, contributing to wall assembly (Franková and Fry 2013). XET: xyloglucan endo-transglucosylase activity, formally xyloglucan:xyloglucosyl transferase, results in the non-hydrolytic cleavage and ligation of xyloglucan chains. XEH: xyloglucan endohydrolase activity, formally xyloglucan specific endo-β-1,4-glucanase yields irreversible chain shortening (Rose et al. 2002). The proteins encoded by XTH genes comprise a subfamily of GH16 (Eklöf and Brumer 2010). GH16 enzymes display a diversity of substrate specificities, with family members cleaving β-1,3 or β-1,4 bonds in various glucans and galactans. Both the XET and XEH structures display the β-jellyroll fold common to all members of GH16, but with notable differences that reflect the specialization of these enzymes toward their highly branched substrate. Compared with the GH16 β-1,3;1,4-glucanases, which hydrolyse unsubstituted polyglucose chains (Planas 2000), the XETs and XEHs have a much wider substrate-binding cleft due to a major loop deletion in the negative subsides of the active-site cleft. The experimentally determined XET (Johansson et al. 2004) and XEH (Baumann et al. 2007) tertiary structures also reveal the structural importance of highly conserved Cys residues (Campbell and Braam 1999), which help to stabilize the C-terminal extension by the formation of two disulphide bonds. Although the inclusion of XETs in a GH family may seem incongruous, the ability of XETs to catalyse transglycosylation is a logical consequence of the canonical “retaining” catalytic mechanism employed by all members of GH16 (Planas 2000). A key feature of the retaining mechanism is the formation of a covalent glycosyl-enzyme intermediate, which can be broken down by water, yielding hydrolysis (XEH activity), or an incoming saccharide substrate, yielding transglycosylation (XET activity; Planas 2000; Gilbert et al. 2008) (Fig 7). In addition to possessing a wide active-site cleft, which is capable of accommodating brush-like xyloglucan chains, GH16 XETs and XEHs have a number of structural features that form specific interactions with the polysaccharide to bring it “in register” for cleavage by the catalytic amino acids. The cutting/pasting of polysaccharide chains by transglycanases occurs not only at the moment of secretion but also between

pairs of polysaccharides which have already been part of the wall architecture for some time. It is difficult to deduce whether such reactions contribute predominantly to wall loosening and thus growth promotion (Thompson and Fry 2001) or to wall strengthening by stitching polymers together (Nishikubo et al. 2007; Eklöf and Brumer 2010).

2.1.3 Expansins

Expansins were first identified as wall-loosening proteins in studies of 'acid-induced growth' (McQueen-Mason et al. 1992; Cosgrove 1989; Li et al. 1993). They are typically 250-275 amino acids long and are made up of two domains (domain 1 and domain 2) preceded by a signal peptide (Cosgrove 2000). A nonenzymatic mechanism has been proposed for expansin action: wall loosening proteins induce wall stress relaxation and irreversible wall extension in a pH-dependent manner, but they do not hydrolyse wall polymers. They induced long-term pH-dependent extension, and they enhanced stress relaxation of isolated walls over a broad time range, also in a pH-dependent manner (Cosgrove 1999). Expansin domain 1 has a distant homology to glycoside hydrolase family 45 (GH45) proteins, most of which are fungal β-1,4-D-endoglucanases: they form a six-stranded β-barrel with a groove for substrate binding. Furthermore, expansin domain 1 shares with GH45 a number of conserved cysteines that form disulphide bridges in the fungal enzymes. It is interesting that several residues that make up the catalytic site of GH45 endoglucanases are also conserved in expansin (Sampedro & Cosgrove 2005). Despite the presence of these conserved GH45 motifs, no hydrolytic activity has been detected for either α-expansin or β-expansin proteins.

2.1.3.1 α-expansins

Primary transcripts of α-expansins are predicted to encode a protein of ~28 kDa, which includes a secretory signal peptide that is removed to make a mature protein of ~25 kDa. α-expansins typically lack motifs for N-linked glycosylation. In terms of abundance, α-expansins are very minor components of the hypocotyl wall. Studies to date indicate that α-expansin lacks significant hydrolytic activity against the major polysaccharides of the wall (McQueen and Cosgrove 1995).

2.1.3.2 β-expansins

They share only 25% amino acid identity with α-expansins, but they appear to be homologous to them in structure. Curiously, grasses (but only grasses) also have an additional group of secreted proteins homologous only to expansin domain 2; these are known in the immunological literature as grass group-2 pollen allergens (G2As). They seem to have evolved from a truncated copy of a β-expansin gene and they share about 35-45% protein identity with their closest β-expansin relatives; their native biological function is uncertain. Although G2As evolved from a β-expansin ancestor, because of the loss of domain 1 they are considered a separate family and not part of the expansin superfamily.

2.1.4 GT

Glycosyltransferases are enzymes that catalyse the formation of the glycosidic linkage to form a glycoside. These enzymes utilize 'activated' sugar phosphates as glycosyl donors, and catalyse glycosyl group transfer to a nucleophilic group, usually an alcohol (www.cazypedia.org).

2.1.4.1 Definition

Glycosyltransferases can utilize a range of donor substrates. Sugar mono- or diphosphonucleotides are sometimes termed Leloir donors; the corresponding enzymes are termed Leloir glycosyltransferases. Glycosyltransferases that utilize non-nucleotide donors, which may be polyprenol pyrophosphates, polyprenol phosphates, sugar-1-phosphates, or sugar-1-pyrophosphates, are termed non-Leloir glycosyltransferases. In the last two cases, the enzymes that catalyse the transfer of a glycosyl group from a glycosyl phosphate or pyrophosphate are more commonly referred to as phosphorylases and pyrophosphorylases.

2.1.4.2 Reaction mechanisms

Glycosyltransferases catalyse the transfer of glycosyl groups to a nucleophilic acceptor with either retention or inversion of configuration at the anomeric centre. This allows the classification of glycosyltransferases as either retaining or inverting enzymes.

2.2 Cell wall degrading enzymes (CWDEs)

Anaerobic microorganisms have evolved a system to break down plant cell walls that involves the formation of a large extracellular enzyme complex called the cellulosome, which consists of a scaffolding protein and many bound cellulases (Doi & Kosugi 2004). It is now well established that phytopathogenic microorganisms secrete a number of hydrolytic enzymes capable of degrading cell-wall polymers to invade the plant tissue and feed on the released nutrients (Walton 1994). Nevertheless, so far only few cell wall–degrading enzymes have been reported to be important pathogenicity factors by demonstrating that their loss had significant impact on the pathogenicity of the microorganism. The majority of these reports refer to pectin-degrading enzymes such as polygalacturonases, pectin methylesterases, and pectate lyases (D'Ovidio et al. 2004). However, unlike in dicotyledonous plants, pectin is only a minor constituent of graminaceous cell walls, which consist mainly of hemicellulose (Carpita 1996). Since xylans represent a large proportion of the hemicellulosic fraction of cereal cell-wall matrices, xylan-degrading enzymes are expected to be important components of the offensive arsenal of cereal pathogens (Cooper et al. 1988; Wanjiru et al. 2002) and may have a role similar to that of pectic enzymes in infection of dicotyledons. The interaction between plants and pathogens induces a diverse battery of defence mechanisms, including cell wall strengthening, de novo production of antimicrobial compounds, ethylene biosynthesis, and rapid localized cell death (Greenberg 1997; Morel and Dangl 1997). These defence responses are triggered by pathogen-derived molecules that have been termed elicitors.

2.2.1 Clan A

The catalytic domain of enzymes from this clan has a (β/α)8-barrel fold, also called a TIM-barrel fold, which resembles the shape of a salad bowl.

2.2.1.1 GH2

At the primary-structure level, GH2 is now divided into four subclasses, each dedicated to the hydrolysis of a particular glycosidic link. These subfamilies encompass taxonomic barriers, as the same subfamily can group enzymes from bacteria as well as from mammals (Côté et al. 2006). In GH2 the catalytic functions are attributed to a nucleophile, which is a glutamate residue located close to the C-terminus of the seventh β-strand, and an acid/base residue, which is a glutamate residue (always preceded by an asparagine residue) located close to the C-terminus of the fourth β-strand. The putative glutamate nucleophiles seemed to be strictly conserved in all the analysed GH2 members, a doublet Ser-Asp instead of Asn-Glu was observed at the location of the putative acid/base residue (Côté et al. 2006).

2.2.1.2 GH5

GH5 is one of the largest families of glycoside hydrolases, containing more than 1800 entries in the CAZy database. In this family, enzymes with varying activities are classified, that is, mostly cellulases (EC 3.2.1.4), glucan 1,3-β-glucosidases (EC 3.2.1.58), and mannan endo-1,4-β-mannosidases (EC 3.2.1.78) (Pollet et al. 2010). GH5 xylanases are multidomain proteins and include a catalytic domain and a C-terminal carbohydrate-binding module (CBM). CBMs are thought to have one or more of the following functions: (i) increase enzyme concentration on the surface of the substrate, (ii) non-hydrolytic substrate disruption and (iii) surface/interfacial modifications (Arantes and Saddler 2010; Shoseyov et al. 2006). Their location within the CD can be both N-terminal and/or C-terminal, being connected to the CD by linker sequences of various length (from a few residues to several dozen), rich in serine or threonine (Gilbert and Hazlewood 1993). The catalytic domain has the (β/α)8-barrel fold typical for clan GH-A, while the smaller second domain—an unclassified putative xylan-binding domain—has a β9-barrel motif. Both domains are connected with two linker peptides and a large interface is formed between α-helices 7 and 8 of the catalytic domain and the β-sheet of the CBM. The active-site cleft with the catalytic residues, Glu165 and Glu253, is located along the C-terminal side of the β-barrel of the catalytic domain. Glu165 is the acid/base while Glu253 serves as the nucleophile during catalysis (Pollet et al. 2010).

2.2.1.3 GH10

Endoxylanases typically have a high molecular weight (≥30 kDa) and a low pI (Subramaniyan and Prema 2002). The overall structure of the catalytic domain of these enzymes is an eightfold β/α barrel resembling a ‘salad bowl,’ with the active

site located in a cleft running across the barrel top. The functional domain may be accompanied by a carbohydrate-binding domain, which is in most cases, however, specific for cellulose (Kulkarni et al. 1999). The active site of GH10 xylanases is located in a shallow groove along the top face of the molecule at the C-terminal side of the β-strands of the barrel. The two conserved catalytic glutamates involved in the retaining mechanism are located near the C-terminal ends of β-strands 4 and 7 (Jenkins et al. 1995). Enzymatic reaction proceeds via a double displacement mechanism (Henrissat and Davies 1997; Rye and Withers 2000). GH10 endoxylanases tend to release shorter oligosaccharides. Typical for GH10 xylanases is that they have a broad substrate specificity. They attack not only linear substrates but also accommodate decorated heteroxylans and xylo-oligosaccharides in their active-site clefts (Pollet et al. 2010). Analysis of hydrolysis products showed that GH10 xylanases can attack the glycosidic linkage next to a single- or double-substituted xylose toward the non-reducing end and require two unsubstituted xylose residues between branched residues (Pell et al. 2004).

2.2.2 Clan C

The GH clan-C consists of 4 major groups of sequences that correspond to the bacterial and fungal members in each of the two families. Despite low sequence identity between the two different subgroups (bacterial, fungal) within each of these GH families, and the virtual absence of identity between the two families in GH-C, the overall three-dimensional structures for all known GH-C protein structures are remarkably similar. Only three amino acids are completely conserved in clan GH-C. Two of the conserved amino acids correspond to the active-site nucleophile Glu 116 and the acid/base Glu 200 which are essential to catalysis. The third conserved residue, Val 160, is located on the inner concave β-sheet at the end of strand β7. Several of the GH-C enzymes contain one or two disulphide bridges, and it is likely that they are mainly needed for a local stabilization of their structures, and not for the overall protein fold, since they are localized at different positions in the different enzymes and are missing in several of the enzymes (Sandgren et al. 2005).

2.2.2.1 GH11

GH11 endoxylanases are generally characterized by a low molecular mass (typically around 22 kDa) and a high pI, although endoxylanases exhibiting acidic pI values also are naturally occurring in this family (Törrönen and Rouvinen 1997). Their structure adopts a β–jelly roll fold, which has been likened to a right hand with a two–β strand ‘thumb’ forming a lid over the active site located in the ‘palm’ (Törrönen et al. 1994). In consonance with GH10 endoxylanases, two glutamates are implicated in the enzymatic hydrolysis, with retention of anomeric configuration via a double-displacement mechanism. However, GH11 endoxylanases are usually more selective and release larger oligosaccharides, since substituents represent a more serious hindrance to their activity (Biely et al. 1997). The catalytic domain displays a β-jelly-roll architecture, composed of two antiparallel β-sheets named A and B sculpting a long and deep cleft. Many

hydrogen bonds between the β-strands stabilize the framework. The β-sheet A is nearly planar, whereas the β-sheet B is partially twisted on itself, forming a perpendicular angle. A unique α-helix is packed under the β-sheet B (Paës et al. 2012). The β-jelly-roll is one of the super-folds (Orengo et al. 1994), since its structure is better conserved than its sequence and it can hold several functions. GH11 catalytic residues are thus two glutamic acids, involved in a double-displacement catalytic mechanism and the existence of a covalent glycosyl-enzyme intermediate (McCarter and Withers 1994; Sinnott 1990). GH11 select the type of decoration present on the heteroxylan backbone, since they can only hydrolyse xylosidic bonds where the two corresponding xylose moieties in subsites (−1) and (+1) are not branched (Biely et al. 1997). GH11 display a high selectivity for heteroxylans, but their specificity is dependent on the type of heteroxylan. Indeed, heteroxylans which are xylose-based backbones polysaccharides from plant cell walls can be substituted in O2 and O3 by other sugars, typically arabinose, galactose, xylose, and by ferulic and glucuronic acids (Saulnier et al. 1995). Heteroxylans with high degree of substitutions are therefore not fully hydrolysed by GH11, contrarily to low substituted heteroxylans and of course xylan.

2.2.2.2 GH12

All GH 12 enzymes hydrolyse glycosidic bonds using the retaining mechanism (Sandgren et al. 2005). It leads to a net retention of the configuration at the anomeric carbon (C1) of the substrate after cleavage, via a double displacement mechanism. Two catalytic carboxylate residues are involved usually at opposite sides of the sugar plane approximately 5.5 A apart. The enzyme has a molecular weight of 25 kDa, has a neutral pI of 7.5, is only sparsely glycosylated, the two catalytic residues GH12 are the two carboxylates E116 and E200 (Okada et al. 2000). GH12 consists of 15 long β-strands that fold into two twisted, largely anti-parallel b-sheets, A and B, which pack on top of one another. The convex β-sheet A consists of six anti-parallel strands (A1–6), and the concave β-sheet B consists of nine largely anti-parallel strands (B1–9) (Fig 8). There is a single α-helix in the structure that packs against the outer convex surface of β-sheet B. The two cysteines Cys 4 and Cys 32, form a disulphide bond that bridges β-strands A1 and A2 (Sandgren et al. 2005). The N-terminal glutamine of GH12 (*H.*

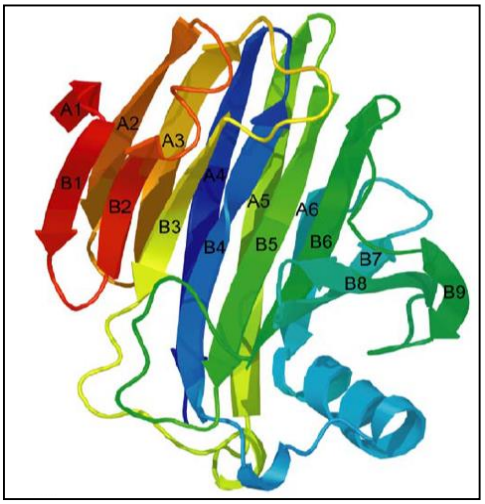


Fig. 8: Schematic ribbon diagram drawing top side view of *H. jecorina* Cel12A crystal structure, starting with red at N-terminus, and ending with blue at the C-terminus. The two β-sheet in the structure are labeled A and B. individal strands are labeled (A1-A6 or B1-B9) according to their position in the two β-sheet.

jecorina Cel12A) undergoes a cyclization and condensation reaction with the amine group of the N-terminus, to produce a cyclic pyro-glutamate. This is common in fungal extracellular enzymes and is thought to make the protein resistant to proteolytic degradation. Glucosyl binding sites frequently include the exposed surfaces of aromatic side-chains of Trp, Phe and Tyr residues. The carboxylate group oxygens of the two glutamates in *H. jecorina* Cel12A are separated by 5.4 Å, a distance frequently observed for the nucleophile/acid–base involved in a retaining mechanism (McCarter and Withers, 1994).

2.2.3 Not classified clan

2.2.3.1 GH74

β-1,4-xyloglucan hydrolase GH74 is an 842-residue protein that consists of a N-terminal catalytic module (residues 1–776) and a C-terminal dockerin module (residues 777– 842) (Martinez-Fleites et al. 2006). Enzymes of the GH74 family catalyse hydrolysis with inversion of the anomeric configuration of the product with respect to the β-linkage of the substrate (Lombard et al. 2014). The inverting mechanism is mediated by a catalytic base residue, which activates water for direct nucleophilic attack of a water molecule by deprotonation, and a catalytic acid residue, which protonates the leaving group facilitating its departure (Davies and Henrissat 1995; McCarter and Withers 1994). The catalytic residues, Asp70 (catalytic base) and Asp483 (catalytic acid), are located on opposite sides in the middle of this cleft, approximately 8 Å apart (Attia et al. 2016). The surface of this cleft is formed by the loops connecting the β-propeller blades in both domains. In the apo GH74 structure some of these loops are disordered, whereas in the ligand complexed forms they become ordered and participate in substrate binding. GH74 family groups enzymes that are able to hydrolyse xyloglucan oligosaccharides but also are active on non-branched substrates like barley β-glucans (β-1,3/1,4 glucan), carboxymethyl cellulose (CMC), Avicel (microcrystalline cellulose) or galactomannan (β-1,4-mannose).

2.2.4 Elicitor

If successful pathogens need wall-degrading enzymes, there would be a selective advantage to plants to recognize them as signals of incipient attack. Indeed, wall depolymerases from microorganisms are recognized by plants in several different ways (Walton 1994). The term "oligosaccharin" has been coined by Albersheim and co-workers to describe biologically active oligosaccharides that are produced as a result of the action of either endogenous or microbial enzymes on larger inactive polysaccharides. Plant cells could perceive xylanase directly by a receptor for this protein (Hanania and Avni 1997) or indirectly via plant cell wall fragments generated by its enzymatic activity (Bucheli et al. 1990). An endoxylanase from *Trichoderma viride*, referred to as the ethylene-inducing xylanase (EIX), has been used most extensively to study elicitor activity of endoxylanases. A protein purified from a commercial

preparation (Cellulysin) 3 of plant cell wall-digesting enzymes secreted by the saprophytic fungus *Trichoderma viride* is a potent inducer of ethylene biosynthesis in tobacco (*Nicotiana tabacum cv Xanthi*) leaf discs. The protein was identified as an endoxylanase whose secretion by *T. viride* can be induced in culture by growth on D-xylose (Fuchs et al. 1989). The apparent size of native EIX may prove very important to the mechanism whereby the enzyme induces ethylene biosynthesis in plant tissues. A 9.2 kDa EIX protein would be capable of penetrating through even the smallest plant cell wall pores to interact directly with the plasmalemma (Dean & Anderson 1991). Some wall polysaccharides undergo enzymic turnover to release biologically active oligosaccharides ('oligosaccharins') with putative signalling roles. In addition to enzymes generating oligosaccharins from polysaccharides, plants also possess enzymes that degrade oligosaccharins, either by hydrolysis or by grafting large polysaccharides to them (Baydoun and Fry 1989; Darvill et al.,1992; García-Romera and Fry 1995). This may be important since biological 'messages' need to be inactivated when the information that they carry is no longer relevant to the plant's environmental or developmental situation. Oligosaccharides released from plant cell walls by fungal hydrolases are capable of inducing ethylene biosynthesis in plants (Tong et al. 1986; Baldwin et al. 1988) and, in some cases, may cause rapid death of plant cells (Bucheli et al. 1990). The enzymatic activity of fungal xylanase is not necessary for its elicitor activity (Enkerli et al 1999). Enzymatic activity is not required for cell death activity (Ma et al. 2015). The mutations drastically reduced the enzyme activity, they did not affect elicitor activity in a comparable manner. The mutated forms of xylanase were still active as elicitors of extracellular alkalinization and ethylene biosynthesis of tomato cell cultures, and of necrosis in tomato and tobacco leaves (Enkerli et al 1999). The elicitor activity is solely based on the specific recognition of the xylanase protein. Oligosaccharides derived from fungal and plant cell wall polysaccharides are one class of well characterized elicitors that, in some cases, can induce defence responses at a very low concentration e.g. nM (Shibuya and Minami 2001). A partial hydrolysate of the *Phytophthora sojae* β-glucan also acts as an active elicitor on various plant cells belonging to the family Fabaceae, indicating the presence of similar perception systems in these plants. Glycoside hydrolase family 12 (GH12) protein produced by the soybean pathogen *Phytophthora sojae* that exhibits xyloglucanase and b-glucanase activity. It acts as an important virulence factor during *P. sojae* infection but also acts as a pathogen-associated molecular pattern (PAMP) in soybean (*Glycine max*) and Solanaceous species, where it can trigger defence responses including cell death. GH12 proteins occur widely across microbial taxa, and many of these GH12 proteins induce cell death in *Nicotiana benthamiana* (Ma et al. 2015).

3. Inhibitor

In response to pathogenic attack, plants produce glycoside hydrolase inhibitor proteins (GHIPs) against the cell wall-degrading enzymes. In response to attack by the multiplicity of xylanases secreted by pathogenic species, plants probably have evolved different GHIPs proteins to counteract the intruding microorganisms. In the course of evolution, antagonistic interactions between pathogens and their hosts could have resulted in an ongoing evolution of glycoside hydrolase that evade inhibition and inhibitors which adapt to these newly produced enzymes. This molecular struggle can lead to a 'diversifying' selection causing variation of residues at the interaction interface between the glycoside hydrolase and the inhibitor. Even single residue substitutions can alter the outcome of the glycoside hydrolase–inhibitor interaction (Dornez et al. 2010).

3.1 Dicots

Xyloglucan-specific endo-β-1,4-glucanase inhibitor protein (XEGIP) inhibits the hydrolytic activity of a xyloglucan-specific β-1,4-endoglucanase isolated from *Aspergillus aculeatus* (XEG from GH12 family) (Pauly et al. 1999) that was detected in the culture medium of suspension-cultured tomato cells. A protein whose presence is correlated with the inhibitory activity was purified from the ethanol-precipitated material by ion-exchange and size-exclusion chromatography (Qin et al. 2003). The molecular weight of the purified protein was approximately 50 kDa. XEGIP inhibits XEG by binding to XEG to form a 1:1 complex and that XEGIP does not proteolytically degrade XEG. Xyloglucan-specific endoglucanase inhibitor protein did not show any capacity to inhibit other plant cell-wall-degrading enzymes tested (Cook et al. 1999). The deduced protein sequence has a putative signal peptide (22 aa); the predicted mature XEGIP protein has 415 amino acids and a molecular weight of 44229.74 Da. The mature protein also has five potential N-glycosylation sites, which are likely to account for some or all of the 6625 Da difference between the predicted molecular weight and the molecular weight measured by MALDI-MS (50 853 Da). XEGIP is present as a single copy in tomato. Northern analysis revealed that XEGIP mRNA was expressed in all the vegetative tissues examined, with lower expression levels in young healthy leaves. XEGIP mRNA abundance increased during fruit expansion, peaked immediately prior to the onset of ripening at the mature green stage, and declined as ripening progressed (Qin et al. 2003). Proteins homologous have been identified in various plants, and several of these proteins have been characterized (Xyloglucan-

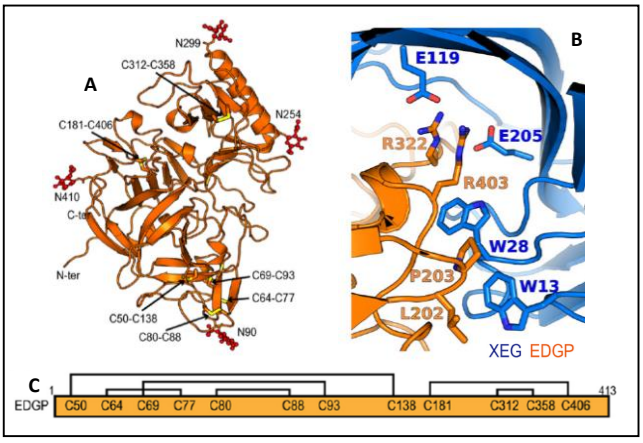


Fig. 9: Crystal structure of EDGP. A) Overall structure of EDGP is represented by a ribbon model. The disulfide bonds are shown by stick models, in which sulfur atoms are colored yellow. N-Linked glycans are represented by ball and stick models and colored red. **B)** Schematic drawing shows disulfide bonds in EDGP. **C)** Detailed interactions between R322 and R403 from EDGP (orange) and E119 and E205 from XEG (light blue).

specific endo-β-1,4-glucanase inhibitor proteins like (XEGIPs-like)). Extracellular dermal glycoprotein (EDGP) (Fig 9) from carrot was purified from carrot callus culture medium (Yoshizawa et al. 2011). EDGP adopts a pepsin-like fold that is β-rich with several α-helices and is roughly divided by a centre cleft comprising the active site (Browner et al. 1995). Despite the structural similarity to pepsin, one of the catalytic aspartates in pepsin is replaced by Ser271 in EDGP, and thus EDGP lacks protease activity. Consistent with this, other GHIPs also lack the catalytic aspartate. EDGP has six disulphide bonds, and these supposedly stabilize the tertiary structure of EDGP in the extracellular environment. EDGP has four putative N-linked glycosylation sites: Asn90, Asn254, Asn299, and Asn410 (Shang et al. 2004). In each putative N-linked glycosylation site, the electron density map indicated at least one sugar moiety linked to the asparagine (Yoshizawa et al. 2012). EDGP completely covers the active cleft of xyloglucan-specific endo-β-1,4-glucanase from *Aspergillus aculeatus* (XEG from GH12 family) like showed in the complex structure; no substantial conformational changes occur in EDGP upon the binding of XEG. In the XEG-EDGP structure, Arg322 and Arg403 of EDGP insert into the active cleft of XEG and form an electrostatic interaction with the catalytic residues, Glu119 and Glu205 (Fig 9). Hydrophobic interactions are made between the aliphatic moiety of Arg403 of EDGP and Trp28 of XEG, Leu202 and Pro203 of EDGP and Trp13 of XEG. EDGP mimics the interaction between XEG and xyloglucan observed in the structure of the XEG xyloglucan complex. Interestingly, EDGP does not interact with Tyr24 of XEG, which is involved in xylose recognition (Yoshizawa et al. 2012). The two arginines in inhibition loop 1 (IL-1) and in inhibition loop 2 (IL-2) are conserved in most GHIPs, including homologous proteins from tomato, tobacco, potato, and *Arabidopsis thaliana*. Of these, tomato and tobacco GHIPs inhibit GH12 enzymes. Furthermore, Leu202 and Pro203 (which contact the -4 subside of the enzyme active site) are also conserved in GHIPs of most plants. NEC4 protein functions as a xyloglucan-specific fungal endoglucanase-inhibitor protein (XEGIP). In ornamental tobacco (*Nicotiana langsdorffii x Nicotiana sanderae*), this defensive function is mediated by a series of nectar-expressed proteins termed nectarins. Nectar proteins are present in nectar to protect the gynoecium and developing embryos therein from microbial attack caused by nonsterile visiting pollinators. This observation is based on the identification of a novel nectar-contained biochemical pathway, termed the Carter-Thornburg nectar redox cycle (Carter and Thornburg 2004). The mature protein sequence contains six potential sites for N-linked glycosylation. These are Asn-24, Asn-114, Asn-152, Asn-278, Asn-321, and Asn-432. The difference between the observed molecular mass of the NEC4 protein (60 kDa) and that predicted from the amino acid sequence (44.6 kDa) suggests that at least some of these potential sites must be glycosylated. Four of these N-glycosylation sites are present in the NEC4 model, and each of them is surface exposed as would be expected if they are glycosylated. Purified NEC4, incubated with a xyloglucan-specific endoglucanase (XEG) from *Aspergillus aculeatus* showed strong inhibition of the XEG (Naqvi et al. 2005). The calculated inhibition constant for the interaction of XEG and NEC4 was not significantly different from that determined by Qin et al. (2003) for the interaction of XEG with XEGIP from tomato cell cultures. Thus, NEC4 has potent XEG inhibition activity, confirming that this protein is an XEGIPs-like. The formation of an NEC4:XEG complex modulates key reactions in the Nectar Redox Cycle, which maintains the antimicrobial environment of

nectar (Harper et al. 2010). Experiments initially designed to evaluate the NEC4:XEG interaction in raw nectar showed that adding XEG to nectar increases its GOX activity, as determined by measuring the accumulation of gluconic acid and H_2O_2. GOX activity of NEC5 is enhanced upon addition of the fungal enzyme XEG. Although the interaction of XEG with NEC4 is well established (Qin et al. 2003; Naqvi et al. 2005), direct interaction of XEG with NEC5 has never been observed. Activation of NEC5 activity by the NEC4:XEG complex plays a key role in the induction of toxic reactive oxygen species (ROS) and their targeting to the fungal surface in floral nectar. The resulting increase in ROS also has the potential to initiate signal transduction pathways that control other plant defence responses. Homologous protein was purified also from potato (XGIP), the observed molecular weight of 55 kDa is greater than the predicted 48 kDa, due to glycosylation (Jones & Perez 2014). XGIP overexpression was capable of protecting potato from disease caused by endo-β-1,4-glucanase GH12 from *Phytophthora infestans* (EGL12s). In a possible case of host-pathogen co-evolution, it has been reported that Solanaceous plants harbour multiple copies of XEGIPs-like and they may have evolved to address the variations in pathogen EGL12s (Jones 2012). In the case of *P. infestans*, it was reported that transient silencing of XGIP in potato by agro-infiltration led to an increase in tissue susceptibility (Jones et al. 2006). The expression new apple defence-related gene, MdXEGIP1, was significantly enhanced in response to infection of *Botryosphaeria dothidea*. As revealed by further investigation, recombinant MdXEGIP1 exhibited inhibitory effect to XEG secreted by *B. dothidea* (Bai et al. 2015). MdXEGIP1 had two arginines required for XEGIPs-like to inhibit the activity of XEG (Yoshizawa et al. 2012). MdXEGIP2 had only one of the two required arginine. The different structure of apple XEGIPs-like suggested that these proteins might have different inhibition activity to XEGs. Although MdXEGIP2 share higher sequence identity with MdXEGIP1, MdXEGIP2 couldn't be induced by *B. dothidea* infection like MdXEGIP1. The similar phenomenon was also observed during interaction of *Nicotiana benthamiana* with bacterial pathogen, in which NbXEGIP1 expression was induced by bacterium *Pseudomonas syringae pv. tabaci*, but the same induction was not observed in NbXEGIP2 (Xie et al. 2008) pathogen. The recombinant MdXEGIP1 exhibited strongly inhibitory activity to XEG from GH12, GH74 families and especially to XEG extract of *B. dothidea*, which indicated that MdXEGIP1 might contribute to the resistance of apple plant to ring rot disease. Distinct roles for the xyloglucan-specific endo-β-1,4-glucanase inhibitor protein from *Humulus lupus* (HlXEGIP) homologues in defence mechanisms. The biochemical and enzymatic properties of the inhibitors were determined, and the activity of HlXEGIP homologues against fungal xyloglucan-degrading endoglucanase was found to be different from: The HlXEGIP homologues genes contain no intron and encode proteins of 443 (HlXEGIP1), 445 (HlXEGIP2), and 441 (HlXEGIP3) amino acids. The HlXEGIP homologues genes were differentially regulated by wounding or biotic stress. HlXEGIP1 was produced during abiotic stress; in contrast, HlXEGIP2 appears to be produced in response to biotic stress. Various xylanase inhibitor isoforms also showed differential expression patterns, depending on the type of the pathogen and/or infected tissue (Habrylo et al, 2012). Shang et al. (2005) postulate that glycosylation is essential for correct carrot EDGP folding and secretion. In fact, all XEGIPs-like describe present several potential N-glycosylation sites, that explain the difference between the predicted molecular weight and the molecular

weight measured. Nevertheless, some xylanase inhibitors were produced and were shown to be active after production in bacterial systems (Fierens et al., 2003; Takahashi-Ando et al.,2007; Weng et al.,2010), proving the unnecessary of glycosylation for the inhibition activity. A xyloglucan-specific endo-b-1,4-glucanase inhibitor protein, CaXEGIP1, was isolated and functionally characterized in pepper (*Capsicum annuum*) plants; it is a protein of 430 amino acids with a predicted molecular mass of 47.2 kDa and a pI of 9.15 (Choi et al. 2013). Recombinant CaXEGIP1 inhibits XEG enzyme from *Clostridium thermocellum* F7/YS, belonging to GH74 family (Martinez-Fleites et al., 2006). The GH12 family XEG enzymes are structurally and mechanistically distinct from GH74 family XEG enzymes (Master et al., 2008; See Introduction 2.2.2.2 and 2.2.2.3). In addiction the gene used to produce a recombinant-CaXEGIP1, encode for a protein (Uniprot: K4FKJ9) that does not have the two conserved Arg residues from IL-1 and IL-2. This work doesn't analyse the other homologous protein (Uniprot: B9VUU9) that present the two conserved arginine (See Appendix sequence and alignment). The CaXEGIP1 overexpression induced a spontaneous cell death response and also increased the expression of some defence-related proteins when it was strongly induced in pepper leaves inoculated with the *Xanthomonas campestris pv vesicatoria* avirulent (incompatible) strain Bv5-4a harbouring avrBsT; with the same mechanism, if CaXEGIP1 was overexpressed by *Agrobacterium tumefaciens* it triggered pathogen-independent, spontaneous cell death in pepper and *Nicotiana benthamiana* leaves; also overexpression of CaXEGIP1 in *Arabidopsis thaliana* enhanced resistance to *Hyaloperonospora arabidopsidis* infection (Choi et al. 2013).

3.2 Cereals

Cell wall-degrading enzymes (CWDEs) are encoded by multigene families, making it difficult to completely eliminate their activity. Cereal cell walls only contain low levels of pectins (Carpita 1996), instead polygalacturonase inhibitor (PGIPs) are especially important in dicotyledons and non graminaceous monocotyledons, having pectin-rich cell walls (De Lorenzo et al., 2001). For these reasons GHIPs in cereal grains do not only occur in high levels, but also in multiple forms (Dornez et al. 2010). Three types of GHIPs occur in a fairly coordinated fashion throughout grain development and germination: *Triticum aestivum* L. endoxylanase inhibitors (TAXIs-like), xylanase inhibitor proteins (XIPs-like), thaumatin-like xylanase inhibitors (TLXIs-like) (Fig 10). TAXIs-like and TLXIs-like are already present in low levels at the water ripe stage [5 days post-anthesis (DPA)], while XIPs-like only appear at the milky stage (12 DPA) (Croes et al. 2009). The levels of TAXIs-like, XIPs-like and TLXIs-like drastically increase between the milky and the soft dough stage (19 DPA), and from then on, they remain high until grain maturity (42 DPA) (Croes et al. 2009). The levels of TAXIs-, XIPs- and TLXIs-like in the germinating grain kernels initially increase, but then decline systematically to typically 31%, 77% and 12% [23 days post-imbibition (DPI)], respectively, of their original level present in ungerminated wheat grains (Croes et al., 2009). The accumulation of GHIPs during the early stages of germination is consistent with the phenomenon of germination-based resistance, which has been reported previously for PR-proteins The three types of GHIPs in cereals are similarly distributed throughout the wheat kernel and they are highly abundant in the envelope of the wheat caryopsis, with the

exception of the outer pericarp. In particular, the highest concentrations of GHIPs occur in the aleuronic layer (Croes et al. 2009). Taking into account the yields of the different kernel

	TAXI	XIP	TLXI
Molecular mass	Form A: 40 kDa Form B: 30 + 10 kDa	30 kDa	18 kDa
pI value	>8.0	>8.7	>9.3
Glycosylation	Yes (limited N-glycosylation at Asn105)	Yes (two N-glycosylation sites Asn89 and Asn265)	Yes (mostly O-glycosylation)
Optimal pH	4.8–5.0	4.5–6.5	4.5–5.5
Optimal T	20–40°C	30°C	30–40°C
pH stability	3.0–12.0	3.5–n.d.	1.0–12.0
Half life (100°C)	3.5–4.5 min	<30 s	> 2 h
Specificity	GH11 Fungal + bacterial	GH10 and GH11 Fungal	GH11 Fungal + bacterial
Cross-inhibition	None	Barley α-amylases (GH13)	None
Mechanism	Competitive	Competitive	Non-competitive
Kinetics	Fast	Fast	Slow
K_i value#	2.2–20 nM	3.4–610 nM	61.5 nM
$[I]/[E]_{50}$ value$	0.6–1.1	0.6–63	1.6
Levels in common wheat cultivars	70–200 ppm 17–138 ppm 81–190 ppm	210–560 ppm 193–355 ppm 156–371 ppm	51–150 ppm
Different genetic forms in common wheat	*Taxi-Ia* *Taxi-IIa* *Taxi-Ib* *Taxi-IIb* *Taxi-III** *Taxi-IV** *Taxi-725ACCN* *Taxi-725ACC* *Taxi-725OS* *Taxi-602OS* *Taxi-801OS* *Taxi-801NEW*	*Xip-I* *Xip-III* 12 *Xip-R* genes (*Xip-R1* and *Xip-R2*)	*Tlxi*

Fig. 10: Three types of GHIPs in *Triticum aestivum*

layers, the starchy endosperm contains about half of the total GHIPs population (Croes et al. 2009). The coding sequences of all three types of GHIPs are preceded by a signal sequence, which presumably directs the proteins outside the plant cell wall. The apoplastic localization of GHIPs may be favourable for their action as inhibitors of microbial xylanases intruding the host plant. Investigation of the gene upstream region revealed the presence of consensus sequences of cis-acting elements implicated in pathogen- and wound-inducible gene expression, i.e., GCC-box and W-box sequences, the core sequence of activation sequence-1 and a consensus sequence of type I Myb-binding sites (Igawa et al., 2004).

3.2.1 TAXIs-like

Triticum aestivum L. endoxylanase inhibitors are widely represented in cereals. In general, they are characterized by molecular masses (MMs) of about 40 kDa and alkaline pI values (> 8.0) (Fig 10) (Goesaert et al. 2004). TAXI exist as two molecular forms. The first consists of a single polypeptide chain. The second is derived from the former and is made up of two polypeptides of approximately 30 and 10 kDa, held together by a disulphide bridge (Debyser et al. 1998; Debyser et al. 1999). TAXI proteins

are only slightly glycosylated (Sansen et al. 2004; Croes et al. 2008). Levels of TAXI in common wheat are high and reported values in different wheat cultivars vary from 17 to 200 ppm. TAXIs-like fold in two β-barrel domains with a few helical segments (Fig 11), divided by an extended cleft. A close structural relationship with pepsin-like aspartic peptidases has been revealed, but sequence similarities are low and TAXIs-like show no proteolytic activity (Sansen et al. 2004). TAXIs-like specifically inhibit bacterial and fungal xylanases belonging to glycoside hydrolase family GH11. The active site of GH family 11 xylanases contains two conserved Glu residues located on either side of the extended open cleft. Histidine is a key residue for inhibition as insertion of its imidazole ring between the two catalytic glutamate residues of the xylanase active site (Fig. 11) (Sansen et al. 2004; Fierens et al. 2005; Raedschelders et al. 2005; Pollet et al. 2009). Mutation of His374 TAXI−IA into alanine, lysine or glutamine confirmed that histidine strongly affects the affinity of TAXI-IA for the xylanases of *Aspergillus niger*, *Bacillus subtilis* and *Hypocrea jecorina* (Fierens et al. 2005). The *Aspergillus niger* xylanase (ANXI)-TAXI-I structure reveals a direct interaction of the inhibitor with the active site region of the enzyme and further substrate-mimicking contacts with binding subsides filling the whole substrate-docking region. TAXI-I displays significant sequence similarity with TAXI like proteins identified in barley (HVXI) (Goesaert et al. 2001), rye (SCXI-I to SCXI-IV) (Goesaert et al. 2002), and durum wheat (TDXI-I and TDXI-II) (Fig 12a and 12b) (Goesaert et al. 2003).

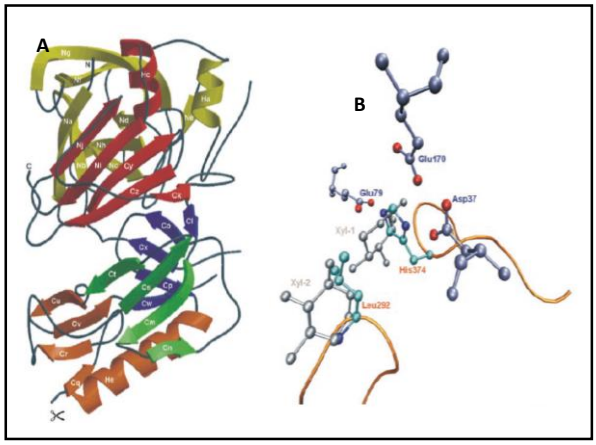


Fig. 11: Crystal structure of TAXI-I: A) Overall structure of TAXI-I. A is represented by a ribbon model. The composition of the β-sheets B, N1, N2, N3, C1, C2, and C3 is indicated. The position of the cleavage site is marked. **B)** Detailed view on the key interactions of the inhibition. The imidazole ring of His374 of TAXI-I (orange) is located in between the two catalytic residues (Glu79 and Glu170) of ANXI and is strongly hydrogen bonded to Asp37.

3.2.2 XIPs-like

Xylanase inhibitor proteins occur as monomeric proteins with MMs of approximately 30 kDa and pI values varying between 5.5 (rice) and > 9.3 (maize) (Fig 10) (Goesaert 2002; Goesaert et al. 2003; Goesaert et al. 2005). XIPs-like possess two N-glycosylation sites and are approximately 2% (w/w) glycosylated (Payan et al. 2003). Levels of XIP-type inhibitors in wheat are 2 to 3 times higher than those of TAXIs-like inhibitors (Bonnin et al. 2005; Dornez et al. 2006; Croes et al. 2009) and range from 156 to 560 ppm in different common wheat cultivars. XIPs-like possess a (β/α)8 barrel fold, the top of which is decorated by loops arranged to form a long depression running along one side of the molecule (Payan et al. 2003). They have structural features and N-terminal amino acid sequences typical for GH18 chitinases. XIPs-like generally inhibit fungal but

not bacterial GH10 and GH11 xylanases (Flatman et al. 2002). They bind GH10 and GH11 xylanases at distinct locations, possibly allowing simultaneous binding of the inhibitor to both target enzymes. For both GHs, substrate mimicry in the active site is the key element of the competitive inhibition mechanism and Lys234 XIP−I and Arg149 XIP−I are the key determinants for inhibition of GH10 and GH11 xylanases, respectively (Payan et al. 2004).

cereal	fraction	total protein (mg)	total activity (InU)[b]	specific activity (InU/mg)[b]
rye	crude extract	6100	71 500	12
durum wheat	crude extract	3300	40 000	12
barley	crude extract	2070	6200	3

Fig. 12a: Purification of TAXI-type endoxylanase inhibitor from Cereals whole meals

	inhibition activity against						
cereal	XAN[a]	XBS[a]	XTL2[a]	XTL3[a]	XTV[a]	XAA[b]	XANid[b]
durum wheat	+	++	+	++	++	0	0
	++	++	++	++	++	0	0
barley	++	++	++	++	++	0	0
rye	++	++	++	++	++	0	0
	++	++	++	++	++	0	0

Fig. 12b: Endoxylanase inhibitor activities against:
XAN, *Aspergillus niger* endoxylanase GH11;
XBS, *Bacillus subtilis* endoxylanase XynA GH11;
XTL2, *Trichoderma longibrachiatum* endoxylanase GH11;
XTL3, *Trichoderma longibrachiatum* endoxylanase GH11;
XTV, *Trichoderma viride* endoxylanase GH11;
XAA, *Aspergillus aculeatus* endoxylanase GH10;
XANid, *Aspergillus nidulans* endoxylanase GH10.

3.2.3 TLXIs-like

Thaumatin-like xylanase inhibitors have a MM of approximately 18 kDa and a pI value of at least 9.3 (Fig 10) (Fierens et al. 2007). In SDS-PAGE, native TLXI appears as a broad band of approximately 18 kDa, consisting of different smaller bands, which are due to the occurrence of varying degrees of glycosylation (Fierens et al. 2007). TLXI contains five disulphide bridges, resulting in a very stable protein that withstands extreme pH and temperature conditions (Fierens et al. 2009). Levels up to 6 times lower than TAXIs-like and XIPs-like (Fierens et al. 2007), Croes and co-workers (2009) reported TLXI levels in different wheat cultivars from 51 to 150 ppm. The 3D crystal structure of TLXI shows that it consists of two domains. The first domain, domain I, consists of a β-sandwich built up of two β-sheets comprising five (A1 to A5) and six (B1 to B6) β-strands. Apart from the N-terminal (A1) and the C-terminal (A5) strand, all these β-strands run antiparallel within each β-sheet. The second domain is formed by a β-hairpin turn followed by an extended loop and corresponds to domain III of thaumatin (Vandermarliere et al. 2010). The α-helices of domain II, present in thaumatin, are absent in TLXI (Vandermarliere et al. 2010). TLXI, like other short-chain thaumatin-like proteins (TLPs) with a MMs between 15 and 19 kDa, lacks about 50 amino acids, corresponding to domain II of thaumatin (Fierens et al. 2009). TLXIs-like specifically inhibit bacterial and fungal xylanases belonging to GH11 and do not inhibit GH10 xylanases. TLXIs-like inhibit xylanases in a non-competitive manner, i.e. by binding outside the active site. This is in sharp contrast to TAXIs-like and XIPs-like for which competitive inhibition kinetics have been demonstrated (Payan et al. 2004; Sansen et al. 2004). His22 TLXI, located at a flexible loop, is a key residue for inhibition activity (Rombouts et al. 2009).

3.2.4 Legumes

Two genes encoding γ-conglutin (Cγ) have been identified in *Lupinus albus*, but only one seems to be quantitatively expressed in the developing seed (Foley et al. 2011; Scarafoni et al. 2001). Also in *Lupinus angustifolius* two genes encoding γ-conglutin (LaCγ) are present and the only γ-conglutin (LaCγ) expressed show about 88% homologies with γ-conglutin (Cγ) from *Lupinus albus.* γ-conglutin is an unusual basic 7S protein, which is equally soluble in water and salt solutions. This protein accounts for about 4-5% of total proteins in mature lupin seeds (Duranti et al. 1981). Lupin γ-conglutin is located in the protein bodies of developing lupin seeds, however, the protein has also been detected in the extracellular apoplastic regions of germinating lupin cotyledons (Duranti et al. 1994). The unusual stability of γ-conglutin during seed germination and in vitro tests with various proteolytic enzymes (Duranti at al. 1995) strongly suggests that it is not a storage protein. The unusual extra-vacuolar location of γ-conglutin adds further evidence that this protein probably plays a non-storage role. γ-conglutin displays further unique properties of its own: it binds divalent metal ions, especially Zn^{2+} and Ni^{2+} (Duranti et al. 2001). Cγ is a hexameric protein (Czubinski at al. 2015), each monomer consisting of two disulphide-linked polypeptides of 30 and 17 kDa (Restani et al. 1981). These two polypeptide chains originate from a single precursor synthesized during seed development and processed by post-translational proteolysis. The γ-conglutin promoter is matured by the formation of six disulphide bridges, one of which links the α and β subunits, and the large subunits are N-glycosylated at a single site (Duranti et al. 1981; Duranti et al. 2008). In a reversible process, the protein forms a pH-

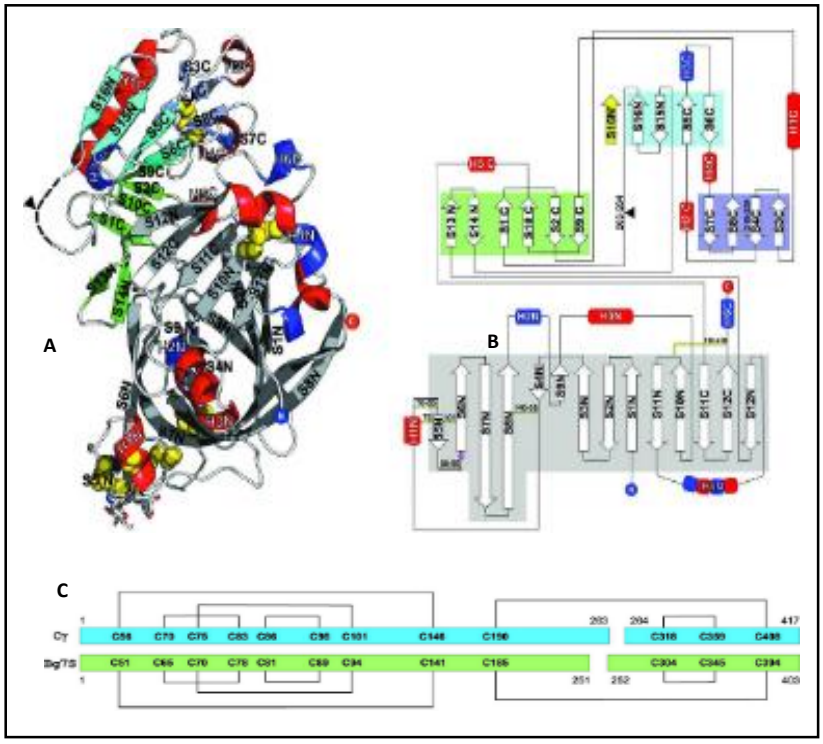


Fig. 13: Crystal structure of γ-conglutin: A) Overall structure of γ-conglutin from *Lupinus angustifolius,* the secondary-structure elements are marked and coloured according to topology. **B)** Topology diagram of one molecule of the of γ-conglutin hexamer: red/blue cylinders represent α-helices and arrows represent β-strands, that are organized into β-sheets. The post-translational cleavage site is marked by an arrowhead and the disulphide bridge in yellow. **C)** Topology of disulphide bonds of γ-conglutin and Bg7S, the post-translational cleavage sites are shown as gaps.

dependent association–dissociation equilibrium between the monomer and an oligomeric assembly (Capraro et al. 2010). The oligomeric form changes into the monomeric form with a dimeric transition state when the pH shifts from neutral to slightly acidic. The core of the γ-conglutin fold consists of four β-sheets flanked on the outside by α-helices. The huge 14-stranded antiparallel β-sheet forms the spine of the γ-conglutin molecule and dominates the N-terminal domain (Czubinski at al. 2015). The N-terminal domain has a rich pattern of (four) disulphide bridges. The glycosylation site is also present in this domain and is rigidified by three of these disulphide bonds. The C-terminal domain consists of three β-sheets, two of which are composed of four antiparallel strands each, while the third is composed of six mixed strands. The longest α-helix of the γ-conglutin fold is found within the C-terminal domain. The two domains are covalently linked by the interdomain Cys190–Cys408 disulphide bridge (Fig 13) (Czubinski at al. 2015). Putative *Lupinus albus* conglutin γ-like proteins (LACGs-like) are present in almost all leguminous seeds, even if they aren't largely expressed or accumulated. Basic 7S globulin from *Glycine max* (Bg7S) adopts a β-rich structure with several α-helices. Bg7S is post-translationally cleaved between Ser251 and Ser252, resulting in the α-chain and β-chain. Although these chains are intricately folded, the structure of Bg7S is roughly divided into the α-domain and β-domain. Bg7S has 12 cysteines in positions similar to those found in the primary structures of other XEGIPs-like and TAXIs-like, and these residues form six disulphide bonds (Fig 13). The disulphide bonds supposedly stabilize the three-dimensional structure of Bg7S. The Cys209–Cys418 bond seems to be significant for stabilization in particular, because it links the α-chain and β-chain (Yoshizawa et al 2011). Bg7S adopt a pepsin fold. The pseudo-active site of Bg7S corresponding to pepsin is located in the cleft between the α-domain and β-domain. However, Bg7S lacks protease activity, because one aspartate corresponding to the catalytic residue of pepsin is replaced by Ser265, like happened in TAXIs-like (Yoshizawa et al 2011). Bg7S exists as a tetramer with a cruciform shape formed by two superposed type I dimers, while γ-conglutin is arranged in a circular hexameric form composed of two three-membered rings. In the Bg7S tetramer, a different `type II' interface (A–B and C–D) is formed between loop elements of two molecules in a head-to-head arrangement. Interestingly, the amino-acid residues that are directly involved in the formation of the intermolecular β-sheet in the type II dimer of the γ-conglutin hexamer differ from the corresponding residues in Bg7S (Fig 13) (Czubinski at al. 2015).

AIM OF THE WORK

Antagonistic interactions between pathogens and their hosts could have resulted in an ongoing evolution: plants present different sugar composition in hemicellulose fraction of cell wall, pathogens secret a collection of glycoside hydrolase to penetrate a specific cell wall, plants defence themselves producing corresponding glycoside hydrolase inhibitor proteins.

Despite the conserved structural characteristic, each plant glycoside hydrolase inhibitor proteins (GHIPs) group presents a different expression pathway and a specific cellular action. As already stated, xyloglucan-specific endo-β-1,4-glucanase inhibitor proteins-like (XEGIPs-like) are typical of dicots inhibit the hydrolytic activity of a xyloglucan-specific β-1,4-endoglucanase isolated from GH12 family. *Triticum aestivum* L. endoxylanase inhibitors (TAXIs-like) have apoplastic localization in cereals promoting their action as inhibitors of microbial xylanases GH11. The *Lupinus albus* γ-conglutin (LACGs-like) remains among GHIPs the less characterized. For this reason, most of the experimental activity focused on a γ-conglutin, a protein member of this group, found also in other legume seeds, which shows peculiar structural and functional characteristics.

Thus, the first of the aim of this project was to provide new insights about the structural features at the basis of inhibitory activity and specificity of LACGs-like GHIPs.

The second goal of the work was to deepen into the cellular responses involving the *ex-novo* synthesis of GHIPs following to pathogens attacks. This part of the work has been carried out by using *Arabidopsis thaliana* seeds germinate under different conditions to mimic pathogen infection.

MATERIALS AND METHODS

1. Bioinformatics analysis

Sequence alignments were generated using the program ClustalW2 (available on-line at http://www.ebi.ac.uk/Tools/clustalw2) (Larkin et al., 2007). Structure superimposition was generation using RasTop 2.2 (Philippe Valadon, http://www.geneinfinity.org/rastop/) using PDB access. The putative homologous protein in *Arabidopsis thaliana* were generating using TAIR BLAST 2.2.8 (http://www.arabidopsis.org/Blast/). γ-Conglutin (Uniprot accession number: Q9FSH9) was chosen like a query sequences.

1.1 Plant material

In this work we used *Arabidopsis thaliana* ecotype Columbia (Col-0). Samples of the plant tissue used for extraction of nucleic acids were harvested directly from the plant, immediately frozen in liquid nitrogen and stored at -80 °C until further use.

1.1.1 Conditions of *Arabidopsis thaliana* growth in soil

Plants of *Arabidopsis thaliana* ecotype Columbia (Col-0) were grown in chambers under controlled conditions of photoperiod and temperature. The temperature was kept constant at 21°C and lighting was from fluorescent tubes that provided cool white light with an intensity of 150 µE m-2 s -2 (Sylvania Standard F58W/133-T8). Usually, the plants were grown under inductive photoperiod conditions, which were 16h light and 8h dark (long day, LD). The seeds were suspended in distilled water and kept in darkness at 4 °C for three days in order to synchronize germination. Seeds were sown in plastic pots of 6 to 15 cm of diameter in a mixture of compost: perlite: vermiculite (2:1:1). The pots were placed in trays and watered twice every week by immersion in distilled water containing a commercial fertilizer (Algoflasth) at a 1:250 dilution. After sowing, the trays were covered with a plastic film to maintain high humidity during germination and to prevent contamination of seeds from other plants nearby. When it appeared the first true leaves, holes were made in the plastic film. The number and size of the holes were gradually increased and after two or three days the film was removed.

1.1.2 Conditions of Arabidopsis thaliana growth in solid media (Petri dishes)

In vitro culture of *Arabidopsis thaliana* ecotype Columbia (Col-0) in Petri dishes was held in growth chambers with constant temperature of 21°C, under long-day photoperiod (16 h light and 8 h dark). Growth medium (MS medium, Murashige and Skoog, 1962) contained 2.2 g/L MS salts (Duchefa), 10 g/L sucrose, 0.1 g/L MES pH 5.9 and 0.6% fitoagar (Pronadisa). The seeds were sterilized by washing them in 70% (v/v) ethanol with 0.05% (v/v) Triton X-100 for 3 min and with 95% ethanol (v/v) for 1 min and then immediately with sterile Milli-Q water in a laminar flow hood prior to sowing. The sterilized seeds were plated in Petri dishes of 9 cm of diameter (100 seeds per box). The

Petri dishes with the seeds were stored for three days at 4°C in darkness and then they were moved into the growth chamber.

1.2 Elicitation of defence responses during seed germination

1.2.1 Chitosan preparation

Low-viscous crab shells chitosan was prepared by dissolving chitosan in 0.25 N HCl. Undissolved particles were removed by centrifugation (15 min, 10,000 g). The pH of the solution was then brought to 9.5 with 2.5 N NaOH and the precipitated chitosan was recovered by filtration, washed extensively with milliQ water and lyophilized. Stock solution (10 mg/ml) was prepared by dissolving the chitosan in 0.05 N HCl and adjusting the pH to 5.5 with diluted NaOH. The solution was then autoclaved and stored at 4°C until used.

1.2.2 GH2

Mannosidase GH2 from *Cellulomonas fimi* by Megazyme (Uniprot: Q9XCV4; Length: 842 aa, Mass: 93696 Da) supplied in 3.2 M ammonium sulphate at ~ 80 U/mL – 13,2 U/mg. One Unit of β-D-mannosidase activity is defined as the amount of enzyme required to release one µmole of p-nitrophenol per minute from pNP-β-D-mannopyranoside (0.8 mM) in sodium maleate buffer (100 mM), pH 6.5 at 35°C, monitored at 400 nm.

1.2.3 GH5

Xyloglucanase GH5 from *Paenibacillus sp* by Megazyme (Uniprot: Q3MUH8; Length: 405 aa, Mass 44865 Da) supplied in 3.2 M ammonium sulphate at ~ 1,000 U/ml – 109 U/mg. One Unit of xyloglucanase activity is defined as the amount of enzyme required to release one mmole of glucose reducing-sugar equivalents per minute from xyloglucan (5 mg/mL) in sodium acetate buffer (100 mM) pH 5.5.

1.2.4 *Paenibacillus polymyxa*

Paenibacillus polymyxa is a Gram-positive, rod-shaped, motile, rhizobacterium. It is non-pathogenic and found in environments such as plant roots in soil and marine sediment. It promotes plant growth and it is demonstrated abilities of its various strains to encourage crop growth via one or more mechanisms, as well as produce lignocellulose-modifying enzymes (Lal & Tabacchioni, 2009).

1.3 Genomic Analysis

1.3.1 Extraction and purification of genomic DNA from *Arabidopsis thaliana*

A. thaliana total genomic DNA were made according to the method of extraction and purification of SureFood® PREP Basic (R-Biopharm AG) following the manufacturer's instructions. 100 mg of tissue (rosette leaves) was snap frozen in liquid nitrogen and rapidly homogenized with liquid nitrogen. Grinding was performed with a micropestle directly inside the Eppendorf tube. DNA quantification was performed in a spectrophotometer (BioPhotometer Eppendorf). Samples were stored at -80 ° C until use.

1.3.2 Extraction and purification of total RNA from Arabidopsis thaliana

To obtain *A. thaliana* total RNA used for RT-PCR, the Aurum™ Total RNA Fatty and Fibrous Tissue Kit (Bio-Rad) was used following the manufacturer's instructions. The extractions were carried out from 100 mg of tissue (seedlings or rosette leaves). Tissue was snap frozen in liquid nitrogen and rapidly homogenized with liquid nitrogen. Grinding was performed with a micropestle directly inside the tube submerged in liquid nitrogen. In addition, a digestion step with DNase (DNase I-RNase free, Qiagen) was performed at 37°C for 15 min to remove any possible genomic DNA. RNA was quantified with a biophotometer (Eppendorf) at 260 nm using a TrayCell (Hellma, Müllheim, Germany). Typically, the yield of total RNA was about 5µg from 100 mg of tissue. Samples were stored at -80 ° C until use.

1.3.3 DNA electrophoresis on agarose gels

Fragments of DNA of different size were separated by agarose gel electrophoresis. To perform the agarose gel, it was used the TBE buffer (89 mM Tris, 89 mM boric acid, 2 mM EDTA pH 8), to which 5% (v/v) of 10 mg/ml ethidium bromide was added and a variable percentage of agarose according to the size of the DNA fragment to be analysed (Kb). The samples were diluted in 6X loading buffer (0.1% bromophenol blue, 5% glycerol), whose function is to colour the sample and prevent that the sample spreads out of the gel (Sambrook et al., 1989). Electrophoresis was performed at constant voltage (80 V) and finally the DNA bands were visualized by illuminating the gel with ultraviolet light. To photograph the gel, it was used the UVITEC system, Cambridge.

1.3.4 Polymerase Chain Reaction (PCR) for genotyping

Amplification reactions were performed following the indications supplied in the kit from iQ Supermix kit (Bio-Rad), in a iCycler thermocycler (Bio-Rad). 2 µl of cDNA were diluted with water up to a final volume of 10 µl. The solution provided by the kit contains dNTPs (dATP, dCTP, dGTP, dTTP, each at a concentration of 0.4 mM), 2.5 U of DNA polymerase thermophilic eubacterium *Thermus aquaticus* BM (Taq polymerase),

$MgCl_2$, enhancers and stabilizers. The amplification cycles were divided into three sections: 30 seconds at 95°C (denaturation), 30 seconds at 55°C (annealing of specific primers) and 1 min at 72°C (elongation). The amplification cycles were repeated for 30 cycle. Finally, we added a final elongation period of 5 min at 72° C. The samples were stored at -20°C. Table 1 lists the primers used in this work, they were design based on Primers3 Input (Untergasser et al., 2012) and synthesized by Biomers.

Gene	Primers	Sequence 5' -> 3'	Tm (°C)
At5g19100	F	CTTTGTTTGCCTTCCACCGA	58.97
	R	AGGAGCCAAAGTGCTGATCT	59.01
At5g19110	F	GTCCGCCACTTCACATTCTC	58.92
	R	AGTAGTGAATGCCGGCGATA	58.96
At5g19120	F	TGTCGGTTGGCTCTGTTACT	58.95
	R	AAAACCTCCTCCACCGAACT	58.86
At1g03220	F	CTCGCTAAAGGAACCGTTGG	58.92
	R	CAGTGCTGACCGGATTGATG	58.99
At1g03230	F	CACATGCTTCTCTCCTCCGA	59.18
	R	CATTCCAGCCATACCAACGG	58.98

Tab. 1: List of primers design based on Primers3 Input, **F:** forward, **R:** reverse; **Tm:** melting temperature.

1.3.5 Synthesis of cDNA by reverse transcription (RT)

This procedure allows obtaining cDNA from RNA by the action of a reverse transcriptase (a viral enzyme that synthesizes DNA from an mRNA template). The reverse transcriptase kit used was the iScript Reverse Transcription Supermix for RT-qPCR and the instructions of Bio-Rad were followed. 50ng of total RNA were transformed into cDNA with random primers in 20 µl total volume. The kit contents also an iScript no-RT control supermix to verify the absence of residue contaminant genomic DNA not removing by DNase I treatment. The following reaction conditions have been adopted: priming for 5 min at 25°C, reverse transcription for 30 min at 42°C and RT inactivation for 5 min at 85 °C. The cDNA samples obtained were diluted 1:100 with sterile water before their use as templates in quantitative PCR. The cDNA obtained was stored at -20°C.

1.3.6 Quantitative Real-Time Polymerase Chain Reaction (PCR) for expression analysis

Amplification reactions were performed following the indications supplied in the kit from SsoAdvanced™ Universal SYBR® Green Supermix (Bio-Rad), in an iCycler thermocycler equipped with the MyiQ detection system (Bio-Rad, Milano, Italy) in 96-well optical reaction plates sealed with optical tapes (Bio-Rad). 2 µl of cDNA were diluted with water up to a final volume of 10 µl. Reactions were set up in a final volume of 20 µl. The solution provided by the kit contains dNTPs (dATP, dCTP, dGTP, dTTP, each at a concentration of 0.4 mM) and 2.5 U of DNA polymerase thermophilic eubacterium Thermus aquaticus BM (Taq polymerase). Each primer was added to a final

concentration of 0.25 µM (view the Tab. 2 containing the lists of primers used in this work and they were synthesized by Biomers). For each primer pairs, reaction efficiencies were determined from the standard dilution series of *Arabidopsis thaliana* DNA spanning five orders of magnitude and were found to range between 98% and 102%. The amplification cycles were divided into three sections: 30 seconds at 95°C (denaturation), 30 seconds at 55°C (annealing of specific primers) and 1 min at 72°C (elongation). The amplification cycles were repeated for 45 cycle, dependent on how long was the mRNA to be amplified. Finally, we added a final elongation period of 5 min at 72° C. The samples were stored at -20°C. The fluorescence signal was captured at the end of each cycle (490 nm excitation wavelength). Melting curves were obtained by progressive heating at 0.5°C every 15 s starting from 65°C for checking the specificity of the analysis. Data were collected and processed, including baselines subtraction and threshold definition, with iQ5 software (Bio-Rad, Milano, Italy). No-template control and not retro-transcribed samples (10 ng of RNA) were also included in the experimental set to detect spurious signals arising from amplification of any DNA contamination or primer dimers formed during the amplification reaction. Raw data were analysed with the iQ5 software with the following parameters: baseline from the 2nd to the 10th cycle, threshold calculated by the software for every reaction. Prior to acceptance of data for quantitative work, the amplicons in all samples were checked for specificity by the analysis of the dissociation curve. Differences in gene expression were calculated by the comparative delta–delta CT method (Pfaffl, 2001) with a dedicated Microsoft Excel macro created by Bio-Rad. cDNA from seeds after overnight vernalisation (t0) were used as control samples for quantification of those corresponding to test samples. All quantifications were normalized to the housekeeping gene 18S.

1.4 Proteomic Analysis

1.4.1 Total protein extraction (TPE)

Dry lupin seeds were ground to a meal with a coffee grinder. The flour was defatted in a Soxhlet apparatus by extraction with n-pentane at 37 °C for 4 h and then sieved through a 60 mesh sieve. The defatted flour was suspended (1:10, w/v) in 50 mM Tris–HCl buffer, pH 8.0, containing 0.5 M NaCl. After stirring at 4 °C for 3 h, the suspension was centrifuged at 10,000 g at 4 °C for 30 min. The supernatant was subsequently desalted on a Sephadex G-50 column equilibrated with 50 mM Tris–HCl buffer, pH 8.0, and quantified. TPE was immediately used for enzymatic trials.

1.4.2 Albumin and globulin purification

Albumins were extracted from each flour in distilled water in a ratio 1:10 under stirring for 4 hours at 4° C (for inhibiting the proteolytic activity). The slurry was centrifuged at 10.000 rpm for 30 minute at 4° C and the supernatant containing the albumin fraction was recovered. The pellet was resuspended with phosphate buffer (50 mM pH 7,5 and

0,5 M NaCl) and the salt soluble globulins were extracted overnight under stirring at 4°C. Then, the slurry was centrifuged at 10.000 rpm for 1 hour at 4° C, the pellet was discarded and the supernatant containing globulin fraction was recovered. Albumin and globulin fractions were purified by pigments, DNA, salts or other small molecules through Sephadex G-50 column, equilibrated in phosphate buffer 50 mM.

1.4.3 γ-Conglutin purification

γ-Conglutin was purified as described previously by using a combination of anion and cation exchange chromatography (Duranti et al., 1994). γ-Conglutin solutions were loaded onto an insulin-agarose column equilibrated in 25 mmol/L Tris-HCl, pH 7.2, and eluted by addition of 0.2 mol/L NaCl to the loading buffer (Magni et al., 2004). The purified protein was lyophilized and resuspended in the appropriate buffers before use. For the estimation of γ-conglutin concentrations, optical measurements at 280 nm were made. The extinction coefficient of 1 for a solution of 1 mg/mL was used.

1.4.4 Proteins exudate purification

Regarding the solution prepared to elicit the defence responses during seed germination, possible proteins, produced by *Arabidopsis thaliana* seeds during incubation (see Materials and Methods 1.2) and released in the medium, were analysed. Equal volumes of TCA (10% v/v) were added and put in ice for 15 min, then the samples ware centrifuged at 4°C for 15 min at 10000g. The pellets were washed with cold acetone and then dissolved directly in 100 µl DEN+.

1.4.5 SDS-polyacrylamide gel electrophoresis (SDS-PAGE)

Proteins from different samples were separated at constant amperage (16 mA each minigel) in mini Protean II (Biorad) by SDS-PAGE, prepared following the Laemmli protocol (1970): Separating gel (12%): 3.35 mL distilled H_2O water, 2.25 mL separating buffer (1.5 M Tris-HCl pH 8.8, 0.4% (w/v) SDS (Biorad)), 4 mL acrylamide 30% (Biorad), 50 µL APS (ammonium persulfate) 5% (Biorad), 10 µL TEMED (N, N, N', N-Tetramethylethylenediamine) (Biorad); Stacking gel (4.5%): 3.1 mL distilled H_2O water, 1.5 mL stacking buffer (0.5 M Tris-HCl pH 6.8, 0.4% (w/v) SDS (Biorad)), 0.9 mL acrylamide 30%, 0.5 mL BBF (Bromophenol Blue) (Biorad), 50 µL APS (ammonium persulfate) 5%, 20 µL TEMED (N, N, N', N-Tetramethylethylenediamine). Samples were prepared in reduction conditions in a ratio 1:1 with the denaturant buffer (100mM Tris-HCl pH 6.8, 4% (w/v) SDS, 0,2% (w/v) blue bromophenol, 20% (v/v) glycerol, 200 mM β-mercaptoethanol) and then denaturated at 100°C for 10 minutes (thermal denaturation). The marker used is LMW (Healthcare) in denaturant buffer. Molecular weight markers were: β phosphorylase (97kDa), BSA (66kDa), egg's albumin (45kDa), carbonic anidrase (30kDa), trypsin inhibitor (20.1kDa) and lysozyme (14.4kDa). The electrophoresis buffer was made with Tris-HCl 25 mM pH 8,3, glycine 192 mM, SDS 0,1% (w/v). At the end of the run, the gel was stained with Coomassie brilliant blue R-

250 (30% (v/v) EtOH, 0.05% (w/v) Coomassie Blue (Sigma), 10% (v/v) CH_3COOH, 0.1% (w/v) $CuSO_4$) and destained with destaining buffer (30% (v/v) EtOH, 10% (v/v) CH_3COOH).

1.4.6 Western blot Analyses

After SDS-PAGE, proteins were transferred to a PVDF membranes (Bio-Rad, Milan, Italy) by blotting according to Towbin et al. (1979) on a Trans-blot Electrophoretic Transfer Cell (Bio-Rad, Milan, Italy). The membranes were blocked for 3 h in a solution containing 1% gelatine in 0.1 M phosphate buffer saline (PBS) pH 7.0 and then washed three times with 0.25% gelatine in 0.1 M phosphate buffer saline (PBS) pH 7.0 with 0,5% Tween 20. The membrane was incubated 2 h in a solution containing anti-conglutin γ in a ratio of 100:1 (v/v). The membrane was washed with 0.1 M phosphate buffer saline (PBS) pH 7.0 with 0,5% Tween 20 and subsequently incubated for 2 h in a solution containing mouse anti-rabbit IgG conjugated to peroxidase diluted 1:300 in PBS. The membrane was washed again and placed in PBS containing 1 mg/ml of 4-cloronaphthol dissolved in methanol, the reaction was visualised by the addition 1% hydrogen peroxide.

1.5 Enzymatic assay

1.5.1 GH11

GH11 from *A. niger* xylanase and *T. longibrachiatum* (purchased from Megazyme) acid and basic xylanases have been tested by method assayed according to Qin et al. (2003) and Pauly et al. (1999) and then modifying by Gebruers et al. (2001). Enzyme assays have been carried out in 50 mM sodium acetate buffer, or 50 mM sodium phosphate buffer, pH 7.0, using 2 U of enzyme and 1.3 mg of beechwood xylan as the substrate, at 25 °C 40 min (final volume 1 mL). One unit was defined as the enzyme amount which produced 1 µmol/min of reducing sugars following beechwood xylan hydrolysis. The amounts of reducing sugars produced following enzyme activities were assayed by the p-hydroxy-benzoic acid hydrazide (PAHBAH) method (Lever, 1972).

1.5.2 GH5

GH5 from *Paenibacillus sp* by Megazime has been tested by method assayed according to Qin et al. (2003) and Pauly et al. (1999) and then modifying by Gebruers et al. (2001). Enzyme assays have been carried out in 100 Mm sodium acetate buffer pH 5.5 with BSA 1mg/ml using 2 U of enzyme and 1.3 mg of beechwood xylan as the substrate, at 25 °C 40 min (final volume 1 mL). The amounts of reducing sugars produced following enzyme activities were assayed by the p-hydroxy-benzoic acid hydrazide (PAHBAH) method (Lever, 1972).

1.5.3 GH2

GH2 from *Cellulomonas fimi* by Megazyme. Enzyme assays have been carried out in 100 mM sodium maleate buffer (100 mM), pH 6.5 with BSA 1mg/ml, using 0,1 U of enzyme and 0.08 mM of pNP-β-D-mannopyranoside. The final volume was incubated at 35°C for. The amounts of p-nitrophenol produced following enzyme activities were monitored at 400 nm. GH2 solution was preincubated for 30min at room temperature with an equal amount of sample, that possibly containing inhibitor activity, in the same buffer as enzyme solution. The mixtures were kept now at 35°C and after 10 min, the difference between the absorbance values of the sample and of the control, prepared by using buffer instead of sample, was use as a measure of the inhibition activity, express as a percentage decrease in mannosidase activity.

1.5.4 p-hydroxy-benzoic acid hydrazide Method

Acid hydrazides react in alkaline solutions with reducing carbohydrates to give yellow anions. The reaction with p-hydroxybenzoic acid hydrazide (PAHBAH), can be used in a simple colorimetric method to detect glucose or similar sugar. When sugars are reacted with excess of PAHBAH and it is treated with heat, it changes color and the solution becomes yellow (Fig. 14). This change can be monitored at 410 nm and the variation is proportional to the initial concentration of monosaccharides.

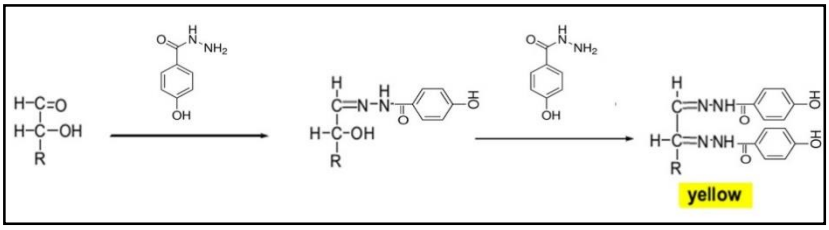


Fig. 14: p-hydroxy-benzoic acid hydrazide method; sugars reacts with excess of PAHBAH, it is treated with heat, it changes color and the solution becomes yellow.

It is necessary to calculate the calibration line, in order to linearly correlate the amount of xylose released by the action of the xylanase. As xylose, glucose possesses an aldehyde group which reacts with hydrazine. Increasing concentrations of glucose are used, all measurements are performed in duplicate to minimize errors. 1 g of 3-hydroxybenzoic acid hydrazide (Sigma) was dissolved in 20 mL of 0.5 N HCl and brought to a final volume of 200 mL with 0.5 N NaOH. The glucose was used at a concentration of 1 mg/m. The hydrazine was added to each sample containing water and glucose. The samples were boiled for 5 minutes and then the absorbance was read at 410 nm.

1.5.5 Inhibitory assays

Prior to use, γ-conglutin was dissolved in the appropriate incubation buffer (Tab. 2), briefly centrifuged and pre-incubated with the enzyme at room temperature for 15 min. In the assays where XEG were used, different amount of γ-conglutin were added

in order to obtain molar enzyme/γ-conglutin ratios of 1:1. Inhibition rates have been calculated as: (AE-AEI)/AE*100, where AE is the measured enzyme activity (µmol/min of reducing sugars produced in 40 min assays), and AEI is the measured activity of the enzyme in the presence of the inhibitor.

GH2	sodium citrate 100 mM pH 6.5
GH2 pH 7.5	50mM sodium phosphate pH 7.5 0.1 M NaCl
GH2 pH 4.8	30mM sodium acetate pH 4.8 0.1 M NaCl
GH2 pH 7.5 Cu	50mM sodium phosphate pH 7.5 0.1 M NaCl 1mM Cu
GH2 pH 4.8 Cu	30mM sodium acetate pH 4.8 0.1 M NaCl 1mM Cu
cγ pH 7.5	50mM sodium phosphate pH 7.5 0.1 M NaCl
cγ pH 4.8	30mM sodium acetate pH 4.8 0.1 M NaCl
cγ pH 7.5 Cu	50mM sodium phosphate pH 7.5 0.1 M NaCl 1mM Cu
cγ pH 4.8 Cu	30mM sodium acetate pH 4.8 0.1 M NaCl 1mM Cu
GLOB	50mM sodium phosphate pH 7.5 0.5 M NaCl

Tab. 2: Appropriate incubation buffer; GH2: enzyme from *Cellulomonas fimi* by Megazyme**; cγ:** γ-conglutin from *Lupinus albus*; **GLOB:** globuline fraction from *Lupinus albus.*

1.5.6 Size exclusion chromatography (SEC)

SEC was carried out using a Waters 625 HPLC and a Superose 12 HR 10/30 column (GE Healthcare) equilibrated in 30mM sodium acetate buffer, pH 4.8, containing 0.1 M NaCl. The flow rate of the mobile phase was 0.5 mL/min. Protein elution was monitored at 280 nm. Before use, lyophylized γ-conglutin were dissolved in elution buffer and pre-incubated with the enzyme at room temperature for 15 min, with an enzyme:inhibitor molar ratio of 1:1 (MW of *Lupinus albus* γ-conglutin: 43000 Da; MW of *Cellulomonas fimi* Mannosidase GH2: 93696 Da; MW of *Paenibacillus sp* Xyloglucanase GH5: 41359 Da).

RESULTS AND DISCUSSIONS

1.1 Bioinformatics Analysis

1.1.1 Sequence and structure analysis

GHIPs were found to be widespread present in the plant kingdom and they have common structural features. According to amino acid sequence alignments, the lupin protein shares a global amino acid sequence similarity of 54% (38% identity) with tomato, potato and tobacco XEGIPs, 55% with carrot EDGP (41% identity) and 42% with wheat TAXI-I (29% identity). In particular, the alignment of the primary structure showed that the position of the 12 cysteines is fully conserved, so it's reasonable to expect that various GHIPs have similar three-dimensional structures. However, it is interesting to note that the cysteine pairs of the disulphide bridges found in EDGP (Shang et al., 2004), TAXI-I (Sansen et al., 2004) and γ-conglutin (Scarafoni et al., 1998) are different. (*Appendix 1 summarizes the alignment*). LAGCs-like and XEGIPs-like share all the S–S bridge, while they share with TAXI-I 3 out of 6 disulphides bridges. But if we observe the alignment, we note that the exchange (Cys2-Cys4, Cys3-Cys7, Cys5-Cys6 vs Cys2-Cys5, Cys3-Cys6, Cys4-Cys7) is useful in order to maintain the three dimensional structure. The other two pairs of cysteines are present across all the proteins. Cys10-Cys11 is located at the C-terminal region of the proteins, the sequence amongst them is called "inhibition loop 1" and a conserved arginine, in XEGIPs-like, or leucine, in TAXI-I, is involved in the bond with a specific GH (Sansen et al., 2004; Naqvi et al., 2005). In the LAGCs-like it's also present the loop, but a deletion of about five amino acids involves this region, otherwise highly conserved. Cys9-Cys12 is also located at the C-terminal region, but as a matter of fact XEGIPs-like consist of single polypeptide chains (Qin et al., 2003; Shang et al., 2004; Naqvi et al., 2005), whereas LAGCs-like consists of two subunits of 29 kDa and 17 kDa and the disulphide bridge Cys9-Cys12 is being involved in the disulphide bridge which links two subunits (Scarafoni et al., 1998; Duranti et al., 2008), and TAXIs-like of, respectively, two forms, A (a polypeptide of 40 kDa) and B (two disulphide-linked subunits of about 29 and 11 kDa). However, in this latter case, both forms bind endo-xylanases and have comparable inhibition activities (Gebruers et al., 2004). This suggests that post-translational proteolysis may not affect the inhibitory capacity of these proteins. The disulphide bridge Cys9-Cys12 define another region called "inhibition loop 2" where a conserved arginine, in XEGIPs-like, or histidine, in TAXIs-like and LACGs-like, is involved in the bond with a specific GH. The sequence of the IL2 loop of γ-conglutin is more similar to the sequence of the IL2 loop of TAXI-I, rather than to the one of XEGIPs. In particular, a His residue, considered a key amino acid for the inhibitory activity of TAXI-I (Sansen et al., 2004; Pollet et al, 2009), is also present in γ-conglutin sequence (Scarafoni et al., 2010) but not in XEGIPs. The superimposition of TAXI-I 3D structure (PDB accession number: 1T6E) and EDGP 3D structure (PDB accession number: 3VLA) with γ-conglutin 3D structure (PDB accession number: 4PPH) confirm the structural difference of the loop responsible for the

interaction with the enzyme as a consequence of the five amino acid long deletion. This may be the cause of an unfavourable local spatial conformation of the protein for the correct interaction with the enzyme. The experimental 3D structures of soybean Bg7S (PDB ID: 1UAP) confirmed the structural variations in the two inhibition loops (Yoshizawa et al., 2011; Czubinski et al., 2015). The number of potential glycosylation sites along the respective amino acid sequence is different for all considered proteins. Indeed, glycosylation does not seem to be pivotal for the GH inhibitory activity. In EDGP the four putative glycosylation sites are all occupied by a glycosyl moieties (Shang et al., 2004). NEC4 and tomato XEGIP sequences show six and five potential sites, respectively. At least some of them are occupied by N-linked glycans, since these proteins are positive to concanavalin A (Qin et al., 2003; Shang et al., 2004; York et al., 2004). In all these cases, one of the sites is placed at the C-terminal end of the sequence. Only a potential N-glycosylation site was predicted of the mature TAXI-I protein (Gebruers et al., 2004), as it is for γ-conglutin. On the other hand, it has been suggested that this post-translational modification may play an important role for other possible functions of this protein, such as in signal transduction, hormone-like peptide binding or invader recognition (Gebruers et al., 2004; Shang et al., 2004; York et al., 2004).

1.1.2 Mutants analysis

Based on the structural sequence, four γ-conglutin mutants were generated. The first two mutants were designed to confer XEGIPs-like and TAXIs-like features to the unmodified recombinant γ-conglutin (rγc), by adding the respective characteristic amino acid stretches to fill the deletion in IL1. The first mutant (rγcMuX), has been designed according to NEC4 sequence (Naqvi et al., 2005), whereas the second mutant (rγcMuT), has been designed according to TAXI-I sequence (Sansen et al., 2004). Two other mutants have been prepared from rγcMuX and rγcMuT sequences, in which His406 residue of IL2 was replaced with a Ser residue (rγcMuX-H/S and rγcMuT-H/S, respectively), to assess the role of this amino acid in rγc (Scarafoni et al., 2016). Unexpectedly, when rγc was tested, inhibitory activity against a GH11, but not against GH12, was observed. The level of inhibition was about 15% (pH 5.3) and 46% (pH 7.0) (Fig. 16). The two experimental conditions were adopted for the following reasons: the acid pH isoptimal for GH activities and γ-conglutin assumes prevalently 240 a monomeric form, whereas as neutral pH, γ-conglutin takes on an hexameric quaternary structure (Capraro et al., 2010; Czubinski et al., 2015) and GH is still active, even if at about half of its maximum activity (not shown). The tests on rγcMuX and rγcMuT proteins revealed that the mutations allowed to increase the inhibitory capacity of rγc against GH11 to about 50% (rγcMuX) (Fig. 15A) and 40% (rγcMuT) (Fig. 15B), when enzyme activity was tested at pH5.3. At pH 7.0, rγcMuT showed essentially the same inhibition level of rγc, whereas rγcMuX was much less efficient, being the inhibition level limited to about 13%. Activity against GH12 was again not observed. From this data, it is clear that the presence of the inserted sequences is not the main reason for the inhibitory activity appearance (Scarafoni et al., 2016). Both rγcMuX-H/S and rγcMuT-H/S mutant proteins are devoid of the inhibitory capacity of rγc, meaning that

His406 is required for a stable interaction between the inhibitor and the target enzyme (Fig. 15C). If the inhibition activity might be attributable to the amino acid substitution or to proteolytic processing it remains to be investigated. Although the variation in PTMs of GHIPs has been previously discussed, their functional role in GHIPs remains controversial (Scarafoni et al., 2010; Yoshizawa et al., 2011). As shown, rγc was not proteolytically cleaved into subunits, since only a single polypeptide of 45 kDa is visible and no bands of 30 kDa and 17 kDa, corresponding to the processed subunits, are present. As expected, the four rγc mutants were not processed into subunits by proteolysis. Therefore, γ-conglutin is able to inhibit a GH11 and GH12 only in the unprocessed form. All active XEGIPs-like consist of a single, not processed polypeptide chain (Qin et al., 2003; Shang et al., 2004; Naqvi et al., 2005), whereas wheat TAXIs-like exists in two isoforms, namely form A (one polypeptide of 40 kDa) and form B (two subunits of 29 and 11 kDa linked by a disulphide bridge). Both isoforms have comparable inhibition properties and activities (Gebruers et al., 2004; Gebruers et al., 2010). In this case, the processing has not significant consequences on the structural characteristics, since it has no effect on the overall architecture, insomuch as both molecular forms can coexist in TAXI-I crystals (Sansen et al., 2004). The 3D structure of *L. albus* γ-conglutin has not been yet determined, but the model of the *L. angustifolius* homologous proteins available (Czubinski et al., 2015). The two lupin protein 274 have more than 90% amino acid s.i. *L. angustifolius* γ-conglutin structure refers to the seed-purified protein, which is also proteotically cleaved (Czubinski et al., 2015). The 3D structures of *L. augustifolius* γ-conglutin and Bg7S monomers, EDGP and TAXI-I are largely overlapping, a fact which reflects the general conserved structural conformation among this kind of GHIPs, regardless the PTM processing (Yoshizawa et al., 2012; Czubinski et al., 2015). IL1 and IL2 of EDGP and TAXI-I show a very similar spatial location. Conversely, superimposition of *L. angustifolius* γ-conglutin structure with EDGP and TAXI-I structures revealed that while IL2s overlap almost completely, IL1s lie with a different geometry in lupin protein (Czubinski et al., 2015), as direct consequence of the sequence deletion. We aimed to evidence possible structural repercussions due to the lack of post translational proteolytic cleavage on the rγc structure and, in the attempt, an in silico analysis has been carried out by using a homology modelling approach. Rγc was modelled using both EDPG (PDB ID: 2VLA) and TAXI-I (PDB ID: 2B42) as templates (Scarafoni et al., 2016). Representations of the two predictive models are reported in Fig. 16A and 16B, respectively (green-coloured). Both structures have been superimposed to that of *L. angustifolius* γ-conglutin (PDB ID: 4PPH, drawn in red). Even considering the intrinsic limitations of this approach, in view of the overall large overlapping, it stands out that in both cases IL1 was predicted to appear spatially closer to IL2 in the unprocessed rγc than in the proteolytically processed *L. angustifolius* γ-conglutin.

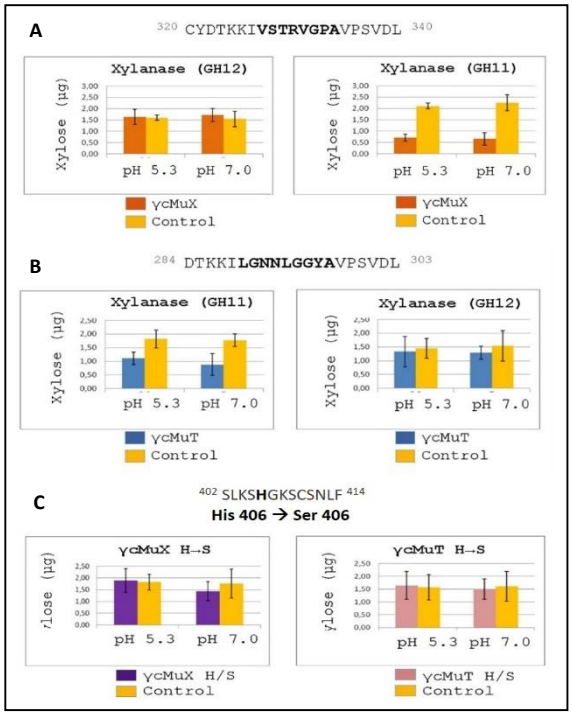
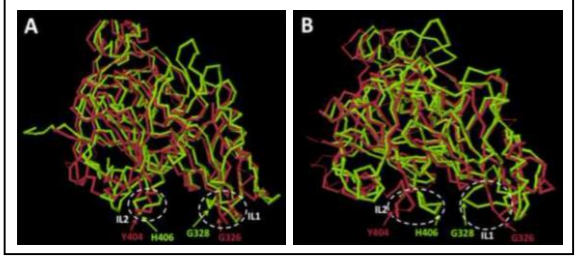


Fig. 15: Inhibitory activities. Four γ-conglutin mutants (see text) were tested against *Thricoderma longibrachiatum* xylanase (GH11) and *Aspergillus aculeatus* XEG (GH12) at pH 5.3 and pH 7.0

Fig. 16: Predictive 3D structures of rγc (reported as green backbone), obtained by homology modelling using carrot EDGP (Panel A) or and wheat TAXI-I (Panel B) proteins as templates (PDB ID: 2VLA and 2B42, respectively). Both models have been superimposed with *L. angustifolius* γ-conglutin (PDB ID: 4PPH) structure (in red). Functional region IL1 and IL2 are encircled by a white dashed line. In loop IL1 the arrows indicate the position of the amino acid stretch deletion. In loop IL2 the green arrow indicates His406 residue of rγc. In the same position of *L. angustifolius* γ-conglutin Tyr404 is instead present (red line). See text for details.



This open the way to argue that the lack of PTM processing compel the two functional loops to lie in a position which may, at least in part, compensate the sequence deletion in IL2. It is worth noting that the two wild-type legume proteins are, amongst all the homologues, not only inhibitory incompetent, but are the sole able to assembly to a quaternary structure, tetrameric in Bg7S and hexameric in γ-conglutin (Yoshizawa et al., 2011; Czubinski et al., 2015). Light scattering experiments indicated that rγc acquires a quaternary structure comparable to that of the natural protein (Capraro et al., 2010). A form of about Mr 240 kDa, compatible with a hexameric quaternary assembly, is prevalent at pH 7.0, whereas the protein is completely disassembled to monomers below pH 5.0 (approx. calculated Mr: 40 KDa). Most likely, the proteolytic processing influences the structural conformation of γ-conglutin and small local rearrangements may be the cause of the observed activity. On the whole, the results provide new insights about structural characteristics at the basis of the lack of inhibitory activity of wild-type γ-conglutin and on the mechanisms of family-specific inhibition of GH11 and GH12 by GHIPs. At the same time, they raise intriguing evolutionistic questions about the significance of the proteolytic processing which these kind of seed legume proteins undergo and about the possible origin of their target enzymes (Scarafoni et al., 2016).

1.1.3 Analysis of putative proteins in a plant model

A. thaliana is a plant widely used as a model organism in plant biology. A careful *in silico* analysis of *A. thaliana* genome, using γ-conglutin (Uniprot: Q9FSH9; EMBL: AJ297490) like a query sequence in TAIR BLAST 2.2.8 (http://www.arabidopsis.org/Blast/), allowed us to identify five different putative sequences homologous to GHIPs. At1g03220 and At1g03230 are closer related to XEGIPs-like proteins: the cysteines pathway is totally conserved, they show arginine in the inhibition loop 1 and inhibition loop 2, which

interact with glycoside hydrolases family 12, and along the sequences there are some potential glycosylation sites. At1g03220 and At1g03230 share a global amino acid sequence similarity of 94% (90% identity), probably caused by a duplication event. So the trend is to conserve these genes, even if they aren't widely expressed and their function remains unknown. In this case it was be impossible to produce a K.O. mutants to study the susceptibility at plant pathogens. At5g19110 is the unique putative proteins that shows leucine and histidine in the inhibition loop 1 and inhibition loop 2, respectively, the same region used by TAXIs-like to interact with glycoside hydrolases family GH11. Also in this case the cysteines pathway is totally conserved. In At5g19100 and At5g19120 the situation is more complicated. The cysteines pathway show substantial difference in particular at N-terminal: in At5g19100 the disulphide bond Cys5-Cy6 disappears, the same happens in At5g19120, where also the disulphide bond Cys2-Cy4 lacks. It is important to note that the active sites are situated totally at C-terminal, where the other two disulphide bond are conserved. In the same way also Cys1-Cys8, that it is responsible of the main structural retreat, are conserved and present the same sequence deletion of LACGs-like. The putative proteins show the characteristic LACGs-like deletion at the IL1, like stressed by the presence of EDGP in the sequence alignment, and in At5g19100 sequence at the IL2 we recover the histidine like both LACGs-like and TAXIs-like. Instead, in the same position in At5g19120 sequence there is a leucine.

1.1.4 Analysis of the 5' upstream region of *Arabidopsis thaliana* GHIPs genes

In silico analysis of the 5' promoter region and UTR region of *Arabidopsis thaliana* five selected genes has been carried out to analyse the regulatory elements driving the expression triggering of the genes. The presence in *Arabidopsis thaliana* genome of different three different GHIP classes suggest a fine tuning of gene expression and thus a different role of the codified proteins. This hypothesis has been also explored in previous work (Habrylo et al, 2012). Several regulatory elements have been found in up-stream region of plant genes involved in defence mechanisms. For example, GCC-box is a stress-response regulatory element triggered by pathogen attack, wounding or other stress-related conditions (Rushton and Somssich, 1998). A GCC-box occurs in the promoter regions of genes encoding PR-proteins such as chitinases and plant defensins in *Arabidopsis thaliana* (Fujimoto et al., 2000; Brown et al., 2003). W-box sequences are found in the promoter regions of many genes, including those that encode PGIPs in bean (D'Ovidio et al., 2004), PR-1 class proteins in maize (Raventós et al., 1995) and PR-10 class proteins in parsley (Rushton et al., 1996), and are known to be the DNA-binding sites of WRKY transcription factors (Rushton and Somssich, 1998; Cormack et al., 2002; Dong et al., 2003). WRKY proteins are encoded by large gene families involved in various stress related conditions. Activation sequence-1 elements in plant promoter regions are activated in reaction to three distinct plant hormones: salicylic acid, jasmonic acid, and auxin (Qin et al., 1994; Redman et al., 2002; Schiermeyer et al., 2003; Thurow et al., 2005). In response to salicylic acid, a MYB transcription factor is expressed, which specifically binds to the Myb-binding consensus sequence and, thereby, participates in transcriptional activation of PR-genes. This has for example

been observed in the mosaic virus inducible myb oncogene homolog (myb1) of tobacco (Yang and Klessig, 1996). The promoter regions of two PGIP genes of *Arabidopsis thaliana* also contain Myb consensus sequences that respond to wounding and specific elicitors (Ferrari et al., 2003). At5g19100 and At5g19110 genes do not show any known regulatory element ascribable to biotic stress. Elements found linked to At1g03220 and At1g03230 genes: "W-box", an octamer group related to cis-regulatory element, found also in promoter of *Arabidopsis thaliana* NPR1 gene, Located between +70 and +79 in tandem. W-box have been described as recognized specifically by salicylic acid (SA)-induced. At5g19120 possess AtREG588, a bZIP-binding motif involved in environmental responses; an ACGT sequence is located at position -155 to -152, and it required for etiolation-induced expression of erd1 (early responsive to dehydration) in *Arabidopsis*. A "G-box" is a binding site which can act as transcriptional repressors. It is essential for expression of beta-phaseolin gene during embryogenesis in bean, tobacco, and *Arabidopsis*. ACACNNG is a novel class of bZIP transcription factors found in the carrot Dc3 gene promoter, whose expression is normally embryo-specific, and also induced by ABA.

1.2 Analysis of proteins patterns

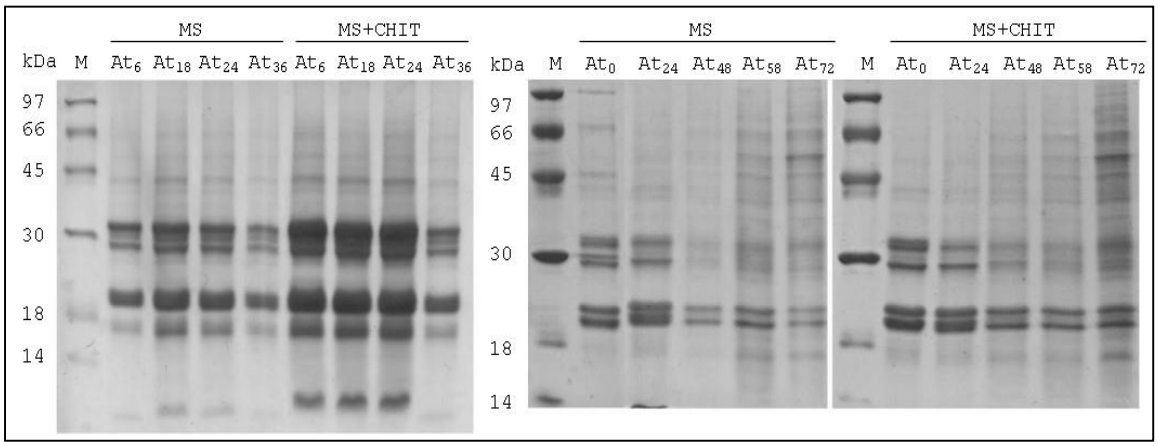


Fig. 17: Analysis of protein pattern *in Arabidopsis thaliana* seeds during germination at different condition.
M: Marker(5µl)
At*:* total protein of *Arabidopsis thaliana* seeds, germinated at different times, extracted in DEN+:H20 (1:10) (5µl)

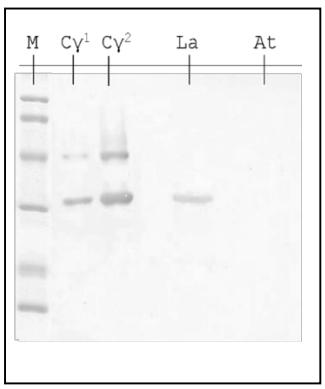


Fig. 18: Western-blot Analyses
M: Marker (5 µl)
Cγ^{1}: γ-conglutin 1mg/ml prepared in sodium acetate buffer 50mM pH 4.8 and 0,5M NaCl, DEN+ 1:1 (1,5 µl)
Cγ^{2}: γ-conglutin 1mg/ml prepared in sodium acetate buffer 50mM pH 4.8 and 0,5M NaCl, DEN+ 1:1 (15 µl)
La: total proteins of *L. albus* dry seeds (15 µl)
At: total proteins of *A. thaliana* seeds germinated for 48 hours (25 µl)

The *Arabidopsis thaliana* seeds were germinated in Petri dishes with MS medium and held in growth chambers with constant temperature of 21°C, under long-day photoperiod, for different times (0-23-48-96-144 hours) and in different conditions simulating biotic and abiotic stresses. In particular, chitosan is a polymer of β-1,4-glucosamine residues and it is a deacetylate derivative of chitin which presents antifungal properties. For this reason, chitosan can elicit plant's natural defense responses against fungal pathogens by triggering the expression of GHIPs in *A. thaliana.* Total seed proteins were extracted in DEN+:H_20 (1:1) and analyzed by SDS-PAGE to evidence changes in the protein expression pathway. The SDS-PAGE of total seed proteins extracted show a degradation during germination and major changes of seed protein fraction occur within 48 hours. But, none difference, from 0 to 72 hours, occurs between seed germinated in only MS medium and seed germinated in MS medium and then incubated for 3 hours in a chitosan (150 mg/L) solution (Fig. 17). In fact, antibody anti-Cγ failed to react with any protein of mature *A. thaliana* seeds (Fig. 8). After 96 hours of germination, the degradation is evident in all condition and through the SDS-PAGE analyses wasn't possible to appreciate any difference (Fig. 19). Instead, western-blot with antibody anti-Cγ show the presence of a similar γ-conglutin protein in *A. thaliana* seeds at 96h of germination in MS medium and then incubate in a chitosan (150 mg/L) solution for 3h (Fig. 20). It is important to note that this protein is processed in two subunits. On the other hand, we know that antibody anti- Cγ don't react with TAXIs-like or XEGIPs-like proteins.

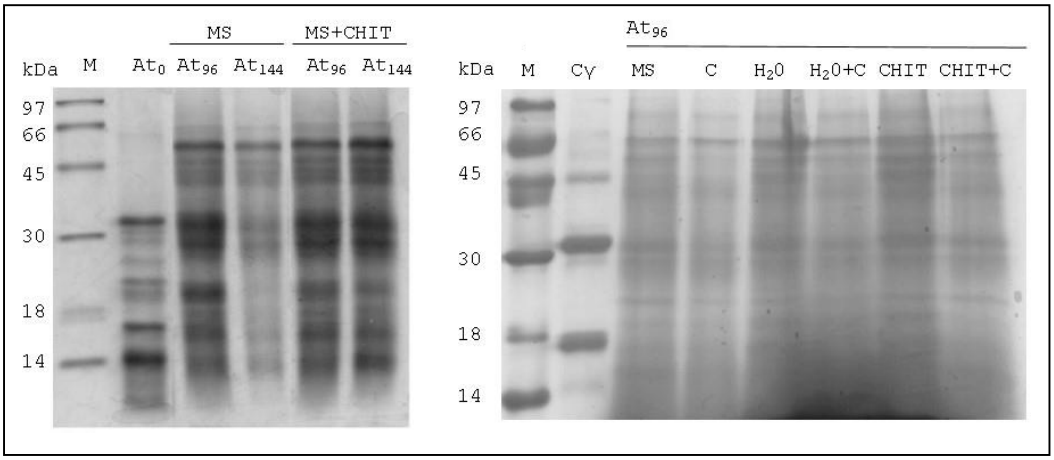


Fig 19: Analyses of protein pattern *in Arabidopsis thaliana* seeds during germination at different condition.
M: Marker(5µl)
At*:* total protein of *Arabidopsis thaliana* seeds, germinated at different times, extracted in DEN+:H20 1:10 (5µl)
Cγ: standard γ-conglutin 1mg/ml prepared in sodium acetate buffer 50mM pH 4.8 and 0,5M NaCl, DEN+ 1:1 (15µl)
At_{96}: total protein of *Arabidopsis thaliana* seeds, germinated for 96 hours in different condition (see description below), extracted rehydration buffer 1:10 (7M thiourea, 6M urea, 2% CHAPS) (25µl)

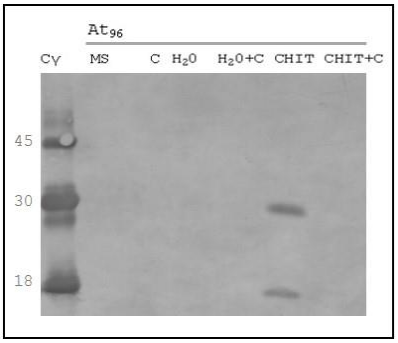


Fig 20: Western-blot Analyses
M: Marker (5 µl)
Cγ: γ-conglutin 1mg/ml prepared in sodium acetate buffer 50mM pH 4.8 and 0,5M NaCl, DEN+ 1:1 (15 µl)
At_{96}: total protein of *Arabidopsis thaliana* seeds, germinated for 96 hours in different condition extracted rehydration buffer 1:10 (7M thiourea, 6M urea, 2% CHAPS) (25µl).
MS: germination in standard condition
C: germination in presence of spontaneous *Paenibacillus polymyxa* contamination
H_20: germination and 3h of incubation in an aqueous solution
H_20+C: germination in presence of spontaneous *Paenibacillus polymyxa* contamination and 3h of incubation in an aqueous solution
CHIT: germination and 3h of incubation in 150mg/l chitosan solution
CHIT+C: germination in presence of spontaneous *Paenibacillus polymyxa* contamination and 3h of incubation in 150mg/l chitosan solution

1.3 Expression of GHIPs homologues in *Arabidopsis thaliana*

Real-time RT-PCR was used to evaluate the expression levels of the GHIPs homologues cDNAs in seeds at different germination time, in response to various biotic and abiotic stresses. The data were normalised to 18S gene expression. In standard condition (see Materials and Methods 1.1.1) the GHIPs homologues genes after 48, 96 and 144 hours of germination aren't expressed in comparison with the no expression of dry seeds (t0) (Fig. 21). In fact, the seeds at t0 were used as the calibrator and the quantification of the transcript was estimated using the comparative Ct method (as 2^(-ΔΔCT)) (Livak and Schmittgen,2001).

	MS t48	MS control	ΔCT TEST	ΔCT control	ΔΔCT	2^(-ΔΔCT)
18s	16,82	28,36	0,00	0,00	0,00	1,00
At5g19100	39,57	38,06	22,74	9,69	13,05	0,00
At5g19110	33,98	35,83	17,15	7,47	9,68	0,00
At5g19120	35,33	36,14	18,51	7,78	10,73	0,00
At1g03220	32,12	37,22	15,30	8,86	6,44	0,01
At1g03230	36,64	33,22	19,82	4,86	14,96	0,00
	MS t96	MS control	ΔCT TEST	ΔCT control	ΔΔCT	2^(-ΔΔCT)
18s	9,17	28,36	0,00	0,00	0,00	1,00
At5g19100	34,17	38,06	25,01	9,69	15,32	0,00
At5g19110	31,55	35,83	22,38	7,47	14,91	0,00
At5g19120	32,76	36,14	23,59	7,78	15,81	0,00
At1g03220	29,07	37,22	19,91	8,86	11,05	0,00
At1g03230	31,16	33,22	22,00	4,86	17,14	0,00
	MS t144	MS control	ΔCT TEST	ΔCT control	ΔΔCT	2^(-ΔΔCT)
18s	16,07	28,36	0,00	0,00	0,00	1,00
At5g19100	33,78	38,06	17,71	9,69	8,02	0,00
At5g19110	32,15	35,83	16,08	7,47	8,61	0,00
At5g19120	33,14	36,14	17,07	7,78	9,29	0,00
At1g03220	28,93	37,22	12,86	8,86	4,00	0,06
At1g03230	31,50	33,22	15,43	4,86	10,57	0,00

Fig. 21: Expression of GHIPs homologues in *Arabidopsis thaliana*
MS t48: 48h of germination; **MS t96:** 48h of germination; **MS t144**: 144h of germination in standard condition.

It is relevant to note that some dishes shown a spontaneous *Paenibacillus polymyxa* contamination after 96 hours of incubation, indicating the plants underwent a biotic stress. Comparing the seeds germinated in that condition with a standard germination on MS field for 96 hours, it is possible to appreciate an increase of expression of all analysed genes (Fig. 22A). To elicit the expression of genes in the seeds germinated in dishes not contaminated, the seeds, after germination for 96 hours, were incubated for 3 hours in an aqueous solution or in a 150mg/l chitosan solution (abiotic stress). Following this treatment, the expression raises (Fig. 22B and Fig. 22C) even if it doesn't reach the level of the expression in presence of *Paenibacillus polymyxa* contamination. Only At5g19120 is five times more expressed in presence of *Paenibacillus polymyxa* contamination and after incubation in a 150mg/l chitosan solution than in standard condition. Also the seeds germinated in presence of *Paenibacillus polymyxa* contamination were incubated for 3 hours in an aqueous solution or in a 150mg/l chitosan solution, to evaluated the combination of biotic and abiotic stress. After the incubation in the aqueous solution (Fig. 22D) At5g19100 and At5g19110 are more expressed than in the separated condition, while At5g19120 maintain the same expression levels. The osmotic pressure or the wounding could facilitate the elicitation power of *Paenibacillus polymyxa.* The gene At1g03220 and At1g03230 are, probably, one the duplication of the other and they codified for two proteins with the 98% of identity. Taken together, they keep the expression constant between two and five times greater than the standard condition. Also after the incubation in a 150mg/l chitosan solution (Fig. 22E) At5g19100 and At5g19110 are more expressed than in the separated condition, but in contrast At5g19120 is almost no expressed, whereas At1g03220 and At1g03230 are down regulated.

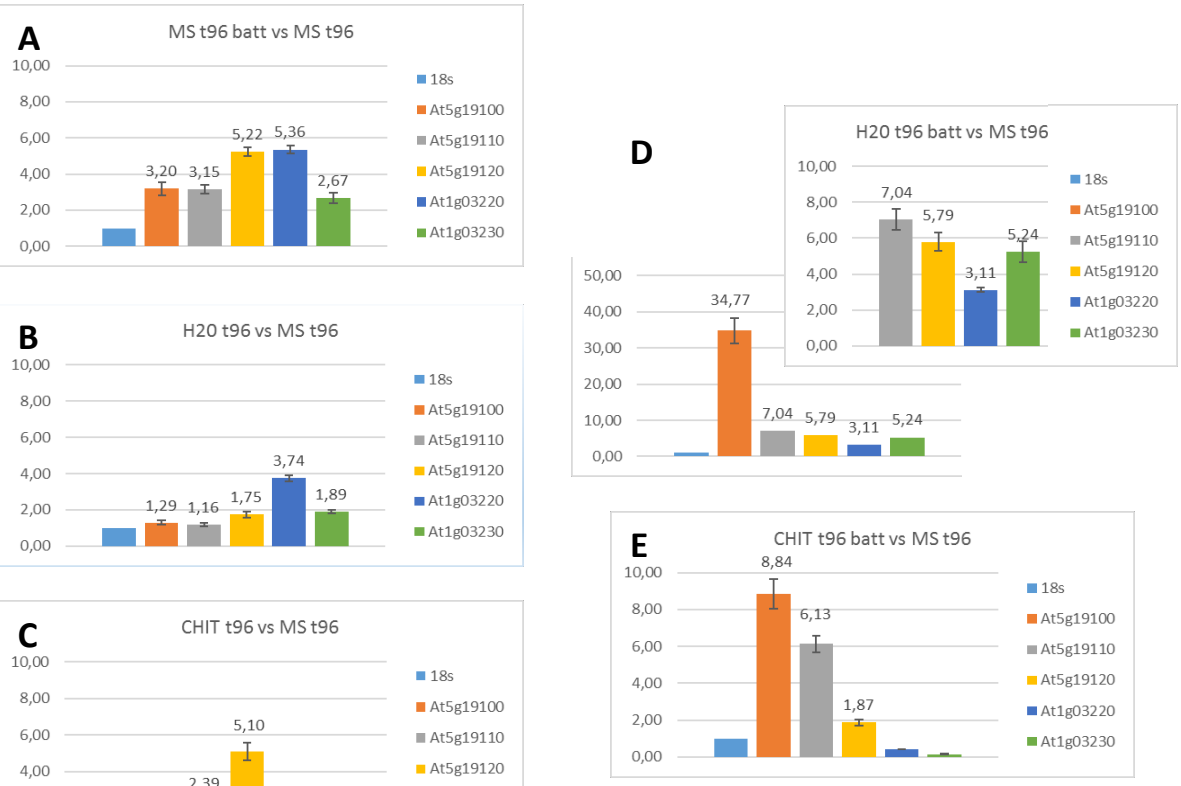


Fig. 22: Expression of GHIPs homologues in *Arabidopsis thaliana*
MS t96: 96h of germination; **MS t96 batt:** 96h of germination in presence of spontaneous *Paenibacillus polymyxa* contamination; **H2O t96:** 96h of germination and 3h of incubation in an aqueous solution; **CHIT t96:** 96h of germination and 3h of incubation in 150mg/l chitosan solution; **H2O t96 batt:** 96h of germination in presence of spontaneous *Paenibacillus polymyxa* contamination and 3h of incubation in an aqueous solution**; CHIT t96 batt:** 96h of germination in presence of spontaneous *Paenibacillus polymyxa* contamination and 3h of incubation in 150mg/l chitosan solution

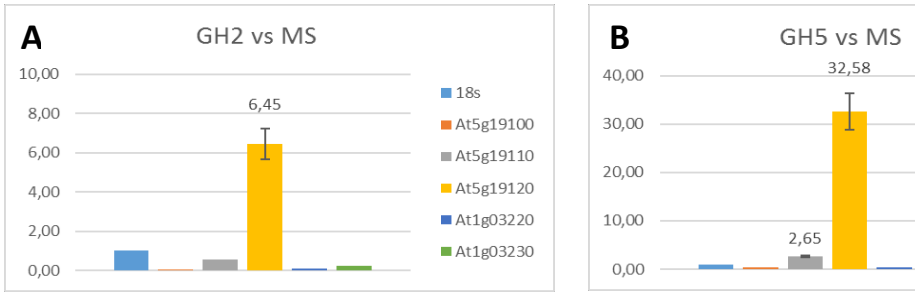
To test others condition to elicited the expression of genes, the seeds were germinated for 7 days on MS dishes and then they were incubated overnight. 7 days after germination, the young *Arabidopsis thaliana* plants present tender leaves that were easily stressed when the samples were recollected with pliers and put in the tubes for the incubation. Different incubation solutions were prepared in a final volume of 2 ml with volumes of inhibitor indicated in the Tab. 3.

GH2	*Cellulomonas fimi*	80	U/ml	10	µl
GH5	*Paenibacillus polymyxa*	1000	U/ml	2	µl
GH11A	*Aspergillus niger*	530	mU/ml	1	ml
GH11T	*Thrichoderma longibrachiatum*	400	mU/ml	1	ml
H20				2	ml

Tab. 3: Incubation solutions: Different incubation solutions were prepared in a final volume of 2 ml

Ct of RT-qPCR were analysed through ΔΔCt - Livak method, using like control cDNA from the seeds germinated on MS field for 7 days, not recollected and incubated. After the treatment

the expression raises (Fig. 23). In particular, the gene LACGs-like At5g19120 was widely expressed in all conditions, mainly after incubation with GH5 (Fig. 23B) and less with GH2 (Fig. 23A). In the control sample (incubation with water only) a strong expression of all five genes was observed (Fig. 23C). If the incubation with water was taken like control, the gene expression in the other condition was silenced (Fig. 23D). Likely, the stress caused by the recover with pliers triggered the expression. What is particularly interesting is the observation that the treatment with the GH enzymes subsequent to the mechanical stress lowered the expressions level. To explain this paradox results more experimentation is needed, even including the study of other genes known to be involved in plant stress response and in cellular events such as programmed cell death.



D H20 -wounding vs GHs

	GH2	GH5	GH11 A	GH11 T
18s	1,00	1,00	1,00	1,00
At5g19100	0,03	0,29	0,01	0,01
At5g19110	0,11	0,52	0,01	0,00
At5g19120	0,09	0,43	0,00	0,00
At1g03220	0,30	0,94	0,01	0,01
At1g03230	0,14	0,40	0,00	0,00

Fig. 23: Expression of GHIPs homologues in *Arabidopsis thaliana*
MS: 7 days (168hours) of germination; **GH2:** 7 days of germination and overnight incubation in GH2 solution; **GH5:** 7 days of germination and overnight incubation in GH5 solution; **H20-wounding:** 7 days of germination and overnight incubation in an aqueous solution: the plants were collected from MS dishes by pliers, therefore the young leaves could be damaged.

1.4 Analysis of protein Expression of GHIPs homologues in *Arabidopsis thaliana*

Total proteins were extracted, after the same treatment described above (Tab. 3), both from the medium in which the plants were incubated, either by blank (solutions identical to

those of inoculum, left in agitation for the same period of time, which, however, are not in contact with seedlings). The extracted proteins are then loaded on gel to observe the difference composition. It seems that GH2 are degraded (Fig. 24), while GH5, GH11 from *Aspergillus niger* and GH11 *Thricoderma longibrachiatum* remain stable over time.

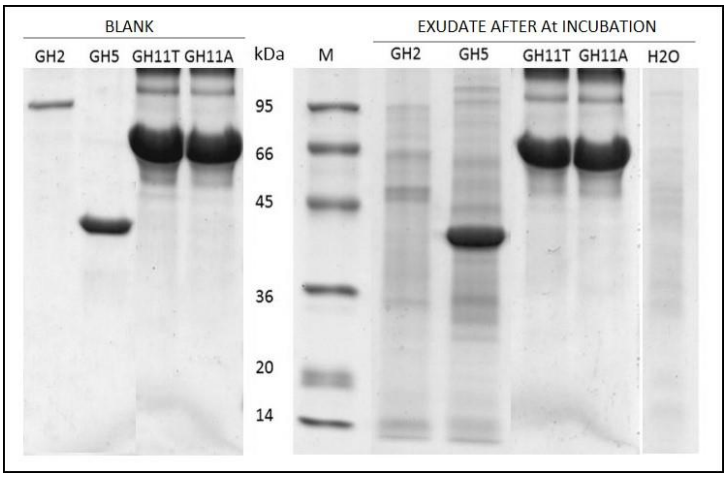


Fig. 24: Analysis of protein pattern in the solution of GHs
M: Marker (5µl).
BLANK: GHs solution incubated overnight without *Arabidopsis thaliana* seeds.
EXUDATE AFTER At INCUBATION: GHs solution incubated overnight in presence of *Arabidopsis thaliana* seeds germinated 7 days old. For the preparation of sample see Materials and Method 1.2 (5µl).

1.5 Analysis of enzymatic assay:

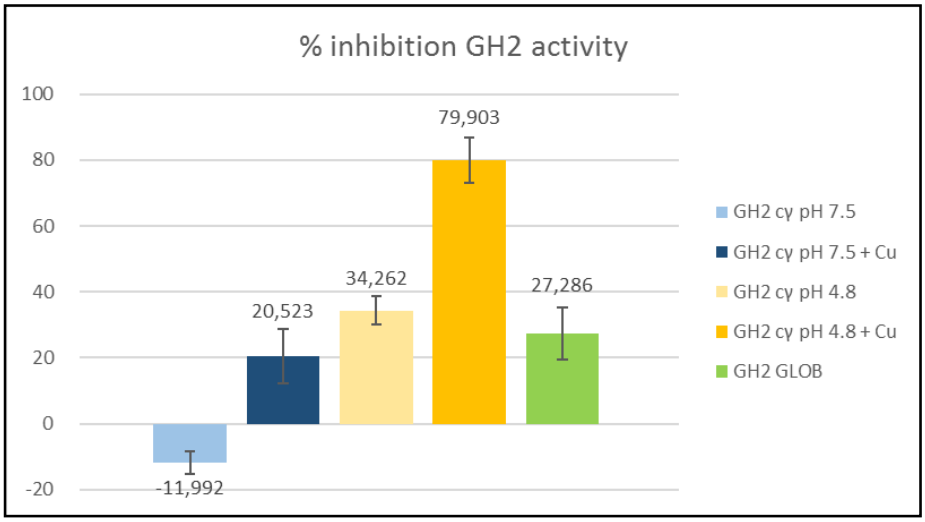


Fig. 25: Inhibition percentage of GH2 activity: γ-conglutin (**cγ**) was tested against GH2 from *Cellulomonas fimi* at pH 4.8 and pH 7.5, in presence or not of copper. Also globulins (**GLOB**) were tested against GH2.

It is best known that γ-conglutin purified from white lupin seeds does not inhibit any endo-β-glucanases against whom similar proteins act, namely members of GH11 and GH12 classes (Scarafoni et al., 2010). γ-conglutin was tested against GH2 from *Cellulomonas fimi*.

The level of inhibition was about 34% at pH 4.8 and 80% at pH 4.8 in presence of copper (Fig. 25). In the same way, the presence of copper promoted the inhibition at pH 7.5. This behaviour may reflect the quaternary structure that the protein assumes at different pH values. At acid pHs γ-conglutin assumes prevalently a monomeric form, whereas as neutral pH, γ-conglutin takes on an hexameric quaternary structure (Capraro et al., 2010; Czubinski et al., 2015). It is likely that the metal ion promotes conformational changes which favour the inhibitory competency. It is worth noting that the treatment with GH2 elicited in particular the expression of At5g19120 LACGs-like protein and the enzyme was degraded during incubation by the proteases secreted into the plant exudates.

γ-conglutin was also tested against GH5 from *Paenibacillus polymyxa*. In this case experimental problems arose, due to the commercial availability of a substrate for measuring enzyme activity. In our experimental setting, the precision of the methodology did not allow us to collect statistically reliable data. Therefore, we did not report such results in the present work.

1.6 Chromatography analysis

Size exclusion chromatography was used to investigate the interaction between cγ and the enzyme. In this experimental setting the adopted inhibitor:enzyme molecular ratio was 1:1. Seed γ-conglutin and GH2 alone were also tested as a control. The seed-extracted protein did not form any stable complex with the enzyme. γ-conglutin didn't interact neither with GH11 (Scarafoni et al., 2016), but in that case any inhibition it was observed.

CONCLUSIONS

It is well known that the fight between plants and pathogens is one of the greater evolutionary force leading to a selective production of weaponry, in pathogen, and defence systems, in plants. Plant cell walls are the first level of defence, it is composed of various polysaccharides and it acts as a physical barrier against microbial pathogens. Microorganisms secrete glycoside hydrolase (GH) to penetrate plant cell walls. In response, plants produce glycoside hydrolase inhibitor proteins GHIPs. According to the amino acid sequences, we found that the GHIPs could be grouped into three different groups: XEGIPs-like: xyloglucan-specific endo-β-1,4-glucanase inhibitor proteins, which inhibit only fungal GH12. They have been found spread across dicots plants, where the major polymers of hemicellulose fraction are xyloglucan. TAXIs-like: *Triticum aestivum* xylanase inhibitors-IA, which inhibit only GH11. They are typical of cereals, where the major polymers of hemicellulose fraction are xylan. LAGCs-like: *Lupinus albus* γ-conglutin seems to fail to inhibit the representative GHs. It is typical of legumes, where the major polymers of hemicellulose fraction are galactomannans.

The LACGs-like group remains among GHIPs the less characterized. To provide new insights about the structural features at the basis of inhibitory activity and specificity of LACGs-like, after an extended bioinformatics analysis, we identify GH2 and GH5 like the possible target of inhibition of *Lupinus albus* γ-conglutin. γ-conglutin inhibits only GH2 activity at acid pH, when it assumes monomeric form, in presence of metal ion, that promotes conformational changes which favour the inhibitory competency. GH2 is a mannosidase, so it is probably able to degrade the lupin seed cell walls, simulating a pathogenic attack.

Moreover, this work describes new experimental findings which aim to deepen into the cellular responses involving the *ex-novo* synthesis of GHIPs following to pathogens attacks. *Arabidopsis thaliana* seeds were germinated under different conditions to mimic pathogen infection. The results of this set of experiment contributed to unveil the role of each and to put the starting point of a definitive classification of GHIPs.

Arabidopsis thaliana is a model organism that present genes coding for GHIPs to three different groups. This occurs only in *A. thaliana*, since all other known major plant present only one GHIPs group.

At1g03220 and At1g03230 (XEGIPs-like) share a global amino acid sequence similarity of 94% (90% identity), probably caused by a duplication event. They are expressed after germination for 96 hours in presence of contamination and were incubated for 3 hours in an aqueous solution (both in presence or not of contamination).

At5g19110 (TAXIs-like) is expressed greater in presence of contamination of *Paenibacillus polymyxa*, in all condition: whether in MS medium than in incubation with water or chitosan.

At5g19100 (LACGs-like) is expressed in presence of contamination like At5g19110 and followed the same trend. But is also expressed in incubation with GH2 enzyme.

At5g19120 (LACGs-like) is expressed after germination for 96 hours in different condition. Western-blot with antibody anti-Cγ show the presence of a similar γ-conglutin protein in *A. thaliana* seeds at 96h of germination in MS medium and then incubate in a chitosan (150 mg/L) solution for 3h (Fig. 20). It is important to note that this protein is processed in two subunits. On the other hand, we know that antibody anti- Cγ don't react with TAXIs-like or XEGIPs-like proteins. At5g19120 (LACGs-like) expression levels are significant in particular after 7 days of incubation and incubation overnight with GH2 enzyme. It is important to note that the enzyme GH2 isn't detectable in the incubation solution. So we can hypothesise that the At5g19120 proteins, largely expressed in presence of GH2, that recognize conglutin gamma antibody, degraded GH2 enzyme in the medium. This is in relation with the inhibition of GH2 from conglutin gamma, without stable interaction it is appreciable.

By and large, this work describes new experimental findings that open new interesting scenarios to better understand some physiological aspects of the plant defence mechanisms, and provides insight of the structural basis of GHIPs inhibitory activity, specificity and repercussions on cellular responses to pathogens attacks.

REFERENCES

Albersheim, P., & Valent, B. S. (1978). Host-pathogen interactions in plants. Plants, when exposed to oligosaccharides of fungal origin, defend themselves by accumulating antibiotics. *J. Cell Biol*.;(United States), 78.

Alberts, B., Johnson, A., Lewis, J., Raff, M., Roberts, K., & Walter, P. (2002). Cell junctions, cell adhesion, and the extracellular matrix. *Molecular Biology of the Cell.* 4th edition.

Anderson, C. T., Carroll, A., Akhmetova, L., & Somerville, C. (2010). Real-time imaging of cellulose reorientation during cell wall expansion in Arabidopsis roots. *Plant physiology*, *152*(2), 787-796.

Andrews, S. R., Charnock, S. J., Lakey, J. H., Davies, G. J., Claeyssens, M., Nerinckx, W., ... & Gilbert, H. J. (2000). Substrate Specificity in Glycoside Hydrolase Family tyrosine 87 and leucine 314 play a pivotal role in discriminating between glucose and xylose binding in the proximal active site *of Pseudomonas cellulose* xylanase 10A. *Journal of Biological Chemistry*, *275*(30), 23027-23033.

Arantes, V., & Saddler, J. N. (2010). Access to cellulose limits the efficiency of enzymatic hydrolysis: the role of amorphogenesis. *Biotechnology for biofuels*, *3*(1), 1.

Attia, M., Stepper, J., Davies, G. J., & Brumer, H. (2016). Functional and structural characterization of a potent GH74 endo-xyloglucanase from the soil saprophyte Cellvibrio japonicus unravels the first step of xyloglucan degradation. *FEBS Journal*, *283*(9), 1701-1719.

Bai, S., Dong, C., Zhu, J., Zhang, Y., & Dai, H. (2015). Identification of a xyloglucan-specific endo-(1-4)-beta-d-glucanase inhibitor protein from apple (Malus× domestica Borkh.) as a potential defense gene against Botryosphaeria dothidea. *Plant Science*, 231, 11-19.

Baldwin, E. A., & Biggs, R. H. (1988). Cell-wall lysing enzymes and products of cell-wall digestion elicit ethylene in citrus. *Physiologia Plantarum*, *73*(1), 58-64.

Baumann, M. J., Eklöf, J. M., Michel, G., Kallas, Å. M., Teeri, T. T., Czjzek, M., & Brumer, H. (2007). Structural evidence for the evolution of xyloglucanase activity from xyloglucan endo-transglycosylases: biological implications for cell wall metabolism. *The Plant Cell*, *19*(6), 1947-1963.

Baydoun, E. A. H., & Fry, S. C. (1989). In vivo degradation and extracellular polymer-binding of xyloglucan nonasaccharide, a naturally-occurring anti-auxin. *Journal of Plant Physiology*, *134*(4), 453-459.

Bellincampi, D., Cervone, F., & Lionetti, V. (2015). Plant cell wall dynamics and wall-related susceptibility in plant–pathogen interactions. *Plant cell wall in pathogenesis, parasitism and symbiosis*, 30.

Biely, P., Vršanská, M., Tenkanen, M., & Kluepfel, D. (1997). Endo-β-1, 4-xylanase families: differences in catalytic properties. *Journal of biotechnology*, *57*(1), 151-166.

Bonnin, E., Daviet, S., Gebruers, K., Delcour, J. A., Goldson, A., Juge, N., & Saulnier, L. (2005). Variation in the levels of the different xylanase inhibitors in grain and flour of 20 French wheat cultivars. *Journal of cereal science*, 41(3), 375-379.

Bower, B., Kodama, K., Swanson, B., Fowler, T., Meerman, H., Collier, K., ... & Ward, M. (1998). Hyperexpression and glycosylation of Trichoderma reesei EGIII. *Special publication-Royal Society of Chemistry*, *219*, 327-334.

Br Breton, C., Šnajdrová, L., Jeanneau, C., Koča, J., & Imberty, A. (2006). Structures and mechanisms of glycosyltransferases. *Glycobiology,* 16(2), 29R-37R.

Bromley, J. R., Busse-Wicher, M., Tryfona, T., Mortimer, J. C., Zhang, Z., Brown, D. M., & Dupree, P. (2013). GUX1 and GUX2 glucuronyltransferases decorate distinct domains of glucuronoxylan with different substitution patterns. *The Plant Journal*, *74*(3), 423-434.

Brown, D. M., Zhang, Z., Stephens, E., Dupree, P., & Turner, S. R. (2009). Characterization of IRX10 and IRX10-like reveals an essential role in glucuronoxylan biosynthesis in Arabidopsis. *The Plant Journal,* 57(4), 732-746.

Brown, R. L., Kazan, K., McGrath, K. C., Maclean, D. J., & Manners, J. M. (2003). A role for the GCC-box in jasmonate-mediated activation of the PDF1. 2 gene of Arabidopsis. *Plant Physiology*, 132(2), 1020-1032.

Browner, M. F., Smith, W. W., & Castelhano, A. L. (1995). Matrilysin-inhibitor complexes: common themes among metalloproteases. *Biochemistry*, 34(20), 6602-6610.

Bucheli, P., Doares, S. H., Albersheim, P., & Darvill, A. (1990). Host-pathogen interactions XXXVI. Partial purification and characterization of heat-labile molecules secreted by the rice blast pathogen that solubilize plant cell wall fragments that kill plant cells. *Physiological and Molecular Plant Pathology*,*36*(2), 159-173.

Burton, R. A., Collins, H. M., Kibble, N. A., Smith, J. A., Shirley, N. J., Jobling, S. A., ... & Bird, A. R. (2011). Over-expression of specific HVCSLF cellulose synthase-like genes in transgenic barley increases the levels of cell wall (1, 3; 1, 4)-β-D-glucans and alters their fine structure. *Plant Biotechnology Journal*, *9*(2), 117-135.

Campbell, P., & Braam, J. (1999). Xyloglucan endotransglycosylases: diversity of genes, enzymes and potential wall-modifying functions. *Trends in plant science*, 4(9), 361-366.

Capraro, J., Galanti, E., Marengo, M., Duranti, M., & Scarafoni, A. (2015). Cysteine-containing peptides are produced by sequential clipping, but not released, from lupin 11S storage globulin during early germination. *Peptidomics,* 2(1).

Capraro, J., Spotti, P., Magni, C., Scarafoni, A., & Duranti, M. (2010). Spectroscopic studies on the pH-dependent structural dynamics of γ-conglutin, the blood glucose-lowering protein of lupin seeds. *International journal of biological macromolecules*, 47(4), 502-507.

Carbohydrate-Active enZYmes (www.cazy.org) A database that describes the families of structurally related catalytic enzymes that degrade, modify, or create glycosidic bonds.

Carpita, N. C. (1996). Structure and biogenesis of the cell walls of grasses. *Annual review of plant biology*, 47(1), 445-476.

Carpita, N. C., & Gibeaut, D. M. (1993). Structural models of primary cell walls in flowering plants: consistency of molecular structure with the physical properties of the walls during growth. *The Plant Journal*, *3*(1), 1-30.

Carter, C., & Thornburg, R. W. (2004). Is the nectar redox cycle a floral defense against microbial attack? *Trends in plant science*, 9(7), 320-324.

Cavalier, D. M., Lerouxel, O., Neumetzler, L., Yamauchi, K., Reinecke, A., Freshour, G., ... & Raikhel, N. V. (2008). Disrupting two Arabidopsis thaliana xylosyltransferase genes results in plants deficient in xyloglucan, a major primary cell wall component. *The Plant Cell*, *20*(6), 1519-1537.

Chiniquy, D., Sharma, V., Schultink, A., Baidoo, E. E., Rautengarten, C., Cheng, K., ... & Pauly, M. (2012). XAX1 from glycosyltransferase family 61 mediates xylosyltransfer to rice xylan. *Proceedings of the National Academy of Sciences*, *109*(42), 17117-17122.

Chiniquy, D., Varanasi, P., Oh, T., Harholt, J., Katnelson, J., Singh, S., ... & Ronald, P. C. (2013). Three novel rice genes closely related to the Arabidopsis 9IRX, 9IRXL, and 14 IRX genes and their roles in xylan biosynthesis. *Frontiers in plant science*, *4*.

Choi, H. W., Kim, N. H., Lee, Y. K., & Hwang, B. K. (2013). The pepper extracellular xyloglucan-specific endo-β-1, 4-glucanase inhibitor protein gene, CaXEGIP1, is required for plant cell death and defense responses. *Plant physiology*, 161(1), 384-396.

Cocuron, J. C., Lerouxel, O., Drakakaki, G., Alonso, A. P., Liepman, A. H., Keegstra, K., ... & Wilkerson, C. G. (2007). A gene from the cellulose synthase-like C family encodes a β-1, 4 glucan synthase. *Proceedings of the National Academy of Sciences*, *104*(20), 8550-8555.

Cook, B. J., Clay, R. P., Bergmann, C. W., Albersheim, P., & Darvill, A. G. (1999). Fungal polygalacturonases exhibit different substrate degradation patterns and differ in their susceptibilities to polygalacturonase-inhibiting proteins. *Molecular Plant-Microbe Interactions*, 12(8), 703-711.

Cooper, R. M., Longman, D., Campbell, A., Henry, M., & Lees, P. E. (1988). Enzymic adaptation of cereal pathogens to the monocotyledonous primary wall. *Physiological and Molecular Plant Pathology*, 32(1), 33-47.

Cormack, R. S., Eulgem, T., Rushton, P. J., Köchner, P., Hahlbrock, K., & Somssich, I. E. (2002). Leucine zipper-containing WRKY proteins widen the spectrum of immediate early elicitor-induced WRKY transcription factors in parsley. *Biochimica et Biophysica Acta (BBA)-Gene Structure and Expression*, 1576(1), 92-100.

Cosgrove, D. J. (1989). Characterization of long-term extension of isolated cell walls from growing cucumber hypocotyls. *Planta*, *177*(1), 121-130.

Cosgrove, D. J. (1997). Assembly and enlargement of the primary cell wall in plants. *Annual review of cell and developmental biology*, *13*(1), 171-201.

Cosgrove, D. J. (1999). Enzymes and other agents that enhance cell wall extensibility. *Annual review of plant biology*, *50*(1), 391-417.

Cosgrove, D. J. (2000). Loosening of plant cell walls by expansins. *Nature*,*407*(6802), 321-326.

Cosgrove, D. J. (2001). Wall structure and wall loosening. A look backwards and forwards. *Plant physiology*, *125*(1), 131-134.

Cosgrove, D. J. (2005). Growth of the plant cell wall. *Nature reviews molecular cell biology*, *6*(11), 850-861.

Côté, N., Fleury, A., Dumont-Blanchette, É., Fukamizo, T., Mitsutomi, M., & Brzezinski, R. (2006). Two exo-β-D-glucosaminidases/exochitosanases from actinomycetes define a new subfamily within family 2 of glycoside hydrolases. *Biochemical Journal*, *394*(3), 675-686.

Croes, E., Gebruers, K., Luyten, N., Delcour, J. A., & Courtin, C. M. (2009). Immunoblot quantification of three classes of proteinaceous xylanase inhibitors in different wheat (Triticum aestivum) cultivars and milling fractions. *Journal of agricultural and food chemistry*, 57(3), 1029-1035.

Croes, E., Gebruers, K., Luyten, N., Delcour, J. A., & Courtin, C. M. (2009). The three classes of wheat xylanase-inhibiting proteins accumulate in an analogous way during wheat ear development and germination. *Journal of plant physiology*, 166(12), 1253-1262.

Croes, E., Gebruers, K., Robben, J., Noben, J. P., Samyn, B., Debyser, G., ... & Courtin, C. M. (2008). Variability of polymorphic families of three types of xylanase inhibitors in the wheat grain proteome. *Proteomics*, 8(8), 1692-1705.

Czubinski, J., Barciszewski, J., Gilski, M., Szpotkowski, K., Debski, J., Lampart-Szczapa, E., & Jaskolski, M. (2015). Structure of γ-conglutin: insight into the quaternary structure of 7S basic globulins from legumes. *Acta Crystallographica Section D: Biological Crystallography*, 71(2), 224-238.

Darvill, A., Augur, C., Bergmann, C., Carlson, R. W., Cheong, J. J., Eberhard, S., ... & Mohnen, D. (1992). Oligosaccharins—oligosaccharides that regulate growth, development and defence responses in plants. *Glycobiology*, *2*(3), 181-198.

Davies, G. J., & Sinnott, M. L. (2008). Sorting the diverse: the sequence-based classifications of carbohydrate-active enzymes. *Biochemical Journal*, c1-c1.

Davies, G., & Henrissat, B. (1995). Structures and mechanisms of glycosyl hydrolases. *Structure*, *3*(9), 853-859.

Davies, G., & Sinnott, M. L. (1997). SG Withers in Comprehensive Biological Catalysis, Vol. 1.

de la Torre, F., Sampedro, J., Zarra, I., & Revilla, G. (2002). AtFXG1, an Arabidopsis gene encoding α-L-fucosidase active against fucosylated xyloglucan oligosaccharides. *Plant Physiology*, *128*(1), 247-255.

De Lorenzo, G., D'Ovidio, R., & Cervone, F. (2001). The role of polygalacturonase-inhibiting proteins (PGIPs) in defense against pathogenic fungi. *Annual review of phytopathology*, 39(1), 313-335.

Dean, J. F., & Anderson, J. D. (1991). Ethylene biosynthesis-inducing xylanase II. Purification and physical characterization of the enzyme produced by Trichoderma viride. *Plant physiology*, *95*(1), 316-323.

Dean, J. F., Gross, K. C., & Anderson, J. D. (1991). Ethylene Biosynthesis-Inducing Xylanase III. Product Characterization. *Plant physiology*, *96*(2), 571-576.

Debyser, W., Delvaux, F., & Delcour, J. A. (1998). Activity of arabinoxylan hydrolyzing enzymes during mashing with barley malt or barley malt and unmalted wheat. *Journal of agricultural and food chemistry*, 46(12), 4836-4841.

Debyser, W., Peumans, W. J., Van Damme, E. J. M., & Delcour, J. A. (1999). Triticum aestivum xylanase inhibitor (TAXI), a new class of enzyme inhibitor affecting breadmaking performance. *Journal of Cereal Science*, 30(1), 39-43.

Doblin, M. S., Pettolino, F. A., Wilson, S. M., Campbell, R., Burton, R. A., Fincher, G. B., ... & Bacic, A. (2009). A barley cellulose synthase-like CSLH gene mediates (1, 3; 1, 4)-β-D-glucan synthesis in transgenic Arabidopsis. *Proceedings of the National Academy of Sciences*, *106*(14), 5996-6001.

Doi, R. H., & Kosugi, A. (2004). Cellulosomes: plant-cell-wall-degrading enzyme complexes. *Nature Reviews Microbiology*, 2(7), 541-551.

Dong, J., Chen, C., & Chen, Z. (2003). Expression profiles of the Arabidopsis WRKY gene superfamily during plant defense response. *Plant molecular biology*, 51(1), 21-37.

Dornez, E., Croes, E., Gebruers, K., De Coninck, B., Cammue, B. P., Delcour, J. A., & Courtin, C. M. (2010). Accumulated evidence substantiates a role for three classes of wheat xylanase inhibitors in plant defense. *Critical reviews in plant sciences*, *29*(4), 244-264.

Dornez, E., Joye, I. J., Gebruers, K., Lenartz, J., Massaux, C., Bodson, B., ... & Courtin, C. M. (2006). Insight into variability of apparent endoxylanase and endoxylanase inhibitor levels in wheat kernels. *Journal of the Science of Food and Agriculture*, 86(11), 1610-1617.

D'Ovidio, R., Mattei, B., Roberti, S., & Bellincampi, D. (2004). Polygalacturonases, polygalacturonase-inhibiting proteins and pectic oligomers in plant–pathogen interactions. *Biochimica et Biophysica Acta (BBA)-Proteins and Proteomics*, *1696*(2), 237-244.

D'Ovidio, R., Raiola, A., Capodicasa, C., Devoto, A., Pontiggia, D., Roberti, S., ... & De Lorenzo, G. (2004). Characterization of the complex locus of bean encoding polygalacturonase-inhibiting proteins reveals subfunctionalization for defense against fungi and insects. *Plant Physiology*, 135(4), 2424-2435.

Duranti, M., Consonni, A., Magni, C., Sessa, F., & Scarafoni, A. (2008). The major proteins of lupin seed: characterisation and molecular properties for use as functional and nutraceutical ingredients. *Trends in Food Science & Technology*, 19(12), 624-633.

Duranti, M., Faoro, F., & Harris, N. (1994). The unusual extracellular localization of conglutin γ in germinating Lupinus albus seeds rules out its role as a storage protein. *Journal of plant physiology*, 143(6), 711-716.

Duranti, M., Gius, C., Sessa, F., & Vecchio, G. (1995). The Saccharide Chain of Lupin Seed Conglutin γ is not Responsible for the Protection of the Native Protein from Degradation by Trypsin, but Facilitates the Refolding of the Acid-Treated Protein to the Resistant Conformation. *European journal of biochemistry*, 230(3), 886-891.

Duranti, M., Restani, P., Poniatowska, M., & Cerletti, P. (1981). The seed globulins of Lupinus albus. *Phytochemistry*, 20(9), 2071-2075.

Duranti, M., Scarafoni, A., Di Cataldo, A., & Sessa, F. (2001). Interaction of metal ions with lupin seed conglutin γ. *Phytochemistry*, 56(6), 529-533.

Ebringerova, A., Hromadkova, Z., Heinze, T., & Hemicellulose, T. H. (2005). *Polysaccharides 1: Structure, characterization and use*, 186, 1-67.

Eklöf, J. M., & Brumer, H. (2010). The XTH gene family: an update on enzyme structure, function, and phylogeny in xyloglucan remodeling. *Plant Physiology*, *153*(2), 456-466.

Elliott, G. O., McLauchlan, W. R., Williamson, G., & Kroon, P. A. (2003). A wheat xylanase inhibitor protein (XIP-I) accumulates in the grain and has homologues in other cereals. *Journal of Cereal Science*, 37(2), 187-194.

Enkerli, J., Felix, G., & Boller, T. (1999). The enzymatic activity of fungal xylanase is not necessary for its elicitor activity. *Plant Physiology*, 121(2), 391-398.

Ferrari, S., Vairo, D., Ausubel, F. M., Cervone, F., & De Lorenzo, G. (2003). Tandemly duplicated Arabidopsis genes that encode polygalacturonase-inhibiting proteins are regulated coordinately by different signal transduction pathways in response to fungal infection. *The Plant Cell*, 15(1), 93-106.

Fierens, K., Brijs, K., Courtin, C. M., Gebruers, K., Goesaert, H., Raedschelders, G., ... & Delcour, J. A. (2003). Molecular identification of wheat endoxylanase inhibitor TAXI-I, member of a new class of plant proteins. *FEBS letters*, 540(1-3), 259-263.

Fierens, E., Gebruers, K., Voet, A. R., De Maeyer, M., Courtin, C. M., & Delcour, J. A. (2009). Biochemical and structural characterization of TLXI, the Triticum aestivum L. thaumatin-like xylanase inhibitor. *Journal of enzyme inhibition and medicinal chemistry*, 24(3), 646-654.

Fierens, E., Rombouts, S., Gebruers, K., Goesaert, H., Brijs, K., Beaugrand, J., ... & Delcour, J. A. (2007). TLXI, a novel type of xylanase inhibitor from wheat (Triticum aestivum) belonging to the thaumatin family. *Biochemical Journal*, 403(3), 583-591.

Fierens, K., Gils, A., Sansen, S., Brijs, K., Courtin, C. M., Declerck, P. J., ... & Campenhout, S. (2005). His374 of wheat endoxylanase inhibitor TAXI-I stabilizes complex formation with glycoside hydrolase family 11 endoxylanases. *Febs Journal*, 272(22), 5872-5882.

Flatman, R., McLAUCHLAN, R. W., Nathalie, J. U. G. E., Furniss, C., Berrin, J. G., Hughes, R. K., ... & Williamson, G. (2002). Interactions defining the specificity between fungal xylanases and the xylanase-inhibiting protein XIP-I from wheat. *Biochemical Journal*, 365(3), 773-781.

Foley, R. C., Gao, L. L., Spriggs, A., Soo, L. Y., Goggin, D. E., Smith, P. M., ... & Singh, K. B. (2011). Identification and characterisation of seed storage protein transcripts from *Lupinus angustifolius*. *BMC plant biology*, 11(1), 1.

Franková, L., & Fry, S. C. (2013). Biochemistry and physiological roles of enzymes that 'cut and paste' plant cell-wall polysaccharides. *Journal of experimental botany*, *64*(12), 3519-3550.

Fry, S. C. (1995). Polysaccharide-modifying enzymes in the plant cell wall. *Annual review of plant biology*, *46*(1), 497-520.

Fry, S. C. (2004). Primary cell wall metabolism: tracking the careers of wall polymers in living plant cells. *New phytologist*, *161*(3), 641-675.

Fry, S. C., Aldington, S., Hetherington, P. R., & Aitken, J. (1993). Oligosaccharides as signals and substrates in the plant cell wall. *Plant Physiology*, *103*(1), 1.

Fry, S. C., York, W. S., Albersheim, P., Darvill, A., Hayashi, T., Joseleau, J. P., ... & Mort, A. J. (1993). An unambiguous nomenclature for xyloglucan-derived oligosaccharides. *Physiologia Plantarum*, 89(1), 1-3.

Fuchs, Y., Saxena, A., Gamble, H. R., & Anderson, J. D. (1989). Ethylene biosynthesis-inducing protein from cellulysin is an endoxylanase. *Plant physiology*, *89*(1), 138-143.

Fujimoto, S. Y., Ohta, M., Usui, A., Shinshi, H., & Ohme-Takagi, M. (2000). Arabidopsis ethylene-responsive element binding factors act as transcriptional activators or repressors of GCC box–mediated gene expression. *The Plant Cell*, 12(3), 393-404.

García-Romera, I., & Fry, S. C. (1995). The longevity of biologically-active oligogalacturonides in rose cell cultures: degradation by exo-polygalacturonase. *Journal of experimental botany*, *46*(12), 1853-1857.

Gebruers, K., Debyser, W., Goesaert, H., Proost, P., Van Damme, J., & DELCOUR, J. A. (2001). Triticum aestivum L. endoxylanase inhibitor (TAXI) consists of two inhibitors, TAXI I and TAXI II, with different specificities. *Biochemical Journal*, 353(2), 239-244.

Gibeaut, D. M., Pauly, M., Bacic, A., & Fincher, G. B. (2005). Changes in cell wall polysaccharides in developing barley (Hordeum vulgare) coleoptiles. *Planta*, *221*(5), 729-738.

Gilbert, H. J., & Hazlewood, G. P. (1993). Bacterial cellulases and xylanases. *Microbiology*, *139*(2), 187-194.

Gilbert, H. J., Stålbrand, H., & Brumer, H. (2008). How the walls come crumbling down: recent structural biochemistry of plant polysaccharide degradation. *Current opinion in plant biology*, *11*(3), 338-348.

Gille, S., & Pauly, M. (2007). O-acetylation of plant cell wall polysaccharides. *Current challenges in plant cell walls*, 9.

Goesaert, H. (2002). Occurrence, purification and properties of endoxylanase inhibitors in different cereals (Doctoral dissertation, Katholieke Universiteit Leuven).

Goesaert, H., Debyser, W., Gebruers, K., Proost, P., Van Damme, J., & Delcour, J. A. (2001). Purification and partial characterization of an endoxylanase inhibitor from barley. *Cereal chemistry*, 78(4), 453-457.

Goesaert, H., Elliott, G., Kroon, P. A., Gebruers, K., Courtin, C. M., Robben, J., ... & Juge, N. (2004). Occurrence of proteinaceous endoxylanase inhibitors in cereals. *Biochimica et Biophysica Acta (BBA)-Proteins and Proteomics*, 1696(2), 193-202.

Goesaert, H., Gebruers, K., Brijs, K., Courtin, C. M., & Delcour, J. A. (2003). XIP-type endoxylanase inhibitors in different cereals. *Journal of cereal science*, 38(3), 317-324.

Goesaert, H., Gebruers, K., Brijs, K., Courtin, C. M., & Delcour, J. A. (2003). TAXI type endoxylanase inhibitors in different cereals. *Journal of agricultural and food chemistry*, 51(13), 3770-3775.

Goesaert, H., Gebruers, K., Courtin, C. M., & Delcour, J. A. (2005). Purification and characterization of a XIP-type endoxylanase inhibitor from rice (Oryza sativa). *Journal of enzyme inhibition and medicinal chemistry*, 20(1), 95-101.

Goesaert, H., Gebruers, K., Courtin, C. M., Proost, P., Van Damme, J., & Delcour, J. A. (2002). A family of 'TAXI'-like endoxylanase inhibitors in rye. *Journal of cereal science*, 36(2), 177-185.

Gorshkova, T. A., Wyatt, S. E., Salnikov, V. V., Gibeaut, D. M., Ibragimov, M. R., Lozovaya, V. V., & Carpita, N. C. (1996). Cell-wall polysaccharides of developing flax plants. *Plant Physiology*, *110*(3), 721-729.

Grabber, J. H. (2005). How do lignin composition, structure, and cross-linking affect degradability? A review of cell wall model studies. *Crop Science*, *45*(3), 820-831.

Greenberg, J. T. (1997). Programmed cell death in plant-pathogen interactions. *Annual review of plant biology*, 48(1), 525-545.

Ha, M. A., Apperley, D. C., & Jarvis, M. C. (1997). Molecular rigidity in dry and hydrated onion cell walls. *Plant Physiology*, *115*(2), 593-598.

Habrylo, O., Forster, A., Jeltsch, J. M., & Phalip, V. (2013). The characterisation of xyloglucanase inhibitors from *Humulus lupulus*. *Phytochemistry*, 90, 70-77.

Hanania, U., & Avni, A. (1997). High-affinity binding site for ethylene-inducing xylanase elicitor on Nicotiana tabacum membranes. *The Plant Journal*, 12(1), 113-120.

Harholt, J., Suttangkakul, A., & Scheller, H. V. (2010). Biosynthesis of pectin. *Plant Physiology*, *153*(2), 384-395.

Harper, A. D., Stalnaker, S. H., Wells, L., Darvill, A., Thornburg, R., & York, W. S. (2010). Interaction of Nectarin 4 with a fungal protein triggers a microbial surveillance and defense mechanism in nectar. *Phytochemistry*, 71(17), 1963-1969.

Hatfield, R., Ralph, J., & Grabber, J. H. (2008). A potential role for sinapyl p-coumarate as a radical transfer mechanism in grass lignin formation. *Planta*,*228*(6), 919-928.

Hayashi, T. (1989). Xyloglucans in the primary cell wall. *Annual review of plant biology*, *40*(1), 139-168.

Henrissat, B. (1991). A classification of glycosyl hydrolases based on amino acid sequence similarities. *Biochemical Journal*, *280*(2), 309-316.

Henrissat, B., & Bairoch, A. (1996). Updating the sequence-based classification of glycosyl hydrolases. *Biochemical Journal*, *316*(Pt 2), 695.

Henrissat, B., & Davies, G. (1997). Structural and sequence-based classification of glycoside hydrolases. *Current opinion in structural biology*, 7(5), 637-644.

Henrissat, B., Callebaut, I., Fabrega, S., Lehn, P., Mornon, J. P., & Davies, G. (1995). Conserved catalytic machinery and the prediction of a common fold for several families of glycosyl hydrolases. *Proceedings of the National Academy of Sciences*, *92*(15), 7090-7094.

Hörnblad, E., Ulfstedt, M., Ronne, H., & Marchant, A. (2013). Partial functional conservation of IRX10 homologs in physcomitrella patens and Arabidopsis thaliana indicates an evolutionary step contributing to vascular formation in land plants. *BMC plant biology*, *13*(1), 1.

Igawa, T., Ochiai-Fukuda, T., Takahashi-Ando, N., Ohsato, S., Shibata, T., Yamaguchi, I., & Kimura, M. (2004). New TAXI-type xylanase inhibitor genes are inducible by pathogens and wounding in hexaploid wheat. *Plant and cell physiology*, *45*(10), 1347-1360.

Jenkins, J., Lo Leggio, L., Harris, G., & Pickersgill, R. (1995). β-Glucosidase, β-galactosidase, family A cellulases, family F xylanases and two barley glycanases form a superfamily of enzymes wit 8-fold β/α architecture and with two conserved glutamates near the carboxy-terminal ends of β-strands four and seven. *FEBS letters*, 362(3), 281-285.

Johansson, P., Brumer, H., Baumann, M. J., Kallas, Å. M., Henriksson, H., Denman, S. E., ... & Jones, T. A. (2004). Crystal structures of a poplar xyloglucan endotransglycosylase reveal details of transglycosylation acceptor binding. *The Plant Cell*, *16*(4), 874-886.

Jones, R. W. (2012). Multiple copies of genes encoding XEGIPs are harbored in an 85-kB region of the potato genome. *Plant Molecular Biology Reporter*, 30(4), 1040-1046.

Jones, R. W., & Perez, F. G. (2014). Constitutive Expression of a XEGIP in Potato Results in Phenotypic Changes Suggesting Endogenous Inhibition of Cell Wall Growth. *Potato research*, 57(2), 133-144.

Keegstra, K. (2010). Plant cell walls. *Plant physiology*, *154*(2), 483-486.

Keegstra, K., Talmadge, K. W., Bauer, W. D., & Albersheim, P. (1973). The structure of plant cell walls III. A model of the walls of suspension-cultured sycamore cells based on the interconnections of the macromolecular components. *Plant physiology*, *51*(1), 188-197.

Keppler, B. D., & Showalter, A. M. (2010). IRX14 and IRX14-LIKE, two glycosyl transferases involved in glucuronoxylan biosynthesis and drought tolerance in Arabidopsis. *Molecular Plant*, *3*(5), 834-841.

Kiefer, L. L., York, W. S., Darvill, A. G., & Albersheim, P. (1989). Xyloglucan isolated from suspension-cultured sycamore cell walls is O-acetylated. *Phytochemistry*, *28*(8), 2105-2107.

Kohorn, B. D., Johansen, S., Shishido, A., Todorova, T., Martinez, R., Defeo, E., & Obregon, P. (2009). Pectin activation of MAP kinase and gene expression is WAK2 dependent. *The Plant Journal*, *60*(6), 974-982.

Kulkarni, N., Shendye, A., & Rao, M. (1999). Molecular and biotechnological aspects of xylanases. *FEMS microbiology reviews*, 23(4), 411-456.

Labavitch, J. M. (1981). Cell wall turnover in plant development. *Annual Review of Plant Physiology*, *32*(1), 385-406.

Lal, S., & Tabacchioni, S. (2009). Ecology and biotechnological potential of Paenibacillus polymyxa: a minireview. *Indian Journal of Microbiology*, 49(1), 2-10.

Larkin, M. A., Blackshields, G., Brown, N. P., Chenna, R., McGettigan, P. A., McWilliam, H., ... & Thompson, J. D. (2007). Clustal W and Clustal X version 2.0. bioinformatics, 23(21), 2947-2948.

Lerouxel, O., Choo, T. S., Séveno, M., Usadel, B., Faye, L., Lerouge, P., & Pauly, M. (2002). Rapid structural phenotyping of plant cell wall mutants by enzymatic oligosaccharide fingerprinting. *Plant Physiology*, *130*(4), 1754-1763.

Lever, M. (1972). A new reaction for colorimetric determination of carbohydrates. *Analytical biochemistry*, 47(1), 273-279.

Li, L. C., & Cosgrove, D. J. (2001). Grass group I pollen allergens (β-expansins) lack proteinase activity and do not cause wall loosening via proteolysis. *European Journal of Biochemistry*, 268(15), 4217-4226.

Li, S., Bashline, L., Lei, L., & Gu, Y. (2014). Cellulose synthesis and its regulation. *The Arabidopsis Book*, e0169.

Li, Z. C., Durachko, D. M., & Cosgrove, D. J. (1993). An oat coleoptile wall protein that induces wall extension in vitro and that is antigenically related to a similar protein from cucumber hypocotyls. *Planta*, *191*(3), 349-356.

Liepman, A. H., Wilkerson, C. G., & Keegstra, K. (2005). Expression of cellulose synthase-like (Csl) genes in insect cells reveals that CslA family members encode mannan synthases. *Proceedings of the National Academy of Sciences of the United States of America*, 102(6), 2221-2226.

Lombard, V., Ramulu, H. G., Drula, E., Coutinho, P. M., & Henrissat, B. (2014). The carbohydrate-active enzymes database (CAZy) in 2013. *Nucleic acids research*, *42*(D1), D490-D495.

Livak, K. J., & Schmittgen, T. D. (2001). Analysis of relative gene expression data using real-time quantitative PCR and the 2− ΔΔCT method. *methods*, 25(4), 402-408.

Ma, Z., Song, T., Zhu, L., Ye, W., Wang, Y., Shao, Y., ... & Tyler, B. M. (2015). A Phytophthora sojae glycoside hydrolase 12 protein is a major virulence factor during soybean infection and is recognized as a PAMP. *The Plant Cell*, 27(7), 2057-2072.

Magni, C., Sessa, F., Accardo, E., Vanoni, M., Morazzoni, P., Scarafoni, A., & Duranti, M. (2004). Conglutin γ, a lupin seed protein, binds insulin in vitro and reduces plasma glucose levels of hyperglycemic rats. *The Journal of nutritional biochemistry*, 15(11), 646-650.

Martinez-Fleites, C., Guerreiro, C. I., Baumann, M. J., Taylor, E. J., Prates, J. A., Ferreira, L. M., ... & Davies, G. J. (2006). Crystal structures of Clostridium thermocellum xyloglucanase, XGH74A, reveal the structural basis for xyloglucan recognition and degradation. *Journal of Biological Chemistry*,*281*(34), 24922-24933.

Master, E. R., Zheng, Y., Storms, R., Tsang, A., & Powlowski, J. (2008). A xyloglucan-specific family 12 glycosyl hydrolase from Aspergillus niger: recombinant expression, purification and characterization. *Biochemical Journal*, 411(1), 161-170.

Matsunaga, T., Ishii, T., Matsumoto, S., Higuchi, M., Darvill, A., Albersheim, P., & O'Neill, M. A. (2004). Occurrence of the primary cell wall polysaccharide rhamnogalacturonan II in pteridophytes, lycophytes, and bryophytes. Implications for the evolution of vascular plants. *Plant Physiology*, *134*(1), 339-351.

McCarter, J. D., & Withers, G. S. (1994). Mechanisms of enzymatic glycoside hydrolysis. *Current opinion in structural biology*, *4*(6), 885-892.

McIntosh, L. P., Hand, G., Johnson, P. E., Joshi, M. D., Körner, M., Plesniak, L. A., ... & Withers, S. G. (1996). The p K a of the General Acid/Base Carboxyl Group of a Glycosidase Cycles during Catalysis: A 13C-NMR Study of Bacillus circulans Xylanase. *Biochemistry*, *35*(31), 9958-9966.

McQueen-Mason, S. J., & Cosgrove, D. J. (1995). Expansin mode of action on cell walls (analysis of wall hydrolysis, stress relaxation, and binding). *Plant Physiology*, *107*(1), 87-100.

McQueen-Mason, S. J., Fry, S. C., Durachko, D. M., & Cosgrove, D. J. (1993). The relationship between xyloglucan endotransglycosylase and in-vitro cell wall extension in cucumber hypocotyls. *Planta*, 190(3), 327-331.

McQueen-Mason, S., Durachko, D. M., & Cosgrove, D. J. (1992). Two endogenous proteins that induce cell wall extension in plants. *The Plant Cell*, *4*(11), 1425-1433.

Minic, Z. (2008). Physiological roles of plant glycoside hydrolases. *Planta*,*227*(4), 723-740.

Mohnen, D. (2008). Pectin structure and biosynthesis. *Current opinion in plant biology*, *11*(3), 266-277.

Moller, I., Sørensen, I., Bernal, A. J., Blaukopf, C., Lee, K., Øbro, J., ... & Bacic, A. (2007). High-throughput mapping of cell-wall polymers within and between plants using novel microarrays. *The Plant Journal*, 50(6), 1118-1128.

Morel, J. B., & Dangl, J. L. (1997). The hypersensitive response and the induction of cell death in plants. *Cell death and differentiation*, 4(8), 671-683.

Murashige, T., & Skoog, F. (1962). A revised medium for rapid growth and bio assays with tobacco tissue cultures. *Physiologia plantarum*, 15(3), 473-497.

Naqvi, S. S., Harper, A., Carter, C., Ren, G., Guirgis, A., York, W. S., & Thornburg, R. W. (2005). Nectarin IV, a potent endoglucanase inhibitor secreted into the nectar of ornamental tobacco plants. Isolation, cloning, and characterization. *Plant physiology*, 139(3), 1389-1400.

Nishikubo, N., Awano, T., Banasiak, A., Bourquin, V., Ibatullin, F., Funada, R., ... & Mellerowicz, E. J. (2007). Xyloglucan endo-transglycosylase (XET) functions in gelatinous layers of tension wood fibers in poplar—a glimpse into the mechanism of the balancing act of trees. *Plant and cell physiology*,*48*(6), 843-855.

Okada, H., Mori, K., Tada, K., Nogawa, M., & Morikawa, Y. (2000). Identification of active site carboxylic residues in Trichoderma reesei endoglucanase Cel12A by site-directed mutagenesis. *Journal of Molecular Catalysis B: Enzymatic*, *10*(1), 249-255.

O'Neill, M. A., Eberhard, S., Albersheim, P., & Darvill, A. G. (2001). Requirement of borate cross-linking of cell wall rhamnogalacturonan II for Arabidopsis growth. *Science*, *294*(5543), 846-849.

O'Neill, M. A., Ishii, T., Albersheim, P., & Darvill, A. G. (2004). Rhamnogalacturonan II: structure and function of a borate cross-linked cell wall pectic polysaccharide. *Annual Review Plant Biology*, *55*, 109-139.

Orengo, C. A., Jones, D. T., & Thornton, J. M. (1994). Protein superfamilies and domain superfolds. *Nature*, *372*(6507), 631-634.

Paës, G., Berrin, J. G., & Beaugrand, J. (2012). GH11 xylanases: structure/function/properties relationships and applications. *Biotechnology advances*, 30(3), 564-592.

Pauly, M., & Keegstra, K. (2016). Biosynthesis of the Plant Cell Wall Matrix Polysaccharide Xyloglucan. *Annual review of plant biology*, *67*, 235-259.

Pauly, M., Andersen, L. N., Kauppinen, S., Kofod, L. V., York, W. S., Albersheim, P., & Darvill, A. (1999). A xyloglucan-specific endo-β-1, 4-glucanase from Aspergillus aculeatus: expression cloning in yeast, purification and characterization of the recombinant enzyme. *Glycobiology*,9(1), 93-100.

Pauly, M., Gille, S., Liu, L., Mansoori, N., de Souza, A., Schultink, A., & Xiong, G. (2013). Hemicellulose biosynthesis. *Planta*, *238*(4), 627-642.

Pauly, M., Qin, Q., Greene, H., Albersheim, P., Darvill, A., & York, W. S. (2001). Changes in the structure of xyloglucan during cell elongation. *Planta*, *212*(5-6), 842-850.

Payan, F., Flatman, R., Porciero, S., Williamson, G., Nathalie, J. U. G. E., & Roussel, A. (2003). Structural analysis of xylanase inhibitor protein I (XIP-I), a proteinaceous xylanase inhibitor from wheat (Triticum aestivum, var. Soisson). *Biochemical Journal*, 372(2), 399-405.

Payan, F., Leone, P., Porciero, S., Furniss, C., Tahir, T., Williamson, G., ... & Roussel, A. (2004). The dual nature of the wheat xylanase protein inhibitor XIP-I structural basis for the inhibition of family 10 and family 11 xylanases. *Journal of Biological Chemistry*, 279(34), 36029-36037.

Pell, G., Taylor, E. J., Gloster, T. M., Turkenburg, J. P., Fontes, C. M., Ferreira, L. M., ... & Gilbert, H. J. (2004). The mechanisms by which family 10 glycoside hydrolases bind decorated substrates. *Journal of Biological Chemistry*, *279*(10), 9597-9605.

Peña, M. J., Zhong, R., Zhou, G. K., Richardson, E. A., O'Neill, M. A., Darvill, A. G., ... & Ye, Z. H. (2007). Arabidopsis irregular xylem8 and irregular xylem9: implications for the complexity of glucuronoxylan biosynthesis. *The Plant Cell*, *19*(2), 549-563.

Persson, S., Paredez, A., Carroll, A., Palsdottir, H., Doblin, M., Poindexter, P., ... & Somerville, C. R. (2007). Genetic evidence for three unique components in primary cell-wall cellulose synthase complexes in Arabidopsis. *Proceedings of the National Academy of Sciences*, *104*(39), 15566-15571.

Pfaffl, M. W. (2001). A new mathematical model for relative quantification in real-time RT–PCR. *Nucleic acids research*, 29(9), e45-e45.

Planas, A. (2000). Bacterial 1, 3-1, 4-β-glucanases: structure, function and protein engineering. *Biochimica et Biophysica Acta (BBA)-Protein Structure and Molecular Enzymology*, *1543*(2), 361-382.

Pollet, A., Delcour, J. A., & Courtin, C. M. (2010). Structural determinants of the substrate specificities of xylanases from different glycoside hydrolase families. *Critical reviews in biotechnology*, *30*(3), 176-191

Pollet, A., Sansen, S., Raedschelders, G., Gebruers, K., Rabijns, A., Delcour, J. A., & Courtin, C. M. (2009). Identification of structural determinants for inhibition strength and specificity of wheat xylanase inhibitors TAXI-IA and TAXI-IIA. *FEBS journal*, 276(14), 3916-3927.

Popper, Z. A., & Fry, S. C. (2003). Primary cell wall composition of bryophytes and charophytes. *Annals of botany*, *91*(1), 1-12.

Popper, Z. A., & Fry, S. C. (2008). Xyloglucan− pectin linkages are formed intra-protoplasmically, contribute to wall-assembly, and remain stable in the cell wall. *Planta*, *227*(4), 781-794.

Qin, Q., Bergmann, C. W., Rose, J. K., Saladie, M., Kolli, V. S., Albersheim, P., ... & York, W. S. (2003). Characterization of a tomato protein that inhibits a xyloglucan-specific endoglucanase. *The Plant Journal*, 34(3), 327-338.

Qin, X. F., Holuigue, L., Horvath, D. M., & Chua, N. H. (1994). Immediate early transcription activation by salicylic acid via the cauliflower mosaic virus as-1 element. *The Plant Cell*, 6(6), 863-874.

Raedschelders, G., Fierens, K., Sansen, S., Rombouts, S., Gebruers, K., Robben, J., ... & Volckaert, G. (2005). Molecular identification of wheat endoxylanase inhibitor TAXI-II and the determinants of its inhibition specificity. *Biochemical and biophysical research communications*, 335(2), 512-522.

Raventós, D., Jensen, A. B., Rask, M. B., Casacuberta, J. M., Mundy, J., & Segundo, B. (1995). A 20 bp cis-acting element is both necessary and sufficient to mediate elicitor response of a maize PRms gene. *The Plant Journal*, 7(1), 147-155.

Redman, J., Whitcraft, J., Johnson, C., & Arias, J. (2002). Abiotic and biotic stress differentially stimulate as-1 element activity in Arabidopsis. *Plant Cell Reports*, 21(2), 180-185.

Restani, P., Duranti, M., Cerletti, P., & Simonetti, P. (1981). Subunit composition of the seed globulins of Lupinus albus. *Phytochemistry*, 20(9), 2077-2083.

Richmond, T. A., & Somerville, C. R. (2000). The cellulose synthase superfamily. *Plant physiology*, *124*(2), 495-498.

Ridley, B. L., O'Neill, M. A., & Mohnen, D. (2001). Pectins: structure, biosynthesis, and oligogalacturonide-related signaling. *Phytochemistry*,*57*(6), 929-967.

Rombouts, S., Fierens, E., Vandermarliere, E., Voet, A., Gebruers, K., Beaugrand, J., ... & Van Campenhout, S. (2009). His22 of TLXI plays a critical role in the inhibition of glycoside hydrolase family 11 xylanases. *Journal of enzyme inhibition and medicinal chemistry*, 24(1), 38-46.

Rose, J. K., Braam, J., Fry, S. C., & Nishitani, K. (2002). The XTH family of enzymes involved in xyloglucan endotransglucosylation and endohydrolysis: current perspectives and a new unifying nomenclature. *Plant and Cell Physiology*, *43*(12), 1421-1435.

Rushton, P. J., & Somssich, I. E. (1998). Transcriptional control of plant genes responsive to pathogens. *Current opinion in plant biology*, 1(4), 311-315.

Rushton, P. J., Torres, J. T., Parniske, M., Wernert, P., Hahlbrock, K., & Somssich, I. E. (1996). Interaction of elicitor-induced DNA-binding proteins with elicitor response elements in the promoters of parsley PR1 genes. *The EMBO journal*, 15(20), 5690.

Ryan, C. A., & Farmer, E. E. (1991). Oligosaccharide signals in plants: a current assessment. *Annual review of plant biology*, *42*(1), 651-674.

Rye, C. S., & Withers, S. G. (2000). Glycosidase mechanisms. *Current opinion in chemical biology*, 4(5), 573-580.

Sambrook, J., Fritsch, E. F., & Maniatis, T. (1989). Molecular cloning: a laboratory manual (No. Ed. 2). *Cold spring harbor laboratory press*.

Sampedro, J., & Cosgrove, D. J. (2005). The expansin superfamily. *Genome biology*, 6(12), 1.

Sampedro, J., Pardo, B., Gianzo, C., Guitián, E., Revilla, G., & Zarra, I. (2010). Lack of α-xylosidase activity in Arabidopsis alters xyloglucan composition and results in growth defects. *Plant physiology*, *154*(3), 1105-1115.

Sancho, A. I., Faulds, C. B., Svensson, B., Bartolomé, B., Williamson, G., & Juge, N. (2003). Cross-inhibitory activity of cereal protein inhibitors against α-amylases and xylanases. *Biochimica et Biophysica Acta (BBA)-Proteins and Proteomics*, 1650(1), 136-144.

Sandgren, M., Ståhlberg, J., & Mitchinson, C. (2005). Structural and biochemical studies of GH family 12 cellulases: improved thermal stability, and ligand complexes. *Progress in biophysics and molecular biology*, *89*(3), 246-291.

Sansen, S., De Ranter, C. J., Gebruers, K., Brijs, K., Courtin, C. M., Delcour, J. A., & Rabijns, A. (2004). Structural basis for inhibition of Aspergillus niger xylanase by Triticum aestivum xylanase inhibitor-I. *Journal of Biological Chemistry*, 279(34), 36022-36028.

Saulnier, L., Marot, C., Chanliaud, E., & Thibault, J. F. (1995). Cell wall polysaccharide interactions in maize bran. *Carbohydrate polymers*, *26*(4), 279-287.

Scarafoni, A., Consonni, A., Pessina, S., Balzaretti, S., Capraro, J., Galanti, E., & Duranti, M. (2016). Structural basis of the lack of endo-glucanase inhibitory activity of Lupinus albus γ-conglutin. *Plant Physiology and Biochemistry*, 99, 79-85.

Scarafoni, A., Di Cataldo, A., Vassilevskaia, T. D., Bekman, E. P., Rodrigues-Pousada, C., Ceciliani, F., & Duranti, M. (2001). Cloning, sequencing and expression in the seeds and radicles of two Lupinus albus conglutin γ genes. *Biochimica et Biophysica Acta (BBA)-Gene Structure and Expression*, 1519(1), 147-151.

Scarafoni, A., Ronchi, A., & Duranti, M. (2010). γ-Conglutin, the Lupinus albus XEGIP-like protein, whose expression is elicited by chitosan, lacks of the typical inhibitory activity against GH12 endo-glucanases. *Phytochemistry*, 71(2), 142-148.

Scheller, H. V., & Ulvskov, P. (2010). Hemicelluloses. *Plant Biology*, *61*(1), 263.

Schiermeyer, A., Thurow, C., & Gatz, C. (2003). Tobacco bZIP factor TGA10 is a novel member of the TGA family of transcription factors. *Plant molecular biology*, 51(6), 817-829.

Schröder, R., Wegrzyn, T. F., Bolitho, K. M., & Redgwell, R. J. (2004). Mannan transglycosylase: a novel enzyme activity in cell walls of higher plants. *Planta*, 219(4), 590-600.

Schultink, A., Liu, L., Zhu, L., & Pauly, M. (2014). Structural diversity and function of xyloglucan sidechain substituents. *Plants*, *3*(4), 526-542.

Shang, C., Sassa, H., & Hirano, H. (2005). The role of glycosylation in the function of a 48-kDa glycoprotein from carrot. *Biochemical and biophysical research communications*, 328(1), 144-149.

Shang, C., Shibahara, T., Hanada, K., Iwafune, Y., & Hirano, H. (2004). Mass spectrometric analysis of posttranslational modifications of a carrot extracellular glycoprotein. *Biochemistry*, 43(20), 6281-6292.

Shibuya, N., & Minami, E. (2001). Oligosaccharide signalling for defence responses in plant. *Physiological and Molecular Plant Pathology*, 59(5), 223-233.

Shoseyov, O., Shani, Z., & Levy, I. (2006). Carbohydrate binding modules: biochemical properties and novel applications. *Microbiology and Molecular Biology Reviews*, *70*(2), 283-295.

Sims, I. M., Gane, A. M., Dunstan, D., Allan, G. C., Boger, D. V., Melton, L. D., & Bacic, A. (1998). Rheological properties of xyloglucans from different plant species. *Carbohydrate polymers*, *37*(1), 61-69.

Sinnott, M. L. (1990). Catalytic mechanism of enzymic glycosyl transfer. *Chemical Reviews*, *90*(7), 1171-1202.

Smith, B. G., & Harris, P. J. (1999). The polysaccharide composition of Poales cell walls: Poaceae cell walls are not unique. *Biochemical Systematics and Ecology*, *27*(1), 33-53.

Subramaniyan, S., & Prema, P. (2002). Biotechnology of microbial xylanases: enzymology, molecular biology, and application. *Critical reviews in biotechnology*, 22(1), 33-64.

Takahashi-Ando, N., Inaba, M., Ohsato, S., Igawa, T., Usami, R., & Kimura, M. (2007). Identification of multiple highly similar XIP-type xylanase inhibitor genes in hexaploid wheat. *Biochemical and biophysical research communications,* 360(4), 880-884.

Takeda, T., Furuta, Y., Awano, T., Mizuno, K., Mitsuishi, Y., & Hayashi, T. (2002). Suppression and acceleration of cell elongation by integration of xyloglucans in pea stem segments. *Proceedings of the National Academy of Sciences*, *99*(13), 9055-9060.

Talbott, L. D., & Ray, P. M. (1992). Molecular size and separability features of pea cell wall polysaccharides Implications for models of primary wall structure. *Plant Physiology*, *98*(1), 357-368.

Thompson, D. S. (2005). How do cell walls regulate plant growth? *Journal of experimental botany*, *56*(419), 2275-2285.

Thompson, J. E., & Fry, S. C. (2001). Restructuring of wall-bound xyloglucan by transglycosylation in living plant cells. *The Plant Journal*, 26(1), 23-34.

Thurow, C., Schiermeyer, A., Krawczyk, S., Butterbrodt, T., Nickolov, K., & Gatz, C. (2005). Tobacco bZIP transcription factor TGA2. 2 and related factor TGA2. 1 have distinct roles in plant defense responses and plant development. *The plant journal*, 44(1), 100-113.

Tong, C. B., Labavitch, J. M., & Yang, S. F. (1986). The induction of ethylene production from pear cell culture by cell wall fragments. *Plant physiology*,*81*(3), 929-930.

Törrönen, A., Harkki, A., & Rouvinen, J. (1994). Three-dimensional structure of endo-1, 4-beta-xylanase II from Trichoderma reesei: two conformational states in the active site. *The EMBO journal*, 13(11), 2493.

Törrönen, A., & Rouvinen, J. (1997). Structural and functional properties of low molecular weight endo-1, 4-β-xylanases. *Journal of biotechnology*, 57(1), 137-149.

Towbin, H., Staehelin, T., & Gordon, J. (1979). Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets: procedure and some applications. *Proceedings of the National Academy of Sciences*, 76(9), 4350-4354.

Tuomivaara, S. T., Yaoi, K., O'Neill, M. A., & York, W. S. (2015). Generation and structural validation of a library of diverse xyloglucan-derived oligosaccharides, including an update on xyloglucan nomenclature. *Carbohydrate research*, *402*, 56-66.

Urbanowicz, B. R., Peña, M. J., Ratnaparkhe, S., Avci, U., Backe, J., Steet, H. F., ... & Darvill, A. G. (2012). 4-O-methylation of glucuronic acid in Arabidopsis glucuronoxylan is catalyzed by a domain of unknown function family 579 protein. *Proceedings of the National Academy of Sciences*,*109*(35), 14253-14258.

Valadon, P. (2004). RasTop. Version 2.0. 3. Philippe Valadon, PO Box 928064, La Jolla, CA 92192-8064, USA.

Vandermarliere, E., Lammens, W., Schoepe, J., Rombouts, S., Fierens, E., Gebruers, K., ... & Courtin, C. M. (2010). Crystal structure of the noncompetitive xylanase inhibitor TLXI, member of the small thaumatin-like protein family. *Proteins: Structure, Function, and Bioinformatics*, 78(10), 2391-2394.

Vega-Sánchez, M. E., Verhertbruggen, Y., Christensen, U., Chen, X., Sharma, V., Varanasi, P., ... & Singh, S. (2012). Loss of cellulose synthase-like F6 function affects mixed-linkage glucan deposition, cell wall mechanical properties, and defense responses in vegetative tissues of rice. *Plant physiology*, *159*(1), 56-69.

Vincken, J. P., York, W. S., Beldman, G., & Voragen, A. G. (1997). Two general branching patterns of xyloglucan, XXXG and XXGG. *Plant Physiology*, *114*(1), 9.

Varki, A., Cummings, R. D., Esko, J. D., Freeze, H. H., Stanley, P., Marth, J. D., ... & Etzler, M. E. (2009). Symbol nomenclature for glycan representation. *Proteomics*, 9(24), 5398-5399.

Untergasser, A., Cutcutache, I., Koressaar, T., Ye, J., Faircloth, B. C., Remm, M., & Rozen, S. G. (2012). Primer3—new capabilities and interfaces. *Nucleic acids research*, 40(15), e115-e115.

Walton, J. D. (1994). Deconstructing the cell wall. *Plant Physiology*, 104(4), 1113.

Wanjiru, W. M., Zhensheng, K., & Buchenauer, H. (2002). Importance of cell wall degrading enzymes produced by Fusarium graminearum during infection of wheat heads. *European Journal of Plant Pathology*, 108(8), 803-810.

Weng, X. Y., Huang, Y. Y., Gao, H., & Sun, J. Y. (2010). Characterization of a xylanase inhibitor TAXI-I from wheat. *Biologia Plantarum*, 54(1), 154-158.

Wu, A. M., Hörnblad, E., Voxeur, A., Gerber, L., Rihouey, C., Lerouge, P., & Marchant, A. (2010). Analysis of the Arabidopsis IRX9/IRX9-L and IRX14/IRX14-L pairs of glycosyltransferase genes reveals critical contributions to biosynthesis of the hemicellulose glucuronoxylan. *Plant physiology*, *153*(2), 542-554.

Wu, A. M., Rihouey, C., Seveno, M., Hörnblad, E., Singh, S. K., Matsunaga, T., ... & Marchant, A. (2009). The Arabidopsis IRX10 and IRX10-LIKE glycosyltransferases are critical for glucuronoxylan biosynthesis during secondary cell wall formation. *The Plant Journal*, *57*(4), 718-731.

Xie, W., Hao, L., & Goodwin, P. H. (2008). Role of a xyloglucan-specific endo-β-1, 4-glucanase inhibitor in the interactions of *Nicotiana benthamiana* with Colletotrichum destructivum, C. orbiculare or Pseudomonas syringae pv. tabaci. *Molecular plant pathology*, 9(2), 191-202.

Yang, Y., & Klessig, D. F. (1996). Isolation and characterization of a tobacco mosaic virus-inducible myb oncogene homolog from tobacco. *Proceedings of the National Academy of Sciences*, 93(25), 14972-14977.

York, W. S., van Halbeek, H., Darvill, A. G., & Albersheim, P. (1990). Structural analysis of xyloglucan oligosaccharides by 1 Hn. mr spectroscopy and fast-atom-bombardment mass spectrometry. *Carbohydrate Research*, *200*, 9-31.

Yoshizawa, T., Shimizu, T., Hirano, H., Sato, M., & Hashimoto, H. (2012). Structural basis for inhibition of xyloglucan-specific endo-β-1, 4-glucanase (XEG) by XEG-protein inhibitor. *Journal of Biological Chemistry*, 287(22), 18710-18716.

Yoshizawa, T., Shimizu, T., Hirano, H., Sato, M., & Hashimoto, H. (2011). Purification, crystallization and X-ray diffraction study of extracellular dermal glycoprotein from carrot and the inhibition complex that it forms with an endo-β-glucanase from Aspergillus aculeatus. *Acta Crystallographica Section F: Structural Biology and Crystallization Communications*, 67(7), 830-832.

Yoshizawa, T., Shimizu, T., Yamabe, M., Taichi, M., Nishiuchi, Y., Shichijo, N., ... & Hashimoto, H. (2011). Crystal structure of basic 7S globulin, a xyloglucan-specific endo-β-1, 4-glucanase inhibitor protein-like protein from soybean lacking inhibitory activity against endo-β-glucanase. *FEBS journal*, 278(11), 1944-1954.

Yuan, S., Wu, Y., & Cosgrove, D. J. (2001). A fungal endoglucanase with plant cell wall extension activity. *Plant Physiology*, 127(1), 324-333.

Zablackis, E., Huang, J., Muller, B., Darvill, A. G., & Albersheim, P. (1995). Characterization of the cell-wall polysaccharides of Arabidopsis thaliana leaves. *Plant Physiology*, *107*(4), 1129-1138.

Zandleven, J., Beldman, G., Bosveld, M., Schols, H. A., & Voragen, A. G. J. (2006). Enzymatic degradation studies of xylogalacturonans from apple and potato, using xylogalacturonan hydrolase. *Carbohydrate polymers*, *65*(4), 495-503.

Zandleven, J., Sørensen, S. O., Harholt, J., Beldman, G., Schols, H. A., Scheller, H. V., & Voragen, A. J. (2007). Xylogalacturonan exists in cell walls from various tissues of Arabidopsis thaliana. *Phytochemistry*, *68*(8), 1219-1226.

Zhao, Y., Thilmony, R., Bender, C. L., Schaller, A., He, S. Y., & Howe, G. A. (2003). Virulence systems of Pseudomonas syringae pv. tomato promote bacterial speck disease in tomato by targeting the jasmonate signaling pathway. *The Plant Journal*, 36(4), 485-499.

SUMMARY

Plant cell walls are composed mostly of polysaccharides and it consist of three layers (the primary cell wall, the secondary cell wall and the middle lamella) that are made up of different percentage of cellulose, pectins and hemicelluloses. These latter are composed of a linear backbone made up of (1,4)-β-D-glycans with an equatorial configuration. Based on type of glycans forming the backbone it is possible to distinguish: mannans contain β-(1,4)-linked mannose; in xyloglucan β-1,4 glucans can be substituted with a diverse array of glycosyl and nonglycosyl residues and xylans are composed by β-(1,4)-linked xylose residues. The seeds of many legumes are known to accumulate galactomannan in their endospermic cell walls. In many dicots xyloglucans constitute the major hemicellulose of growing cell walls, comprising ~20% of the dry mass of primary cell walls. Grasses - but not monocots in general - have a reduced xyloglucan content. Cell wall polysaccharide biogenesis includes polymer synthesis, secretion, assembly, and rearrangement during development. All of these modification demands the reversible 'loosening' of the cellulose–hemicellulose–pectin network. Glycoside hydrolase (GH) enzymes located in the wall or in the plasma membrane play a crucial role in the degradation of different cell wall polysaccharides. On the other hand, pathogenic microorganisms secrete glycoside hydrolase to penetrate plant cell walls. As a response, plants produce glycoside hydrolase inhibitor proteins (GHIPs). Xyloglucan-specific endo-β-1,4-glucanase inhibitor proteins-like (XEGIPs-like) are typical of dicots, they inhibit the hydrolytic activity of a xyloglucan-specific β-1,4-endoglucanase isolated from GH12 family. XEGIPs-like have been found widespread in dicots: they were detected in the medium of cultured tomato cells, purified from carrot callus, isolated from the nectar of ornamental tobacco, when overexpressed they were capable of protecting potato from disease caused by endo-β-1,4-glucanase GH12 from *Phytophthora infestans,* enhanced in apple in response to infection of *Botryosphaeria dothidea,* they have distinct roles in defence mechanisms in *Humulus lupus.* In cereals three types of GHIPs occur in a fairly coordinated fashion throughout grain development and germination: *Triticum aestivum* L. endoxylanase inhibitors (TAXIs-like), xylanase inhibitor proteins (XIPs-like), thaumatin-like xylanase inhibitors (TLXIs-like). The accumulation of GHIPs during the early stages of germination is consistent with the phenomenon of germination-based resistance and their highest concentrations occur in the aleuronic layer. The apoplastic localization of GHIPs in cereals may be favourable for their action as inhibitors of microbial xylanases GH10 and/or GH11 from *Aspergillus niger*, *Bacillus subtilis* and *Hypocrea jecorina* intruding the host plant. GHIPs homologous are also present in legume (LACGs-like): γ-conglutin is largely expressed and accumulated in *Lupinus* spp. and Bg7S in *Glycine* spp.

GHIPs have common structural features. In particular, the alignment of the primary structurer showed that the position of the 12 cysteines is fully conserved, so various GHIPs have similar three-dimensional structures. Cys10-Cys11 is located at the C-terminal region of the proteins, the sequence amongst them is called "inhibition loop 1" and a conserved arginine, in XEGIPs-like, or leucine, in TAXI-I, is involved in the bond GH12 or GH11, respectively. In the LAGCs-like it's also present the loop, but a deletion of about five amino acids involve this region, otherwise highly conserved. The superimposition of TAXI-I 3D structure (PDB accession number: 1T6E) and EDGP 3D structure (PDB accession number: 3VLA) with γ-conglutin 3D structure (PDB accession number:

4PPH) confirm the structural difference of the loop responsible for the interaction with the enzyme as a consequence of the five amino acid long deletion. This may be the cause of an unfavourable local spatial conformation of the protein for the correct interaction with the enzyme. The disulphide bridge Cys9-Cys12 define another region called “inhibition loop 2” where a conserved arginine, in XEGIPs-like, or histidine, in TAXIs-like and LACGs-like, is involved in the bond with a specific GH. The sequence of the IL2 loop of γ-conglutin is more similar to the sequence of the IL2 loop of TAXI-I, rather than to the one of XEGIPs. In particular, a His residue, considered a key amino acid for the inhibitory activity of TAXI-I is also present in γ-conglutin sequence, but not in XEGIPs. γ-conglutin was expressed in *Pichia pastoris*. Unexpectedly, this recombinant γ-conglutin (rγc) was able to inhibit a GH11 enzyme, but not GH12. In lupin, γ-conglutin is naturally cleaved in two subunits, whereas in *P. pastoris* it is not. Most likely, the proteolytic processing influences the structural conformation of γ-conglutin and small local rearrangements may be the cause of the observed activity. Also a set of γ-conglutin mutants was designed upon TAXIs-like and XEGIPs-like sequences and expressed in *Pichia pastoris*. The mutants were able to modulate the inhibition capacity. The enzymatic assays and the bioinformatics analysis confirmed that the presence of IL1 is not strictly required to manifest inhibition, even if the specifically inserted amino acid stretches enhanced the activity. On the other hand, histidine in IL2 is confirmed to be necessary and sufficient to manifest the inhibitory competence of rγc.

The LACGs-like among GHIPs remains the less characterized. For this reason, we undertook experiments aimed to study the inhibitory specificity of wild type γ-conglutin. Inhibitory capacity was tested against under different condition using selected GH2 and GH5 members. GH11 and GH12 were not tested since previous results evidenced no capacity.

In *Arabidopsis thaliana*, which is a small flowering plant widely used as a model organism in plant biology, we have found five genes coding for GHIPs belonging to each of three aforementioned groups. At1g03220 and At1g03230 are closer related to XEGIPs-like proteins: the cysteines pathway is totally conserved, they show arginine in the inhibition loop 1 and inhibition loop 2, which interact with glycoside hydrolases family 12, and along the sequences there are some potential glycosylation sites. At5g19110 is the unique putative proteins that shows leucine and histidine in the inhibition loop 1 and inhibition loop 2, respectively, the same region used by TAXIs-like to interact with glycoside hydrolases family GH11. At5g19100 and At5g19120 show the characteristic LACGs-like deletion at the IL1 and only in At5g19100 sequence at the IL2 we recover the histidine like both LACGs-like and TAXIs-like. We have studied the expression of these genes to deep into the biology of the plant response to pathogen attack. The results of this set of experiment contributed to unveil the role of each.

The seeds were germinated in Petri dishes with MS medium and held in growth chambers with constant temperature of 21°C, under long-day photoperiod, for different times (0-23-48-96-144 hours) and in different conditions simulating biotic stresses. Western-blot with antibody anti-Cγ show the presence of a similar γ-conglutin protein in *A. thaliana* seeds at 96h of germination in MS medium and then incubate in a chitosan (150 mg/L) solution for 3h. For the first time we show that proteolytic processing of LACGs-like occurs in organism other than legume. In parallel, total RNAs were extracted and RT-qPCR has been set up to quantify the relative expression levels of gene expression. The GHIPs homologues genes after 48, 96 and 144 hours of germination aren’t expressed under basal condition. Seeds germinated for 96 hours and then exposed to incubation

with aqueous solution or 150mg/l chitosan solution (mimic a biotic stress) shown an expression of gene, greater if the biotic stress was applied at seeds contaminated. In the seeds germinated in dishes that after 96 hours shown a spontaneous growth of *Paenibacillus polymyxa*, an endophytic bacterium exploited as biocontrolling agent. In this case the expression of all analysed genes increased. In another experimental set we tested the direct effect of GH2, GH5, GH11 treatment on *A. thaliana* seedling 7 days old. In this contest, it has been evaluated both the expression of the five selected genes and the effects on the enzymes due to plants exudates.

By and large, this work describes new experimental findings that open new interesting scenarios to better understand some physiological aspects of the plant defence mechanisms, and provides insight of the structural basis of GHIPs inhibitory activity, specificity and repercussions on cellular responses to pathogens attacks.

Appendix_Sequence

R7W0U4 – *Aegilops tauschii*

```
        10         20         30         40         50
MARLPVLVLA VSLAVLAWPA SCKSVRSVLA PVTKDPATRL YTIPFHYGAN
        60         70         80         90        100
IVVDTAGPLV WSTCAPDHLP AAFPCKSDTC RLANKYHVPS CSESAADKLC
       110        120        130        140        150
DPSHKVCRAF PYNPVTGACA AGDLIHTRFV ANTTDGKNPV SQVNVRAVAA
       160        170        180        190        200
CAPSKLLESL PQGASGVAGL AGSDLALPAQ VASAQKVSNK FLLCLPRGLS
       210        220        230        240        250
SDPGVAVFGG GPLHFMAQPE RDYTKELAYT PLVAKKGNPA HYITIKSIAV
       260        270        280        290        300
ESASVPVPAQ ALATGGAVLC TRSPFTLLRS DVFLPLVDAF TKALAGQGAQ
       310        320        330        340        350
GGPVAKAVKP YAPFQLCYDT RTLANTRTGY LVPAVTLTLG GGKNWRMDGL
       360        370        380        390        400
SLMVDMGPTT ACLAFVQMQG VKGGDGSAPS VLIGGFQMEN TVLEFDMKKK
       410        420
RLGFARLPSF TQCGQFNFTT RSA
```

Q9ZVS4 – *Arabidopsis thaliana 03220*

```
        10         20         30         40         50
MAPSPIIFSV LLLFIFSLSS SAQTPFRPKA LLLPVTKDQS TLQYTTVINQ
        60         70         80         90        100
RTPLVPASVV FDLGGRELWV DCDKGYVSST YQSPRCNSAV CSRAGSTSCG
       110        120        130        140        150
TCFSPPRPGC SNNTCGGIPD NTVTGTATSG EFALDVVSIQ STNGSNPGRV
       160        170        180        190        200
VKIPNLIFDC GATFLLKGLA KGTVGMAGMG RHNIGLPSQF AAAFSFHRKF
       210        220        230        240        250
AVCLTSGKGV AFFGNGPYVF LPGIQISSLQ TTPLLINPVS TASAFSQGEK
       260        270        280        290        300
SSEYFIGVTA IQIVEKTVPI NPTLLKINAS TGIGGTKISS VNPYTVLESS
       310        320        330        340        350
IYNAFTSEFV KQAAARSIKR VASVKPFGAC FSTKNVGVTR LGYAVPEIEL
       360        370        380        390        400
VLHSKDVVWR IFGANSMVSV SDDVICLGFV DGGVNARTSV VIGGFQLEDN
       410        420        430
LIEFDLASNK FGFSSTLLGR QTNCANFNFT STA
```

Q9ZVS5 – *Arabidopsis thaliana 03230*

```
        10         20         30         40         50
MASSRIIIFS VLLLSIFSLS SSAQPSFRPK ALLLPVTKDP STLQYTTVIN
        60         70         80         90        100
QRTPLVPASV VFDLGGREFW VDCDQGYVST TYRSPRCNSA VCSRAGSIAC
       110        120        130        140        150
GTCFSPPRPG CSNNTCGAFP DNSITGWATS GEFALDVVSI QSTNGSNPGR
       160        170        180        190        200
FVKIPNLIFS CGSTSLLKGL AKGAVGMAGM GRHNIGLPLQ FAAAFSFNRK
       210        220        230        240        250
FAVCLTSGRG VAFFGNGPYV FLPGIQISRL QKTPLLINPG TTVFEFSKGE
       260        270        280        290        300
KSPEYFIGVT AIKIVEKTLP IDPTLLKINA STGIGGTKIS SVNPYTVLES
       310        320        330        340        350
SIYKAFTSEF IRQAAARSIK RVASVKPFGA CFSTKNVGVT RLGYAVPEIQ
       360        370        380        390        400
LVLHSKDVVW RIFGANSMVS VSDDVICLGF VDGGVNPGAS VVIGGFQLED
       410        420        430
NLIEFDLASN KFGFSSTLLG RQTNCANFNF TSTA
```

Q3E9C8 – *Arabidopsis thaliana 19100*

```
        10         20         30         40         50
MAPRVIFLLL SLVFLYLANT SHSLRKFQSF LHPIYKDTAK NIYTIPLSIG
        60         70         80         90        100
```

```
STSSEKFVLD LNGAAPLLQN CPTAAKSTTY HPIRCGSTRC KYANPNFPCP
       110         120         130         140         150
NNVIAKKRTV CLSSDNSRLF RDTVPLLYTF NGVYTRDSEM SSSLTLTCTD
       160         170         180         190         200
GAPALKQRTI GLANTHLSIP SQLISMYQLP HKIALCLPST ERSQSHNGDL
       210         220         230         240         250
WIGKGEYYYL PYDKDVSKIF ASTPLIGNGK SGEYLIDVKS IQIGAKTVPI
       260         270         280         290         300
PYGATKISTL APYTVFQTSL YKALLTAFTE NIKIAKAPAV KPFGACFYSN
       310         320         330         340         350
GGRGVPVIDL VLSGGAKWRI YGSNSLVKVN KNVVCLGFVD GGVKPKYPIV
       360         370         380         390
IGGFQMEDNL VEFDLEASKF SFSSSLLLHN TSCSVQRLSP F
```

F4JZM1 – *Arabidopsis thaliana 19110*

```
        10          20          30          40          50
MGSSLTRLLV FLSIFAAIAL KSNSQYLLPI TKHEPTNLFY TTFNVGSAAK
        60          70          80          90         100
SPVNLLLDLG TNLTWLDCRK LKSLSSLRLV TCQSSTCKSI PGNGCAGKSC
       110         120         130         140         150
LYKQPNPLGQ NPVVTGRVVQ DRASLYTTDG GKFLSQVSVR HFTFSCAGEK
       160         170         180         190         200
ALQGLPPPVD GVLALSPGSS SFTKQVTSAF NVIPKFSLCL PSSGTGHFYI
       210         220         230         240         250
AGIHYFIPPF NSSDNPIPRT LTPIKGTDSG DYLITVKSIY VGGTALKLNP
       260         270         280         290         300
DLLTGGAKLS TVVHYTVLQT DIYNALAQSF TLKAKAMGIA KVPSVAPFKH
       310         320         330         340         350
CFDSRTAGKN LTAGPNVPVI EIGLPGRIGE VKWGFYGANT VVKVKETVMC
       360         370         380         390         400
LAFIDGGKTP KDLMVIGTHQ LQDHMLEFDF SGTVLAFSES LLLHNTSCST

WPSQK
```

Q0WU92 – *Arabidopsis thaliana 19120*

```
        10          20          30          40          50
MASSSCLNLF FFSFLSALII SKSQISDSVN GVVFPVVKDL PTGQYLAQIR
        60          70          80          90         100
LGDSPDPVKL VVDLAGSILW FDCSSRHVSS SRNLISGSSS GCLKAKVGNE
       110         120         130         140         150
RVSSSSSSRK DQNADCELLV KNDAFGITAR GELFSDVMSV GSVTSPGTVD
       160         170         180         190         200
LLFACTPPWL LRGLASGAQG VMGLGRAQIS LPSQLAAETN ERRRLTVYLS
       210         220         230         240         250
PLNGVVSTSS VEEVFGVAAS RSLVYTPLLT GSSGNYVINV KSIRVNGEKL
       260         270         280         290         300
SVEGPLAVEL STVVPYTILE SSIYKVFAEA YAKAAGEATS VPPVAPFGLC
       310         320         330         340         350
FTSDVDFPAV DLALQSEMVR WRIHGKNLMV DVGGGVRCSG IVGGGSSRVN
       360         370         380
PIVMGGLQLE GFILDFDLGN SMMGFGQRTR SDSTSL
```

Q9LV70 – *Arabidopsis thaliana* At5g48430

```
        10          20          30          40          50
MEKSLLVLCL ILFFTYSYVS ANYYPPKALV STVSKNTILP IFTFTLNTNQ
        60          70          80          90         100
EFFIHIGGPY LVRKCNDGLP RPIVPCGSPV CALTRRFTPH QCSLPSNKII
       110         120         130         140         150
NGVCACQATA FEPFQRICNS DQFTYGDLSI SSLKPISPSV TINNVYYLCI
       160         170         180         190         200
PQPFLVDFPP GVFGLAGLAP TALATWNQLT RPRLGLEKKF ALCLPSDENP
       210         220         230         240         250
LKKGAIYFGG GPYKLRNIDA RSMLSYTRLI TNPRKLNNYF LGLKGISVNG
       260         270         280         290         300
NRILFAPNAF AFDRNGDGGV TLSTIFPFTM LRSDIYRVFI EAFSQATSGI
       310         320         330         340         350
PRVSSTTPFE FCLSTTTNFQ VPRIDLELAN GVIWKLSPAN AMKKVSDDVA
       360         370         380         390         400
CLAFVNGGDA AAQAVMIGIH QMENTLVEFD VGRSAFGFSS SLGLVSASCG
```

DFQTRP

P13917 – *Glycine max*

```
        10         20         30         40         50
MASILHYFLA LSLSCSFLFF LSDSVTPTKP INLVVLPVQN DGSTGLHWAN
        60         70         80         90        100
LQKRTPLMQV PVLVDLNGNH LWVNCEQQYS SKTYQAPFCH STQCSRANTH
       110        120        130        140        150
QCLSCPAASR PGCHKNTCGL MSTNPITQQT GLGELGEDVL AIHATQGSTQ
       160        170        180        190        200
QLGPLVTVPQ FLFSCAPSFL VQKGLPRNTQ GVAGLGHAPI SLPNQLASHF
       210        220        230        240        250
GLQRQFTTCL SRYPTSKGAI IFGDAPNNMR QFQNQDIFHD LAFTPLTITL
       260        270        280        290        300
QGEYNVRVNS IRINQHSVFP LNKISSTIVG STSGGTMIST STPHMVLQQS
       310        320        330        340        350
VYQAFTQVFA QQLPKQAQVK SVAPFGLCFN SNKINAYPSV DLVMDKPNGP
       360        370        380        390        400
VWRISGEDLM VQAQPGVTCL GVMNGGMQPR AEITLGARQL EENLVVFDLA
       410        420
RSRVGFSTSS LHSHGVKCAD LFNFANA
```

B9VUU9 – *Caspicum annuum*

```
        10         20         30         40         50
MASSSCFHVI LFCSFLFFTS TIAQNQTSFR PKGLIIPVMK DGSTLQYLTQ
        60         70         80         90        100
IQQRTPLVPV SLTLDLGGQF LWVDCDQGYV SSSYKPARCR SAQCSLAGAT
       110        120        130        140        150
GCGECFSPPR PGCNNNTCGL FPDNTVTRTA TSGELASDVV SVQSSNGKNP
       160        170        180        190        200
GRNVSDKNFL FVCGATFLLQ GLASGVKGMA GLGRTRISLP SQFSAEFSFP
       210        220        230        240        250
RKFAVCLSSS KSKGVVLFGD GPYFFLPNTE FSNNDFQYTP LLINPVSTAS
       260        270        280        290        300
AFSAGQPSSE YFIGVKSVKI NQKVVPINTT LLSIDNQGVG GTKISTVNPY
       310        320        330        340        350
TVLETSLYNA ITNFFVKELA NVTRVASVAP FGACFDSRNI GSTRVGPAVP
       360        370        380        390        400
QIDLVLQNEN VIWTIFGANS MVQVSENVLC LGFVDGGVNS RTSIVIGGHT
       410        420        430
IEDNLLQLDI ARSRLGFTSS ILFRQTTCAN FNFTSIA
```

K4FKJ9 – *Caspicum annuum*

```
        10         20         30         40         50
MVTFKLPLSV LLLFLQLNIF LCSAEVLYIP VTKDTTTLRY ITEVGQRTPL
        60         70         80         90        100
VPIKLLVHLG GRSLLVDCDK GYKSSTYKSA VCNSTQCSFA KSHACGDCIF
       110        120        130        140        150
KSQLQPGCNN NTCYIWGENP LINSFHDRAE IAEDVLTIGS TPGVHVTWSR
       160        170        180        190        200
FIFTCLLDQD MMRLLAKGVT GIAGFGRESP ISLPNQLALD PRFTRKFGLC
       210        220        230        240        250
LSSSTRSRGV IFIGSGPYNI YNPKKIDISK DLVYTKLIAN KRGGFVASEE
       260        270        280        290        300
YYIQVSSIRV AGKDVPLNKT LLSINKKNGV GGTRISTATP FTILHTSIYD
       310        320        330        340        350
AFKTAFIKAL PKNVTLVDPP IKQFGVCFSS KNIKSTNTGP DLPVIDVVLH
       360        370        380        390        400
KPSAFWRIYG TNSVVQVNKD VMCLAFVGQD QTWEPSIVIG GHQMEENLLV
       410        420        430
FDLPGKNIGF SSSLKLQQTS CSKYDNTTLG
```

Q9FSH9 – *Lupinus albus*

```
        10         20         30         40         50
MAKNMAPILH ILVISLSYSF LFVTSSSQNS QSLYHNSQPT SSSKPNLLVL
        60         70         80         90        100
PIQQDASTKL HWGNILKRTP LMQVPVLLDL NGKHLWVTCS QHYSSSTYQA
       110        120        130        140        150
PFCHSTQCSR ANTHQCFTCT DSTTSRPGCH NNTCGLISSN PVTQESGLGE
       160        170        180        190        200
LAQDVLALHS THGSKLGSLV KIPQFLFSCA PTFLTQKGLP NNVQGALGLG
       210        220        230        240        250
HAPISLPNQL FSHFGLKRQF TMCLSSYPTS NGAILFGDIN DPNNNNYIHN
       260        270        280        290        300
SLDVLHDMVY TPLTISKQGE YFIQVSAIRV NKHMVIPTKN PSMFPSSSSS
       310        320        330        340        350
SYHESSEIGG AMITTTNPYT VLRHSIFEVF TQVFANNVPK QAQVKAVGPF
       360        370        380        390        400
GLCYDTKKIS GGVPSVDLIM DKSDVVWRIS GENLMVQAQD GVSCLGFVDG
       410        420        430        440        450
GVHTRAGIAL GTHQLEENLV VFDLARSRVG FNTNSLKSHG KSCSNLFDLN

NP
```

Q42369 – *Lupinus angustifolius*

```
        10         20         30         40         50
MARNMAHILH ILVISLSYSF LFVSSSSQDS QSLYHNSQPT SSKPNLLVLP
        60         70         80         90        100
VQEDASTGLH WANIHKRTPL MQVPLLLDLN GKHLWVTCSQ HYSSSTYQAP
       110        120        130        140        150
FCHSTQCSRA NTHQCFTCTD STTTRPGCHN NTCGLLSSNP VTQESGLGEL
       160        170        180        190        200
AQDVLAIHST HGSKLGPMVK VPQFLFSCAP SFLAQKGLPN NVQGALGLGQ
       210        220        230        240        250
APISLQNQLF SHFGLKRQFS VCLSRYSTSN GAILFGDIND PNNNNYIHNS
       260        270        280        290        300
LDVLHDLVYT PLTISKQGEY FIQVNAIRVN KHLVIPTKNP FISPSSTSYH
       310        320        330        340        350
GSGEIGGALI TTTHPYTVLS HSIFEVFTQV FANNMPKQAQ VKAVGPFGLC
       360        370        380        390        400
YDSRKISGGA PSVDLILDKN DAVWRISSEN FMVQAQDGVS CLGFVDGGVH
       410        420        430        440
ARAGIALGAH HLEENLVVFD LERSRVGFNS NSLKSYGKTC SNLFDLNNP
```

Q05929 – *Daucus carota*

```
        10         20         30         40         50
ATSLQITLFS LLFIFTITQA QPSFRPSALV VPVKKDASTL QYVTTINQRT
        60         70         80         90        100
PLVSENLVVD LGGRFLWVDC DQNYVSSTYR PVRCRTSQCS LSGSIACGDC
       110        120        130        140        150
FNGPRPGCNN NTCGVFPENP VINTATGGEV AEDVVSVEST DGSSSGRVVT
       160        170        180        190        200
VPRFIFSCAP TSLLQNLASG VVGMAGLGRT RIALPSQFAS AFSFKRKFAM
       210        220        230        240        250
CLSGSTSSNS VIIFGNDPYT FLPNIIVSDK TLTYTPLLTN PVSTSATSTQ
       260        270        280        290        300
GEPSVEYFIG VKSIKINSKI VALNTSLLSI SSAGLGGTKI STINPYTVLE
       310        320        330        340        350
TSIYKAVTEA FIKESAARNI TRVASVAPFG ACFSTDNILS TRLGPSVPSI
       360        370        380        390        400
DLVLQSESVV WTITGSNSMV YINDNVVCLG VVDGGSNLRT SIVIGGHQLE
       410        420        430
DNLVQFDLAT SRVGFSGTLL GSRTTCANFN FTS
```

K9JA99 – *Humulus lupus 2*

```
        10         20         30         40         50
MASFTHFVLF CSLLFPILIT PTIAETPSFR PKALLLPVTK DASTKQYLTQ
        60         70         80         90        100
INQRTPLVPV KLTVNLGGEF LWVDCEKGYV SSTYKPARCR SAQCNLAGSK
       110        120        130        140        150
SCGECFDGPK PGCNNNTCGL FPYNPFIRTS TSGELAQDII SIQSTNGSNP
       160        170        180        190        200
SKVVSFPNVI FTCGSTFLLE GLASGVTGIA GLGRKKIALP SQFAAAFSFK
```

```
       210        220        230        240        250
RKFALCLSSS TRATGVVFFG DGPYIMLPNK DVSQNLIYTP LILNPVSTAG
       260        270        280        290        300
ASFEGEPSAD YFIGVKGIKV NGEDVKLNTS LLSIAKDGTG GTKISTTQPY
       310        320        330        340        350
TSLETSIYKA VIGAFGKAVA KVPRVTAVAP FELCFNSTSF SSTRVGPGVP
       360        370        380        390        400
QIDLVLPNNK AWTIFGANSM VQVSDDVLCL GFVDGGPLHF VDWGIPFTPT
       410        420        430        440
AIVIGGHQIE DNLLQFDLGS STLGFSSSLL FRQTTCSNFN FTSIA
```

K9JA25 – *Humulus lupus 1*

```
       10         20         30         40         50
MASSFSFKLF FFLFLFSALS SHLATAKTAA FPKALVLPVT KDTTTRQYIT
        60         70         80         90        100
QITQRTPPVQ LKVVLDVGGE FLWIDCEKGY KSSTKRPVPC GSPQCVLSGS
       110        120        130        140        150
GACTTSDNPS DVGVCGVMPN NPFSSVGTSG DLFEDILYIQ STNGFNPGKQ
       160        170        180        190        200
VSVPNLLFSC APNSLLEGLA SGIIGMAGFG RNKVALPSLF SSAFSFPRKF
       210        220        230        240        250
GVCLSSSNGV IFFGKEPYVL LPGIDVSDPT SLTYTPLIQN PRSLVSSFEG
       260        270        280        290        300
NPSAEYFIGV KSIKVDGKPL RLNTTLLTFD NEGGHGGTKI STVDPFTTLE
       310        320        330        340        350
TSIYKAVVGA FVKALGPKVP RVKAVAPFGA CFNAKYIGNT RVGPAVPQID
       360        370        380        390        400
LVLRNDKLWS IFGANSMVSV GDDVLCLGFV DGGPLNFVDW GVKFTPTAVV
       410        420        430        440
IGGHQIENNF LLFDLGASRL GFSSSLLFRQ TTCSNFNFNS STY
```

K9JA06 – *Humulus lupus 3*

```
        10         20         30         40         50
MSSNSFHHLL FCSLLLLIIS PSISQTISFR PKALVLQVTK DSATHQYYTH
        60         70         80         90        100
ITQRTPPVQV KVAIDLGGEF LWVDCEKGFN SSTKKPVPCR SAQCNLAKSK
       110        120        130        140        150
ACSTNGNPSE DVCGEFPHNP FISTSTSGDL SQDIIYIQST NGSRPGKVVS
       160        170        180        190        200
VPKFIFTCAP TFLLKGLTSG AVGVAGLGRN KIALPSLFSA AFSFPKKMAV
       210        220        230        240        250
CLSSTNGVVF FGNGPYELSS GIDVSKSLTY TPLILNPVNL IGGFQGESSS
       260        270        280        290        300
EYFIGVKSIK VDGKPVSVNS SLLSFDVDGN GGTKISTVDP YTTLETSIYN
       310        320        330        340        350
TVVNAFVNAL AVRNVHKVAA VAPFSACFNA KDIGLSRAGP IVPPIEFVLQ
       360        370        380        390        400
SEKVVWRVTG ANSMVRVSNE VLCLGFVDGG PLHFVDWGIK FTPTAIVIGG
       410        420        430        440
RQIEDNLLQF DLATSRLGFS SSLLSRQLSC SNFKFNRSTV D
```

Q6KE44 – *Horderum vulgare*

```
        10         20         30         40         50
MARVLLLALA ATLAAQASSK ALPVLAPVTK DAATSLYTIP FHDGANLVLD
        60         70         80         90        100
VAGPLVWSTC DGGQRPPPAE ITCSSPTCLL ANAYPAPGCP APSCGSDRHD
       110        120        130        140        150
KPCTAYPSNP VTGACAAGSL FRARLVANIT DGNRPVSAVT VGVLAACAPT
       160        170        180        190        200
KLLASLPRGS TGVAGLAGSG LALPAQVASA QKVSHRFLLC LPTGGAGVAI
       210        220        230        240        250
LGGGPLPWPQ FTQSMAYTPL VAKQGSPAHY VSGTSIRVED TRVPVPDRAL
       260        270        280        290        300
ATGGVMLSTR LPYVLLRRDV YRPFVDAFAK ALAAQHANGA LAARGVNPVA
       310        320        330        340        350
```

```
PFGLCYDAKT LGNNLGGYSV PNVVLALDGG GEWAMTGKNS MVDVKPGTAC
       360        370        380        390        400
VAFVEMEAGD GGAPAVILGG AQMEDFVLDF DMEKKRLGFI RLPHFTGCGN

LNF
```

Q2HTN2 – *Medicago truncatula*

```
        10         20         30         40         50
MASISILHFL LISLFCSFLL VSSRHQQQPN SNPKPNLLVL PVQQDASTGL
        60         70         80         90        100
HWANIHKRTP LMQVPVLLDL NGKHLWVNCE QHYASSTYQA PYCHSTQCSR
       110        120        130        140        150
ANAHTCHTCV SSFRPGCHNN TCGLMSANPV TQQTAMGELA QDVLAIYAIN
       160        170        180        190        200
GPKPGPMVTI PQFLFSCAPS FLAQKGLPNN VQGVVGLAHS PISLQNQLSS
       210        220        230        240        250
HFGLKRQFTM CLSRHPNSNG AILFGDAPNN MHFGQGNNYN NKNNPNLFNN
       260        270        280        290        300
LVYTPLTITQ QGEYRIHVTS IRLNQHTVVP VSAPMLSSYP EGVMGGTLIS
       310        320        330        340        350
TSIPYTILQH SLFEAFTQVF AKQYPRQAQV NAVGPFGMCF DSKRINQALS
       360        370        380        390        400
VEFVMDRPDV VWRISGENLM VQPRNGVSCL AFVNGGLHPK AAITIGSRQL
       410        420        430
EENMMMFDLA RSRLGFTNSL NSHGMKCSDL FDFTNAP
```

W9QL21 – *Morus notabilis*

```
        10         20         30         40         50
MASFFLHHCP FLLFLLIFFS LCFSVSPSQS AQTGPIFRPN SLALPVRKDP
        60         70         80         90        100
ATGLHVANVS KRTPPLQVPL TIDLNGRFLW ANCEGGSYLS STYNAPLCHS
       110        120        130        140        150
TQCSQAVGPN HYCRTCSSRA RPGCHNGTCG VTVTNPVTGR SAIGELAQDS
       160        170        180        190        200
LSVRSAQGPR QGQGPIARVR QFLFVCAPSA LLQPGLPRKA QGVVGLGHSH
       210        220        230        240        250
VSLPSQLASH FGFQQKFATC LPRGNGNGAV FFGEGPYFFP PGIDVTRKLI
       260        270        280        290        300
YTPLTISQDG EYAINLSGIK INNHPVGTPV SRAIITTTNP YTFLDHSLFV
       310        320        330        340        350
ALTNVFANQL KIPRVQPVAP FGACFDAKGI ASTRIGPAVP PVDLSLHDQS
       360        370        380        390        400
TRWRILGANS MIEARPGVMC LAFVDGGARP HGSSMVIGAY QLEDNLVQFD
       410        420        430
LVKSMLGFSS SLLFRRTSCS NFNFTSSTTS P
```

Q3KU27 – *Nicotiana langsdorffii x Nicotiana sanderae*

```
        10         20         30         40         50
MAYSCLHTIL LCSLLFITST TAQNQTSFRP KGLILPITKD ASTLQYLTQI
        60         70         80         90        100
HQRTHLVPVS LTLDLGGQFL WVDCDQGYVS SSYKPARCRS AQCSLAGAGG
       110        120        130        140        150
CGQCFSPPKP GCNNNTCSLL PDNTITRTAT SGELASDIVQ VQSSNGKNPG
       160        170        180        190        200
RNVTDKDFLF VCGSTFLLEG LASGVKGMAG LGRTRISLPS QFSAEFSFPR
       210        220        230        240        250
KFAVCLSSST NSKGVVLFGD GPYSFLPNRE FSNNDFSYTP LFINPVSTAS
       260        270        280        290        300
AFSSGEPSSE YFIGVKSIKI NQKVVPINTT LLSIDNQGVG GTKISTVNPY
       310        320        330        340        350
TILETSMYNA VTNFFVKELV NITRVASVAP FGACFDSRTI VSTRVGPAVP
       360        370        380        390        400
QIDLVLQNEN VFWTIFGANS MVQVSENVLC LGFVDGGINP RTSIVIGGYT
       410        420        430
IEDNLLQFDL ASSRLGFTSS ILFRQTTCAN FNFTSIA
```

B9VUV0 – *Petunia hybrida*

```
        10         20         30         40         50
MASSCLHAIL LFSLLFISST IVHAQTSFRP KGLILPVTKD ASTLQYLTQI
```

```
        60         70         80         90        100
SQRTPLVPVS LTLDLGGQFL WVDCDQGYVS SSYIPARCRS AKCSLAGSSG
       110        120        130        140        150
CGDCFSPPSP GCNNNTCGAF PDNSITRTAT SGELASDIVS VQSSNGKNPG
       160        170        180        190        200
RNVSDKDFLF VCGATFLLNG LASGVKGMAG LGRTRISLPS QFSAEFSFPR
       210        220        230        240        250
KFAVCLSSTS NSKGVVLFGD GPYSFLPNRE YSSDDFSYTP LFINPVSTAS
       260        270        280        290        300
AFSSGTPSSE YFIGVKSIKI NEKVVPINTT LLSIDSQGVG GTKISTVNPY
       310        320        330        340        350
TILETSIYNA VTNFFVKELA IPTVPSVAPF GVCFDSRNIT STRVGPGVPS
       360        370        380        390        400
IDLVLQNENV FWRIFGANSM VLVSENVLCL GFVDGGVNPR TSIVIGGHTI
       410        420        430
EDNLLQFDLA ASRLGFTSSI LFRQTTCANF NFTSIA
```

V7BPV5 – *Phaseolus vulgaris*

```
        10         20         30         40         50
MASILYFLVF SLSCSFLFLL SESNYVSNPA YLLVLPTQKD VSTGLHWTTL
        60         70         80         90        100
LKRTPLIQVP VLVDLNGNQL WLNCEQHYTS KTYEAPFCHS AQCFRANTHQ
       110        120        130        140        150
CLSCPAAARP GCHKNTCGLM STNPVTQQNG LGELGQDVLA IHVSLGTQLG
       160        170        180        190        200
ELFTVPHFLF SCAPSFLLQK GLPMNVEGVA GLGHGPISLP NQLASHFGLQ
       210        220        230        240        250
RQFTTCLSRH SSSSKGAIIF GDAPNNLHEL HGHAIFQDLA YTPLTITPQG
       260        270        280        290        300
EYNVRVNSIR INQHSVTPVK KTSSTIVGHS GGTMISTSTP HMVLQQSLYE
       310        320        330        340        350
SFIQVFAQQL PTQAHQVKAV APFQLCFHSK NTSEYPGVEL VMDKPNGPVW
       360        370        380        390        400
RISGEALTVQ TQPGVWCLAV VNGGMQPRAE ITIGARQLEE NLVVFDLAKS
       410        420
RVGFGTSPLA SHGMKCADLF NFVDA
```

Q6KE41 – *Secale cereale*

```
        10         20         30         40         50
MPPVLLLVLA ASLVALPSCR SLPVQAPVTK DPATSLYTIP FHDGASLVLD
        60         70         80         90        100
AAGPLVWSTC EAGQPPAGIP CGSPTCLLAN AYPAPGCPAP TCGSDKPCTA
       110        120        130        140        150
FPSNPVTGAC AAGSLFHTSF VANTTDGTKP VSEVKVGVLA ACAPSKLLAS
       160        170        180        190        200
LPRGSTGVAG LANSGLALPA QVASAQKVAN RFFLCLPTGG AGVAIFGGGP
       210        220        230        240        250
LPWPQFTQSM PYTPLVTKGG SPAHYISLKS IKVDNTRVPV SEATGGVMLS
       260        270        280        290        300
TRLPYALLRR DVYRPLVDAF TKALAAQPAN GAPVARAVQP VAPFGVCYDT
       310        320        330        340        350
KTLGNNLGGY AVPNVLLALD GGGEWAMTGK NSMVDVKPGT ACVAFVEMKG
       360        370        380        390
VEAGDGRAPA VILGGAQMED FVLDFDMEKK RLGFTRLPHF TGCGSA
```

D6QUQ0 – *Solanum melongena*

```
        10         20         30         40         50
MASSCCLHAI LLCCLLFFTS TIAQNQTSFR PKGLIIPVTK DASTLQYLTQ
        60         70         80         90        100
IQQRTPLVPI SLTLDLGGQF LWVDCDQGYV SSSYKPARCR SAQCSLAGAS
       110        120        130        140        150
ACGECFSPPR PGCNNNTCSL FPDNTVTGTA TGGELASDIV SVQSSNGKNP
       160        170        180        190        200
GRNVSDKNFL FVCGATFLLQ GLASGVKGMA GLGRTRISLP SQFSAEFSFP
       210        220        230        240        250
RKFALCLTSS NSKGVVLFGD GPYFFLPNKE FSNNDFQYTP LFINPVSTAA
       260        270        280        290        300
AFSSGQPSSE YFIGVKSIKI NQKVVPINTT LLSIDNQGVG GTKLSTVNPY
       310        320        330        340        350
TVMETSLYNA ITNFFVKELA NVTRVAPVTP FGACFDSRNI GSTRVGPAVP
```

```
       360        370        380        390        400
WIDLVLQNQN VVWTIFGANS MVQVSENVLC LGIVDGGVNA RTSIVIGGHT
       410        420        430
IEDNLLQFDH AASRLGFTSS ILFRQTTCAN FNFTSVA
```

A0A0B2S9A1 – *Glycine soja*

```
        10         20         30         40         50
MQVPVLVDLN GNHLWVNCEQ HYSSKTYQAP FCHSTQCSRA NTHQCLSCPA
        60         70         80         90        100
ASRPGCHKNT CGLMSTNPIT QQTGLGELGQ DVLAIHATQG STQQLGPLVT
       110        120        130        140        150
VPQFLFSCAP SFLLQKGLPR NIQGVAGLGH APISLPNQLA SHFGLQHQFT
       160        170        180        190        200
TCLSRYPTSK GALIFGDAPN NMQQFHNQDI FHDLAFTPLT VTPQGEYNVR
       210        220        230        240        250
VSSIRINQHS VFPPNKISST IVGSSGGTMI STSTPHMVLQ QSLYQAFTQV
       260        270        280        290        300
FAQQLEKQAQ VKSVAPFGLC FNSNKINAYP SVDLVMDKPN GPVWRISGED
       310        320        330        340        350
LMVQAQPGVT CLGVMNGGMQ PRAEVTLGTR QLEEKLMVFD LARSRVGFST
       360
SSLHSHGVKC GDLFNFANA
```

Q8H0K8 – *Triticum aestivum*

```
        10         20         30         40         50
MPPVLLLVLA ASLVALPSCQ SLPVLAPVTK DPATSLYTIP FHDGASLVLD
        60         70         80         90        100
VAGPLVWSTC DGGQPPAEIP CSSPTCLLAN AYPAPGCPAP SCGSDKHDKP
       110        120        130        140        150
CTAYPYNPVS GACAAGSLSH TRFVANTTDG SKPVSKVNVG VLAACAPSKL
       160        170        180        190        200
LASLPRGSTG VAGLANSGLA LPAQVASAQK VANRFLLCLP TGGPGVAIFG
       210        220        230        240        250
GGPVPWPQFT QSMPYTPLVT KGGSPAHYIS ARSIVVGDTR VPVPEGALAT
       260        270        280        290        300
GGVMLSTRLP YVLLRPDVYR PLMDAFTKAL AAQHANGAPV ARAVEAVAPF
       310        320        330        340        350
GVCYDTKTLG NNLGGYAVPN VQLGLDGGSD WTMTGKNSMV DVKQGTACVA
       360        370        380        390        400
FVEMKGVAAG DGRAPAVILG GAQMEDFVLD FDMEKKRLGF SRLPHFTGCG

GL
```

S0F0Q8 – *Triticum turgidum*

```
        10         20         30         40         50
MARVLLLVLA ASLVALASSK GLPVLAPVTK DTATSLYTIP FHDGASLVLD
        60         70         80         90        100
VAGPLVWSTC EGSQPPAEIP CSSPTCLLSN AYPAPGCPAP SCGSDRHDKP
       110        120        130        140        150
CTAYPSNPVT GACAAGSLFH TKFAANTTDG NKPVSEVNVG VLAACAPSKL
       160        170        180        190        200
LASLPRGSTG VAGLANSGLA LPAQVASTQK VANRFLLCLP TGGLGVAIFG
       210        220        230        240        250
GGPLPWPQFT QSMDYTPLVA KGGSPAHYIS LKSIKVENTR VPVSERALAT
       260        270        280        290        300
GGVMLSTRLP YVLLRRDVYR PFVGAFTKAL AAQPANGAPV ARAVKPVAPF
       310        320        330        340        350
ELCYDTKSLG NNLGGYWVPN VGLAVDGGSD WAMTGKNSMV DVKPGTACVA
       360        370        380        390        400
FVEMKGVEAG DGRAPAVILG GAQMEDFVLD FDMEKKRLGF LRLPHFTGCG

S
```

Q8GT67 – *Solanum lycopersicum*

```
        10         20         30         40         50
MASSNCLHAI LLCSLLFITS TIAQNQTSFR PKGLIIPVTK DASTLQYLTQ
```

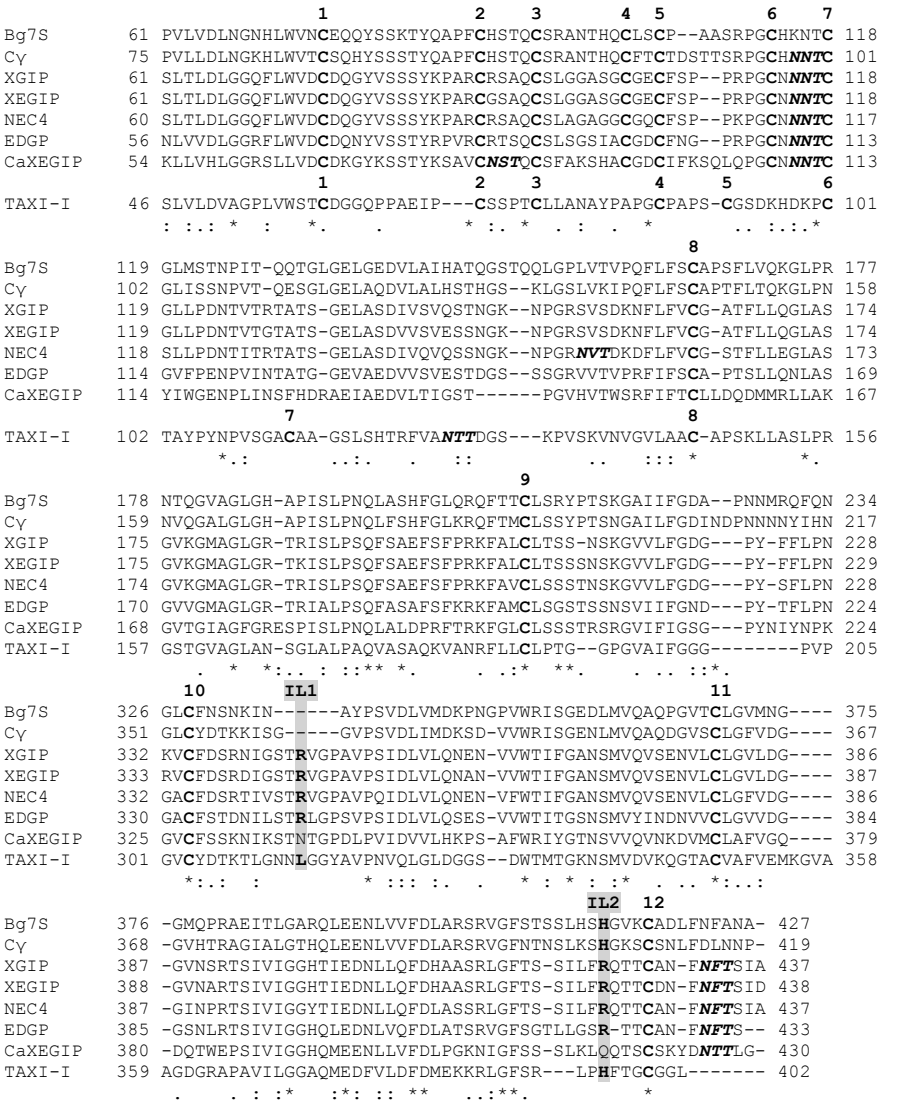
```
        60          70          80          90         100
IQQRTPLVPI SLTLDLGGQF LWVDCDQGYV SSSYKPARCG SAQCSLGGAS
       110         120         130         140         150
GCGECFSPPR PGCNNNTCGL LPDNTVTGTA TSGELASDVV SVESSNGKNP
       160         170         180         190         200
GRSVSDKNFL FVCGATFLLQ GLASGVKGMA GLGRTKISLP SQFSAEFSFP
       210         220         230         240         250
RKFALCLTSS SNSKGVVLFG DGPYFFLPNR QFSNNDFQYT PLFINPVSTA
       260         270         280         290         300
SAFSSGQPSS EYFIGVKSIK INQKVVPINT TLLSIDNQGV GGTKISTVNP
       310         320         330         340         350
YTILETSLYN AITNFFVKEL ANVTRVAVVA PFRVCFDSRD IGSTRVGPAV
       360         370         380         390         400
PSIDLVLQNA NVVWTIFGAN SMVQVSENVL CLGVLDGGVN ARTSIVIGGH
       410         420         430
TIEDNLLQFD HAASRLGFTS SILFRQTTCD NFNFTSID
```

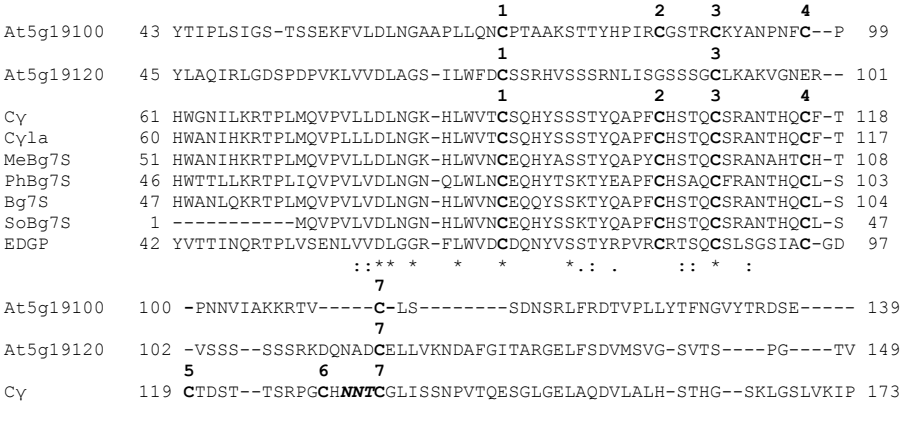
Q7XJE7 – *Solanum tuberosum*

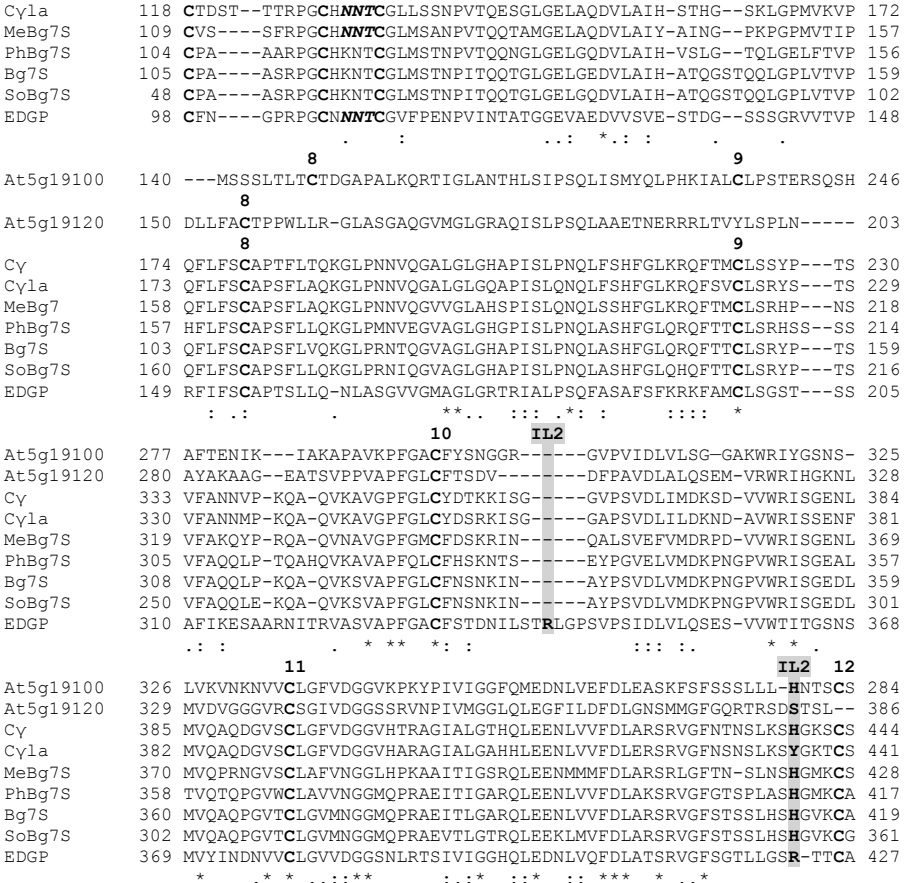
```
        10          20          30          40          50
MASSYCLYAI LLCSLLFITS TIAQNQTSFR PKGLIIPVTK DASTLQYLTQ
        60          70          80          90         100
IQQRTPLVPI SLTLDLGGQF LWVDCDQGYV SSSYKPARCR SAQCSLGGAS
       110         120         130         140         150
GCGECFSPPR PGCNNNTCGL LPDNTVTRTA TSGELASDIV SVQSTNGKNP
       160         170         180         190         200
GRSVSDKNFL FVCGATFLLQ GLASGVKGMA GLGRTRISLP SQFSAEFSFP
       210         220         230         240         250
RKFALCLTSS NSKGVVLFGD GPYFFLPNRE FSNNDFQYTP LFINPVSTAS
       260         270         280         290         300
AFSSGQPSSE YFIGVKSIKI NQKVVPINTT LLSIDNQGVG GTKISTVNPY
       310         320         330         340         350
TILETSLYNA ITNFFVKELA NVTRVAAVAP FKVCFDSRNI GSTRVGPAVP
       360         370         380         390         400
SIDLVLQNEN VVWTIFGANS MVQVSENVLC LGVLDGGVNS RTSIVIGGHT
       410         420         430
IEDNLLQFDH AASRLGFTSS ILFRQTTCAN FNFTSIA
```

	ENTRY	PROTEIN NAME	ORGANISM
AeBg7S	R7W0U4	Aegilops tauschii	Basic 7S globulin 2
At1g03220	Q9ZVS4	Arabidopsis thaliana	Aspartyl protease-like protein
At1g03230	Q9ZVS5	Arabidopsis thaliana	Aspartyl protease-like protein
At5g19100	Q3E9C8	Arabidopsis thaliana	Eukaryotic aspartyl protease family protein
At5g19110	F4JZM1	Arabidopsis thaliana	Eukaryotic aspartyl protease family protein
At5g19120	Q0WU92	Arabidopsis thaliana	Conglutin gamma-like protein
Bg7S	P13917	Glycine max	Basic 7S globulin
CaXEGIP	B9VUU9	Capsicum annuum	Xyloglucan-specific endo-beta-1,4-glucanase inhibitor
CaXEGIP	K4FKJ9	Capsicum annuum	Xyloglucan-specific endo-beta-1,4-glucanase inhibitor
Cγ	Q9FSH9	Lupinus albus	Conglutin gamma
Cyla	Q42369	Lupinus angustifolius	Conglutin gamma
EDGP	Q05929	Daucus carota	Extracellular dermal glycoprotein
HlXEGIP	K9JA99	Humulus lupulus	Xyloglucanase inhibitor 2
HVXI	Q6KE44	Hordeum vulgare var. distichum	Xylanase inhibitor
MeBg7S	Q2HTN2	Medicago truncatula	Basic 7S globulin-like protein
MoBg7S	W9QL21	Morus notabilis	Basic 7S globulin
NEC4	Q3KU27	Nicotiana langsdorffii x Nicotiana sanderae	Nectarin IV
PeXEGIP	B9VUV0	Petunia hybrida	Xyloglucanase-specific endoglucanase inhibitor protein
PhBg7S	V7BPV5	Phaseolus vulgaris	Uncharacterized protein
SCXI	Q6KE41	Secale cereale	Xylanase inhibitor
SmXEGIP	D6QUQ0	Solanum melongena	Xyloglucan specific endoglucanase inhibitor
SoBg7S	A0A0B2S9A1	Glycine soja	Basic 7S globulin
TAXI-I	Q8H0K8	Triticum aestivum	Xylanase inhibitor protein
TDXI	S0F0Q8	Triticum turgidum subsp. durum	Xylanase inhibitor
XEGIP	Q8GT67	Solanum lycopersicum	Xyloglucan-specific fungal endoglucanase inhibitor protein
XGIP	Q7XJE7	Solanum tuberosum	Putative xyloglucanase inhibitor

Appendix_GHIPs Alignment






Cγla 118 CTDST--TTRPGCHNNTCGLLSSNPVTQESGLGELAQDVLAIH-STHG--SKLGPMVKVP 172
MeBg7S 109 CVS----SFRPGCHNNTCGLMSANPVTQQTAMGELAQDVLAIY-AING--PKPGPMVTIP 157
PhBg7S 104 CPA----AARPGCHKNTCGLMSTNPVTQQNGLGELGQDVLAIH-VSLG--TQLGELFTVP 156
Bg7S 105 CPA----ASRPGCHKNTCGLMSTNPITQQTGLGELGEDVLAIH-ATQGSTQQLGPLVTVP 159
SoBg7S 48 CPA----ASRPGCHKNTCGLMSTNPITQQTGLGELGQDVLAIH-ATQGSTQQLGPLVTVP 102
EDGP 98 CFN----GPRPGCNNNTCGVFPENPVINTATGGEVAEDVVSVE-STDG--SSSGRVVTVP 148
8
9
At5g19100 140 ---MSSSLTLTCTDGAPALKQRTIGLANTHLSIPSQLISMYQLPHKIALCLPSTERSQSH 246
8
At5g19120 150 DLLFACTPPWLLR-GLASGAQGVMGLGRAQISLPSQLAAETNERRRLTVYLSPLN----- 203
8
9
Cγ 174 QFLFSCAPTFLTQKGLPNNVQGALGLGHAPISLPNQLFSHFGLKRQFTMCLSSYP---TS 230
Cγla 173 QFLFSCAPSFLAQKGLPNNVQGALGLGQAPISLQNQLFSHFGLKRQFSVCLSRYS---TS 229
MeBg7 158 QFLFSCAPSFLAQKGLPNNVQGVVGLAHSPISLQNQLSSHFGLKRQFTMCLSRHP---NS 218
PhBg7S 157 HFLFSCAPSFLLQKGLPMNVEGVAGLGHGPISLPNQLASHFGLQRQFTTCLSRHSS--SS 214
Bg7S 103 QFLFSCAPSFLVQKGLPRNTQGVAGLGHAPISLPNQLASHFGLQRQFTTCLSRYP---TS 159
SoBg7S 160 QFLFSCAPSFLLQKGLPRNIQGVAGLGHAPISLPNQLASHFGLQHQFTTCLSRYP---TS 216
EDGP 149 RFIFSCAPTSLLQ-NLASGVVGMAGLGRTRIALPSQFASAFSFKRKFAMCLSGST---SS 205
10
IL2
At5g19100 277 AFTENIK---IAKAPAVKPFGACFYSNGGR------GVPVIDLVLSG-GAKWRIYGSNS- 325
At5g19120 280 AYAKAAG--EATSVPPVAPFGLCFTSDV--------DFPAVDLALQSEM-VRWRIHGKNL 328
Cγ 333 VFANNVP-KQA-QVKAVGPFGLCYDTKKISG-----GVPSVDLIMDKSD-VVWRISGENL 384
Cγla 330 VFANNMP-KQA-QVKAVGPFGLCYDSRKISG-----GAPSVDLILDKND-AVWRISSENF 381
MeBg7S 319 VFAKQYP-RQA-QVNAVGPFGMCFDSKRIN------QALSVEFVMDRPD-VVWRISGENL 369
PhBg7S 305 VFAQQLP-TQAHQVKAVAPFQLCFHSKNTS------EYPGVELVMDKPNGPVWRISGEAL 357
Bg7S 308 VFAQQLP-KQA-QVKSVAPFGLCFNSNKIN------AYPSVDLVMDKPNGPVWRISGEDL 359
SoBg7S 250 VFAQQLE-KQA-QVKSVAPFGLCFNSNKIN------AYPSVDLVMDKPNGPVWRISGEDL 301
EDGP 310 AFIKESAARNITRVASVAPFGACFSTDNILSTRLGPSVPSIDLVLQSES-VVWTITGSNS 368
11
IL2
12
At5g19100 326 LVKVNKNVVCLGFVDGGVKPKYPIVIGGFQMEDNLVEFDLEASKFSFSSSLLL-HNTSCS 284
At5g19120 329 MVDVGGGVRCSGIVDGGSSRVNPIVMGGLQLEGFILDFDLGNSMMGFGQRTRSDSTSL-- 386
Cγ 385 MVQAQDGVSCLGFVDGGVHTRAGIALGTHQLEENLVVFDLARSRVGFNTNSLKSHGKSCS 444
Cγla 382 MVQAQDGVSCLGFVDGGVHARAGIALGAHHLEENLVVFDLERSRVGFNSNSLKSYGKTCS 441
MeBg7S 370 MVQPRNGVSCLAFVNGGLHPKAAITIGSRQLEENMMMFDLARSRLGFTN-SLNSHGMKCS 428
PhBg7S 358 TVQTQPGVWCLAVVNGGMQPRAEITIGARQLEENLVVFDLAKSRVGFGTSPLASHGMKCA 417
Bg7S 360 MVQAQPGVTCLGVMNGGMQPRAEITLGARQLEENLVVFDLARSRVGFSTSSLHSHGVKCA 419
SoBg7S 302 MVQAQPGVTCLGVMNGGMQPRAEVTLGTRQLEEKLMVFDLARSRVGFSTSSLHSHGVKCG 361
EDGP 369 MVYINDNVVCLGVVDGGSNLRTSIVIGGHQLEDNLVQFDLATSRVGFSGTLLGSR-TTCA 427
1
2
3
4
5
6
At5g19110 54 LLLDLGTNLTWLDCRKLKSLSSLRLVTCQSSTCKS---IPGNGCAGKSCLY--------- 102
AeBg7S 51 IVVDTAGPLVWSTCAPDH--L-PAAFPCKSDTCRLANKYHVPSCSESAADKLCDPSHKVC 107
HVXI 47 LVLDVAGPLVWSTCDGGQRPP-PAEITCSSPTCLLANAYPAPGCPAPSCG--SDRHDKPC 103
TAXI-I 47 LVLDVAGPLVWSTCDGGQ--P-PAEIPCSSPTCLLANAYPAPGCPAPSCG--SDKHDKPC 101
SCXI 47 LVLDAAGPLVWSTCEAGQ--P-PAGIPCGSPTCLLANAYPAPGCPAPTCG--SD---KPC 98
TDXI 47 LVLDVAGPLVWSTCEGSQ--P-PAEIPCSSPTCLLSNAYPAPGCPAPSCG--SDRHDKPC 101
7
8
At5g19110 103 --KQPNPLGQNPVVTGRVVQDRASLYTTDGGKFLSQVSVRHFTFSCAGEKALQGLPPPVD 160
AeBg7S 108 RAFPYNPV-TGACAAGDLIHTRFVANTTDGKNPVSQVNVRAV-AACAPSKLLESLPQGAS 165
HVXI 104 TAYPSNPV-TGACAAGSLFRARLVANITDGNRPVSAVTVGVL-AACAPTKLLASLPRGST 161
TAXI-I 102 TAYPYNPV-SGACAAGSLSHTRFVANTTDGSKPVSKVNVGVL-AACAPSKLLASLPRGST 159
SCXI 99 TAFPSNPV-TGACAAGSLFHTSFVANTTDGTKPVSEVKVGVL-AACAPSKLLASLPRGST 156
TDXI 102 TAYPSNPV-TGACAAGSLFHTKFAANTTDGNKPVSEVNVGVL-AACAPSKLLASLPRGST 159
9
At5g19110 161 GVLALSPGSSSFTKQVTSAFNVIPKFSLCLPSS---GTGHFYIAG--IHYFIPPFNSSDN 215
AeBg7S 166 GVAGLAGSDLALPAQVASAQKVSNKFLLCLPRGLSSDPGVAVFGGGPLHFMAQPERDYTK 226
HVXI 162 GVAGLAGSGLALPAQVASAQKVSHRFLLCLPTG---GAGVAILGGGPLPW-----PQFTQ 213
TAXI-I 160 GVAGLANSGLALPAQVASAQKVANRFLLCLPTG---GPGVAIFGGGPVPW-----PQFTQ 211
SCXI 157 GVAGLANSGLALPAQVASAQKVANRFFLCLPTG---GAGVAIFGGGPLPW-----PQFTQ 208
TDXI 160 GVAGLANSGLALPAQVASTQKVANRFLLCLPTG---GLGVAIFGGGPLPW-----PQFTQ 211
10
IL1
At5g19110 275 ALAQSFTLKAKAMG------IAKVPSVAPFKHCFDSRTAGKNLTAGPNVPVIEIGLPGRI 328
AeBg7S 284 PLVDAFTKALAGQGAQGGPVAKAVKPYAPFQLCYDTRTLANTRT-GYLVPAVTLTLG--- 340
HVXI 273 PFVDAFAKALAAQHANGALAARGVNPVAPFGLCYDAKTLGNNLG-GYSVPNVVLALD--- 328
TAXI-I 271 PLMDAFTKALAAQHANGAPVARAVEAVAPFGVCYDTKTLGNNLG-GYAVPNVQLGLD--- 326
SCXI 265 PLVDAFTKALAAQPANGAPVARAVQPVAPFGVCYDTKTLGNNLG-GYAVPNVLLALD--- 320
TDXI 271 PFVGAFTKALAAQPANGAPVARAVKPVAPFELCYDTKSLGNNLG-GYWVPNVGLAVD--- 326
11
At5g19110 329 GEVKWGFYGANTVVKVKETVMCLAFID-----GGKTPKDLMVIGTHQLQDHMLEFDFSGT 383
AeBg7S 341 GGKNWRMDGLSLMVDMGPTTACLAFVQMQGVKGGDGSAPSVLIGGFQMENTVLEFDMKKK 400
HVXI 329 GGGEWAMTGKNSMVDVKPGTACVAFVEME---AGDGGAPAVILGGAQMEDFVLDFDMEKK 385
TAXI-I 327 GGSDWTMTGKNSMVDVKQGTACVAFVEMKGVAAGDGRAPAVILGGAQMEDFVLDFDMEKK 386
SCXI 321 GGGEWAMTGKNSMVDVKPGTACVAFVEMKGVEAGDGRAPAVILGGAQMEDFVLDFDMEKK 380
TDXI 327 GGSDWAMTGKNSMVDVKPGTACVAFVEMKGVEAGDGRAPAVILGGAQMEDFVLDFDMEKK 386

```
                      IL2  12
At5g19110   384 VLAFSESLLLHNTSCSTWPSQK--- 405
AeBg7S      401 RLGFARL--PSFTQCGQFNFTTRSA 423
HVXI        386 RLGFIRL--PHFTGCGNLNF----- 403
TAXI-I      387 RLGFSRL--PHFTGCGGL------- 402
SCXI        381 RLGFTRL--PHFTGCGSA------- 396
TDXI        387 RLGFLRL--PHFTGCGS-------- 401
                 *.* .      * *.


                                      1              2    3         4  5
MoBg7S          61 KRTPPLQVPLTIDLNGRFLWANCEGGSYLSSTYNAPLCHSTQCSQAVGPNHYCRTCSSRA 116
At1g03220       50 QRTPLVPASVVFDLGGRELWVDCDKG-YVSSTYQSPRCNSAVCSRAG--STSCGTCFSPP 106
At1g03230       51 QRTPLVPASVVFDLGGREFWVDCDQG-YVSTTYRSPRCNSAVCSRAG--SIACGTCFSPP 107
EDGP            48 QRTPLVSENLVVDLGGRFLWVDCDQN-YVSSTYRPVRCRTSQCSLSG--SIACGDCFNGP 104
HlXEGIP         53 QRTPLVPVKLTVNLGGEFLWVDCEKG-YVSSTYKPARCRSAQCNLAG--SKSCGECFDGP 109
CaXEGIP         53 QRTPLVPVSLTLDLGGQFLWVDCDQG-YVSSSYKPARCRSAQCSLAG--ATGCGECFSPP 109
XEGIP           53 QRTPLVPISLTLDLGGQFLWVDCDQG-YVSSSYKPARCGSAQCSLGG--ASGCGECFSPP 109
SmXEGIP         53 QRTPLVPISLTLDLGGQFLWVDCDQG-YVSSSYKPARCRSAQCSLAG--ASACGECFSPP 109
NEC4            52 QRTHLVPVSLTLDLGGQFLWVDCDQG-YVSSSYKPARCRSAQCSLAG--AGGCGQCFSPP 108
PeXEGIP         52 QRTPLVPVSLTLDLGGQFLWVDCDQG-YVSSSYIPARCRSAKCSLAG--SSGCGDCFSPP 108
                  :**  :   :..:*.*. :*.:*: . *:*::*    * :: *. .      *  * .
                     6    7                                          8
MoBg7S        117 RPGCHNGTCGVTVTNPVTGRSAIGELAQDSLSVRSAQGPRQGQGPIARVRQFLFVCAPSA 180
At1g03220     107 RPGCSNNTCGGIPDNTVTGTATSGEFALDVVSIQSTNGSNPG--RVVKIPNLIFDCGATF 164
At1g03230     108 RPGCSNNTCGAFPDNSITGWATSGEFALDVVSIQSTNGSNPG--RFVKIPNLIFSCGSTS 165
EDGP          105 RPGCNNNTCGVFPENPVINTATGGEVAEDVVSVESTDGSSSG--RVVTVPRFIFSCAPTS 162
HlXEGIP       110 KPGCNNNTCGLFPYNPFIRTSTSGELAQDIISIQSTNGSNPS--KVVSFPNVIFTCGSTF 167
CaXEGIP       110 RPGCNNNTCGLFPDNTVTRTATSGELASDVVSVQSSNGKNPG--RNVSDKNFLFVCGATF 167
XEGIP         110 RPGCNNNTCGLLPDNTVTGTATSGELASDVVSVESSNGKNPG--RSVSDKNFLFVCGATF 167
SmXEGIP       110 RPGCNNNTCSLFPDNTVTGTATGGELASDIVSVQSSNGKNPG--RNVSDKNFLFVCGATF 167
NEC4          109 KPGCNNNTCSLLPDNTITRTATSGELASDIVQVQSSNGKNPG--RNVTDKDFLFVCGSTF 166
PeXEGIP       109 SPGCNNNTCGAFPDNSITRTATSGELASDIVSVQSSNGKNPG--RNVSDKDFLFVCGATF 166
                   *** *.**.    * .   :: **.* * :.:.*::*   .    .    .:* *. :
                                                        9
MoBg7S        181 LLQPGLPRKAQGVVGLGHSHVSLPSQLASHFGFQQKFATCLPR-GNGNGAVFFGEGPYFF 239
At1g03220     165 LL-KGLAKGTVGMAGMGRHNIGLPSQFAAAFSFHRKFAVCLTS---GKGVAFFGNGPYVF 220
At1g03230     166 LL-KGLAKGAVGMAGMGRHNIGLPLQFAAAFSFNRKFAVCLTS---GRGVAFFGNGPYVF 221
EDGP          163 LL-QNLASGVVGMAGLGRTRIALPSQFASAFSFKRKFAMCLSGSTSSNSVIIFGNDPYTF 221
HlXEGIP       168 LL-EGLASGVTGIAGLGRKKIALPSQFAAAFSFKRKFALCLSSSTRATGVVFFGDGPYIM 226
CaXEGIP       168 LL-QGLASGVKGMAGLGRTRISLPSQFSAEFSFPRKFAVCLSS-SKSKGVVLFGDGPYFF 225
XEGIP         168 LL-QGLASGVKGMAGLGRTKISLPSQFSAEFSFPRKFALCLTSSSNSKGVVLFGDGPYFF 226
SmXEGIP       168 LL-QGLASGVKGMAGLGRTRISLPSQFSAEFSFPRKFALCLTS-SNSKGVVLFGDGPYFF 225
NEC4          167 LL-EGLASGVKGMAGLGRTRISLPSQFSAEFSFPRKFAVCLSSSTNSKGVVLFGDGPYSF 225
PeXEGIP       167 LL-NGLASGVKGMAGLGRTRISLPSQFSAEFSFPRKFAVCLSSTSNSKGVVLFGDGPYSF 225
                   **  .*   . *:.*:*: .:.** *::: *.* :*** **     . .. :**:.** :
                                                                    10
MoBg7S        279 ---PVSRAIITTTNPYTFLDHSLFVALTNVFANQL---KIPRVQPVAPFGACFDAKGIAS 332
At1g03220     279 ASTGIGGTKISSVNPYTVLESSIYNAFTSEFVKQAAARSIKRVASVKPFGACFSTKNVGV 338
At1g03230     280 ASTGIGGTKISSVNPYTVLESSIYKAFTSEFIRQAAARSIKRVASVKPFGACFSTKNVGV 339
EDGP          282 -SAGLGGTKISTINPYTVLETSIYKAVTEAFIKESAARNITRVASVAPFGACFSTDNILS 340
HlXEGIP       286 -KDGTGGTKISTTQPYTSLETSIYKAVIGAFGKAV--AKVPRVTAVAPFELCFNSTSFSS 342
CaXEGIP       286 -NQGVGGTKISTVNPYTVLETSLYNAITNFFVKEL--ANVTRVASVAPFGACFDSRNIGS 342
XEGIP         287 -NQGVGGTKISTVNPYTILETSLYNAITNFFVKEL--ANVTRVAVVAPFRVCFDSRDIGS 343
SmXEGIP       286 -NQGVGGTKLSTVNPYTVMETSLYNAITNFFVKEL--ANVTRVAPVTPFGACFDSRNIGS 342
NEC4          286 -NQGVGGTKISTVNPYTILETSMYNAVTNFFVKEL--VNITRVASVAPFGACFDSRTIVS 342
PeXEGIP       286 -SQGVGGTKISTVNPYTILETSIYNAVTNFFVKEL--A-IPTVPSVAPFGVCFDSRNITS 341
                      . : ::: :*** :: *:: *.   * .      :  *  * **  **.:  .
                  IL1                                  11
MoBg7S        333 TRIGPAVPPVDLSLHDQSTRWRILGANSMIEARPGVMCLAFVDGGARP--------HGSS 384
At1g03220     339 TRLGYAVPEIELVLHSKDVVWRIFGANSMVSVSDDVICLGFVDGGVN---------ARTS 389
At1g03230     340 TRLGYAVPEIQLVLHSKDVVWRIFGANSMVSVSDDVICLGFVDGGVN---------PGAS 390
EDGP          341 TRLGPSVPSIDLVLQSESVVWTITGSNSMVYINDNVVCLGVVDGGSN---------LRTS 391
HlXEGIP       343 TRVGPGVPQIDLVLPN-NKAWTIFGANSMVQVSDDVLCLGFVDGGPLHFVDWGIPFTPTA 401
CaXEGIP       343 TRVGPAVPQIDLVLQNENVIWTIFGANSMVQVSENVLCLGFVDGGVN---------SRTS 393
XEGIP         344 TRVGPAVPSIDLVLQNANVVWTIFGANSMVQVSENVLCLGVLDGGVN---------ARTS 394
SmXEGIP       343 TRVGPAVPWIDLVLQNQNVVWTIFGANSMVQVSENVLCLGIVDGGVN---------ARTS 393
NEC4          343 TRVGPAVPQIDLVLQNENVFWTIFGANSMVQVSENVLCLGFVDGGIN---------PRTS 393
PeXEGIP       342 TRVGPGVPSIDLVLQNENVFWRIFGANSMVLVSENVLCLGFVDGGVN---------PRTS 392
                  **:* .** ::* * . .  * * *:***:    .*:**..:***            ::
                                          IL2  12
MoBg7S        385 MVIGAYQLEDNLVQFDLVKSMLGFSSSLLF-RRTSCSNFNFTSSTTSP 401
At1g03220     390 VVIGGFQLEDNLIEFDLASNKFGFSSTLLG-RQTNCANFNFTSTA--- 433
At1g03230     391 VVIGGFQLEDNLIEFDLASNKFGFSSTLLG-RQTNCANFNFTSTA--- 434
EDGP          392 IVIGGHQLEDNLVQFDLATSRVGFSGTLLGSR-TTCANFNFTS----- 433
HlXEGIP       402 IVIGGHQIEDNLLQFDLGSSTLGFSSSLLF-RQTTCSNFNFTSIA--- 445
CaXEGIP       394 IVIGGHTIEDNLLQLDIARSRLGFTSSILF-RQTTCANFNFTSIA--- 437
XEGIP         395 IVIGGHTIEDNLLQFDHAASRLGFTSSILF-RQTTCDNFNFTSID--- 438
SmXEGIP       394 IVIGGHTIEDNLLQFDHAASRLGFTSSILF-RQTTCANFNFTSVA--- 437
NEC4          394 IVIGGYTIEDNLLQFDLASSRLGFTSSILF-RQTTCANFNFTSIA--- 437
PeXEGIP       393 IVIGGHTIEDNLLQFDLAASRLGFTSSILF-RQTTCANFNFTSIA--- 436
                  :***.. :****:::*   . .**:.::*  *:*.* ******
```

PUBBLICATIONS OF THE CANDIDATE

Abbasi Parizad P., **Galanti E.**, Scarafoni A., Lavelli V., Bonomi F., Iametti S. and Marengo M.; Pigmented grains as a source of immunomodulating bioactives. XIV FISV Congress, Rome, 20-23 September 2016, poster session.

Galanti E. and Scarafoni A; Deeping into the structural and functional features of selected putative proteins sequences using a cell free protein system: the example of the GHIPs-like proteins from Arabidopsis thaliana. XI National Conference of the Italian Proteomics Association, University of Perugia, 16-19 May 2016, poster session.

Galanti E., Scarafoni A., Duranti M.; Redefining the role of the seed storage proteins. Biochemistry Conference, University of Milan, Palazzo Feltrinelli in Gargnano (BS) 20-22 March 2016, oral communication.

Scarafoni A., Consonni A., Pessina S., Balzaretti S., Capraro J., **Galanti E.** and Duranti M.; Structural basis of the lack of endo-glucanase inhibitory activity of Lupinus albus γ-conglutin; Plant Physiology and Biochemistry, DOI 10.1016/j.plaphy.2015.11.008.

Capraro J., **Galanti E.**, Marengo M., Duranti M. and Scarafoni A.; Cysteine-containing peptides are produced by sequential clipping, but not released, from lupin 11S storage globulin during early germination; Peptidomics, DOI: 10.1515/ped-2015-0005.

Galanti E., Scarafoni A., Duranti M.; Glycoside hydrolase inhibitor proteins: biochemical features and molecular biology. Biochemistry Conference, University of Milan, Palazzo Feltrinelli in Gargnano (BS) 12-14 April 2015, poster session.

Guerrieri N., De Benedetti E., **Galanti E.**, Scarafoni A., Scaglioni L.; Arsenico e Lepidium sativum: la risposta della proteomica. Convegno nazionale “Arsenico nelle catene alimentari", Istituto Superiore della Sanità, Roma, 4-5 June 2015.

Plant Physiology and Biochemistry 99 (2016) 79–85

Contents lists available at ScienceDirect

Plant Physiology and Biochemistry

journal homepage: www.elsevier.com/locate/plaphy

Research article

Structural basis of the lack of endo-glucanase inhibitory activity of *Lupinus albus* γ-conglutin

Alessio Scarafoni*, Alessandro Consonni, Stefano Pessina, Silvia Balzaretti, Jessica Capraro, Elisabetta Galanti, Marcello Duranti

Department of Food, Environmental and Nutritional Sciences (DeFENS), Università degli Studi di Milano, Milano, Italy

ARTICLE INFO

Article history:
Received 9 August 2015
Received in revised form
30 October 2015
Accepted 16 November 2015
Available online 24 November 2015

Keywords:
Seed proteins
Enzyme inhibitors
Mutagenesis
Plant defence

ABSTRACT

Lupin γ-conglutin and soybean BG7S are two legume seed proteins strongly similar to plant endo-β-glucanases inhibitors acting against fungal GH11 and GH12 glycoside hydrolase. However these proteins lack inhibitory activity. Here we describe the conversion of lupin γ-conglutin to an active inhibitor of endo-β-glucanases belonging to GH11 family. A set of γ-conglutin mutants was designed and expressed in *Pichia pastoris*, along with the wild-type protein. Unexpectedly, this latter was able to inhibit a GH11 enzyme, but not GH12, whereas the mutants were able to modulate the inhibition capacity. In lupin, γ-conglutin is naturally cleaved in two subunits, whereas in *P. pastoris* it is not. The lack of proteolytic cleavage is one of the reasons at the basis of the inhibitory activity of recombinant γ-conglutin. The results provide new insights about structural features at the basis of the lack of inhibitory activity of wild-type γ-conglutin and its legume homologues.

1. Introduction

One of the main functions of plant cell wall is to act as a barrier against microbial pathogens. Cell walls are essentially made up by two polysaccharide networks, one containing mainly pectins and the other one containing cellulose and hemicellulose. The major polymer in cereals hemicellulose fractions is xylan, whereas in dicotyledonous plants the predominant hemicellulose polysaccharides are xyloglucans (Juge, 2006).

To penetrate the cell wall, pathogens secrete a variety of glycoside hydrolases (GHs) that digest the polysaccharides chains. The wall degradation requires the synergistic action of several hydrolytic enzymes. Endo-β-1,4-xylanases (XYs) are pivotal enzymes that operate the initial breakdown of the xylan backbone of monocots, whereas xyloglucan-specific endoglucanases (XEGs) act in dicots plant. XYs belong to the GH10 and GH11 families whereas XEGs are part of the GH12 and GH74 families (Juge, 2006; Cantarel et al., 2009).

In response to pathogen attacks, plants produce specific GH inhibitory proteins (GHIPs) that contrast the action of degrading enzymes. *Triticum aestivum* xylanase inhibitor-IA (TAXI-I) and xyloglucan-specific endo-β-1,4-glucanase inhibitor proteins (XEGIPs) are two of such kind of proteins. They have common structural features but they are effective against different classes of xylanases: TAXI-I inhibits GH11 members, whereas XEGIPSs inhibit GH12 members (Qin et al., 2003; Sansen et al., 2004; Juge, 2006). It has been recently suggested a possible involvement of XEGIPs in programmed cell death, triggered by the interaction of the inhibitory protein with the respective XEG (Naqvi et al., 2005; Choi et al., 2013).

γ-Conglutin is a glycoprotein of *Lupinus* spp. It accounts for about 4% of the total seed proteins and is made up of six monomers, each of them formed by two subunits of about 30 and 17 kDa linked by a disulphide bridge. These two polypeptide chains originate from a single precursor synthesized during seed development and processed by post-translational proteolysis. Only the large subunits are N-glycosylated at a single site (Duranti et al., 2008). The complete amino acid sequence is available (UniProtKB/TrEMBL: Q9FSH_LUPAL) (Scarafoni et al., 2010). It is similar to both TAXI-I and XEGIPs proteins (Czubinski et al., 2015), with which it shares an array of 12 cysteines and amino acid sequence identity (s.i.),

Abbreviations: DLS, dynamic light scattering; GH, glycosyl hydrolase; PTM, post-translational modification; s.i., sequence identity; γ-c, seed-extracted γ-conglutin; rγc, recombinant γ-conglutin; rγcMu, recombinant γ-conglutin mutant; XEG, xyloglucan-specific endoglucanase; XY, Endo-β-1,4-xylanases.

* Corresponding author. Department of Food, Environmental and Nutritional Sciences (DeFENS), Università degli Studi di Milano, via G. Celoria 2, 20133, Milano, Italy.

E-mail address: alessio.scarafoni@unimi.it (A. Scarafoni).

http://dx.doi.org/10.1016/j.plaphy.2015.11.008
0981-9428/

ranging from 28,3% with TAXI-I, to 36–43% with pepper CaXEGIP1, ornamental tobacco NEC4, tomato XEGIP and carrot EDGP (Qin et al., 2003; Sansen et al., 2004; Naqvi et al., 2005; Yoshizawa et al., 2012; Choi et al, 2013). However, the protein failed to inhibit either GH11 and GH12 and other cell wall-degrading enzymes (Scarafoni et al., 2010). γ-Conglutin is also highly similar to Bg7S protein (64% s.i.), a soybean seed protein that does not display any inhibitory activity towards GHs, too (Scarafoni et al., 2010; Yoshizawa et al., 2011). Despite the full characterization at molecular level, the physiological functions of γ-conglutin and Bg7S are still unknown. The expression of γ-conglutin can be elicited during seed germination by chitosan treatment, thus indicating an actual role in defence response mechanisms against fungal attacks (Scarafoni et al., 2010). The crystal structure which as been has been recently determined (Czubinski et al., 2015) revealed molecular features that make γ-conglutin unique among all other homologues, despite the fact that the general fold is preserved. This opened the way to new researches aimed at deepen the structural basis behind the peculiar lack of inhibitory capacity of this legume proteins.

The comprehension of the structural requirements at the basis of GHIPs actions is of great importance, because of their relevance for plant defence and because it can drive specific programs aimed to strengthen plant natural resistance against microbial pathogens.

The results of the present work, carried out by sequence mutagenesis approach, provide new details about structural features and functionality behind the lack of inhibitory capacity of these legume proteins, and, more in general, new insights about the molecular determinants affecting the capacity and specificity of XEGIPs to recognize the target GHs.

2. Material and methods

2.1. General

Protein concentrations were determined according to Bradford (1976), using bovine serum albumin as standard. Seed γ-conglutin were purified according to a previous procedure (Capraro et al., 2010).

2.2. Yeast, bacterial strains and growth conditions

Pichia pastoris X33 strain, *E.coli* XL1-Blue and pPICZaB plasmid were from Invitrogen (San Diego, CA). P. pastoris X33 was cultured in Yeast Peptone Dextrose (YPD) medium (2% peptone, 1% yeast extract, 2% glucose), at 30 °C. Transformants were selected in YPD plates, containing 100 μg/mL zeocin (Invitrogen, San Diego, CA). For protein production, selected transformants were cultured in Buffered Yeast Peptone Sorbitol (BYPS) medium (2% peptone, 1% yeast extract, 2% sorbitol, buffered at pH 7.5 with 100 mM potassium phosphate), at 30 °C, under vigorous shaking. *E. coli* cells were grown in Luria–Bertani (LB) medium, with 25 μg/mL zeocin when needed, under standard conditions.

2.3. Mutants construction

Four mutant genes have been prepared by PCR gene fusion. The experimental strategy is outlined in Supplementary Figure S1. The full set of primers used is listed in Supplementary Table T1. The coding sequence of γ-conglutin gene (EMBL: AJ297490), except the signal peptide, was excised from plasmid pPR12 (Scarafoni et al., 2001) with XhoI and XbaI and cloned into pPICZaB vector, giving pPICZaB-rγc. For preparation of rγcMuX and rγcMuT mutants (giving plasmids pPICZaB-rγcMuX and pPICZaB-rγcMuT), PCR mixes (20 μL, final volume) consisted of 0.5 mM of each primer, 0.8 mM dNTPs, 30 ng of pPICZaB-rγc plasmid DNA, 2.5 U Pfu DNA polymerase, 1 x buffer, as supplied by the manufacture's. Amplifications were carried out on an iQ thermocycler (Bio-Rad, Hercules, CA) adopting the following conditions: 97 °C for 4 min; 30 cycles at 97 °C for 40 s, 60 °C for 40 s, 72 °C for 40 s. The fusion step was performed with the same conditions, but adopting an annealing temperature of 50 °C. rγcMuX-H/S and rγcMuT-H/S mutants were prepared from plasmids pPICZaB-rγcMuX and pPICZaB-rγcMuT, respectively. PCRs were carried out as follows: 97 °C for 4 min; 20 cycles at 97 °C for 40 s, 58 °C for 40 s, 72 °C for 40 s; final extension: 72 °C for 10 min, using Taq polymerase. Plasmids were used to transform XL1-Blue *E. coli* cells, selected on solid LB media containing tetracycline and zeocin. Two of each clones were sequenced to ensure that no point or frame-shift mutation occurred in the constructs sequence.

2.4. Transformation of P. pastoris

Yeast cells were transformed by electroporation using 35 μg of recombinant plasmids or 30 μg of non-recombinant pPICZaB, as previously described (Consonni et al., 2010). Prior to transformation, all plasmids were linearized with SacI restriction enzyme. Insertion of the constructs at the alcohol oxidase promoter (AOX1) site was verified by PCR using 5′AOX1 and 3′AOX1 primers (Consonni et al., 2011).

2.5. Expression and purification

Five recombinants for each construct and one recombinant containing pPICZaB expression vector only (negative control) were grown in 50 mL of BYPS medium at 30 °C to an OD_{600} of 5, under shaking (180 rpm). Expression was induced by adding methanol to 1% final concentration and incubation for 24 h, under shaking. Aliquots of the supernatants were analysed by SDS-PAGE. The clones showing the highest level of expression of each recombinant protein were selected for large-scale production (1 L), adopting the above conditions. Cells were centrifuged at 11,000 *g* at 4 °C for 15 min and the proteins contained in the supernatant were precipitated with ammonium sulphate (70% saturation). The pellet was dissolved in 50 mL sterile water and the proteins again precipitated with cold (−20 °C) acetone (1:1 v/v). After centrifugation at 4 °C for 30 min, the pellet was dissolved in 25 mL 50 mM Tris–HCl (pH 7.5) and loaded on to a DEAE-cellulose column (2.5 × 10 cm, Whatman, Maidstone, UK) equilibrated with the same buffer. Wild-type and mutant γ-conglutins were found in the unbound fraction. Finally, this latter fraction was loaded onto an insulin–agarose immobilized affinity column as previously described (Capraro et al., 2010).

2.6. SDS-PAGE and western blotting

SDS-PAGE was carried out according to Laemmli (1970), under reducing conditions, using a MiniProteanIII electrophoresis cell (Bio-Rad). Runs were carried out at 16 mA constant for each gel. CBB was used for gel staining. Molecular weight markers were from GE Healthcare (Milan, Italy). Western blottings were carried out according to Towbin et al. (1979). Anti-γ-conglutin antibodies (Scarafoni et al., 2013) were diluted 1:1500 in PBS buffer containing 0.02% Tween20. To visualize the bands, goat anti-rabbit antibodies conjugated with horseradish peroxidase (1:3000; BioRad) and H_2O_2 with 4-chloronaphthol as substrate were used.

2.7. Dynamic light scattering (DSC)

DSC has been performed as previously described (Capraro et al.,

2010), using a multi-angle light scattering device (Dawn Heleos, Wyatt), equipped with a fast photon counter (QELS). Data have been analysed with ASTRA V software (Wyatt Technology).

2.8. Size exclusion chromatography (SEC)

SEC was carried out using a Waters 625 HPLC and a Superose 12 HR 10/30 column (GE Healthcare) equilibrated in 30 mM sodium acetate buffer, pH 5.0, containing 0.1 M NaCl. The flow rate of the mobile phase was 0.5 mL/min. Protein elution was monitored at 280 nm. Before use, lyophylized γ-conglutin mutants were dissolved in elution buffer and pre-incubated with the enzyme at room temperature for 15 min, with an enzyme:inhibitor molar ratio of 1:1 (Mr of rγc: 43,000 Da; Mr of *Trichoderma longibrachiatum* XEG: 20.000Da).

2.9. Enzyme activity and inhibition

Aspergillus aculeatus XEG belonging to the GH12 family (Pauly et al., 1999) was from Novozymes (Copenhagen, Denmark) and has been purified as previously described (Qin et al., 2003; Scarafoni et al., 2010). Activity was assayed according to Pauly et al. (1999) and Qin et al. (2003), using tamarind xyloglucan as substrate (Scarafoni et al., 2010). Ten units of XEG were incubated in 50 mM sodium acetate buffer, pH 5.0, or 50 mM sodium phosphate buffer, pH 7.0, with 1.2 mg of substrate at 25 °C, for 40 min (final volume 1 mL). GH11 *Trichoderma longibrachiatum* xylanase M3 (Toerroenen et al., 1992) was purchased from Megazyme (Wicklow, Ireland) and was tested by modifying a previous method (Gebruers et al., 2001; Fierens et al., 2005). Enzyme assays were carried out in 50 mM sodium acetate buffer, pH 5.3, or 50 mM sodium phosphate buffer, pH 7.0, using 2 U of enzyme and 1.3 mg of beechwood xylan as the substrate, at 25 °C 40 min (final volume 1 mL). One unit of enzyme was defined as the enzyme amount which produced 1 μmol/min of reducing sugars following xylan or xyloglucan hydrolysis.

The amount of reducing sugars produced following enzyme activity was assayed by the *p*-hydroxy-benzoic acid hydrazide method (Lever, 1972). Before use, lyophylized γ-conglutin mutants were dissolved in incubation buffer, briefly centrifuged and pre-incubated with the enzyme at room temperature for 15 min, with a molar enzyme:inhibitor ratio of 1:1.

Inhibition rates have been calculated as: $(A_E\text{-}A_{EI})/A_E{}^*100$, where A_E is the measured enzyme activity (μmol/min of reducing sugars produced in 40 min assays), and A_{EI} is the measured activity of the enzyme in the presence of the inhibitor, namely rγc or its mutants.

2.10. Statistical analyses

Enzymes activities have been expressed as means ± S.E. Data were analysed by t-test. P values < 0.05 were considered as statistically significant.

2.11. Bioinformatic analysis

Predictive structures of recombinant γ-conglutin were generated by automated homology modelling, at the Swiss-Model server (http://swissmodel.expasy.org) (Arnold et al., 2006; Biasini et al., 2014). Experimentally determined structures of carrot EDGP and wheat TAXI_I (PDB ID: 2VLA and 2B42, respectively) have been used as templates for modelling.

Structural alignments between two protein structures were carried out by pairwise comparison using DaliLite (Hasegawa and Holm, 2009; Holm and Rosenström, 2010), available at http://ekhidna.biocenter.helsinki.fi/dali_lite/start. The resulting superimposed coordinate files were either downloaded or viewed interactively in Jmol.

3. Results

P. pastoris cultured cells were used to express a γ-conglutin gene (EMBL: AJ297490) and a series of its mutants prepared for this work. This yeast has been widely used for the production of a variety of plant proteins (Macauley-Patrick et al., 2005). TAXI-I was among them and found as active and specific as the natural protein (Fierens et al., 2004).

As already stated, γ-conglutin purified from white lupin seeds does not inhibit any endo-β-glucanases against whom similar proteins act, namely members of GH11 and GH12 classes (Scarafoni et al., 2010). Alignments of γ-conglutin with homologous sequences suggested that the lack of inhibitory activity of wild type lupin protein might be due to a sequence deletion in the first of two main functional loops (IL1 and IL2) of the inhibitor's structure (Scarafoni et al., 2010), which are crucial for the interaction with GHs (Sansen et al., 2004; Yoshizawa et al., 2012). The experimental 3D structures of soybean Bg7S (PDB ID: 1UAP) and *Lupinus angustifolius* γ-conglutin (PDB ID: 4PPH) later confirmed the structural variations in the two inhibition loops (Yoshizawa et al., 2011; Czubinski et al., 2015).

Sequence alignments of γ-conglutin, XEGIPs and TAXI-I evidenced interesting local variations (Supplementary Fig. S2). In general, γ-conglutin regions located upstream and downstream the deletion in IL1 loop, show the highest sequence similarity with XEGIPs proteins. Conversely, the sequence of the IL2 loop of γ-conglutin is more similar to the sequence of the IL2 loop of TAXI-I, rather than to the one of XEGIPs. In particular, a His residue (His_{374} in TAXI-I sequence), considered a key amino acid for the inhibitory activity of TAXI-I (Sansen et al., 2004; Pollet et al., 2009), is also present in γ-conglutin sequence (Scarafoni et al., 2010) but not in XEGIPs. Structural studies of the complex EDGP:XEG evidenced two critical Arg residues in both IL1 and IL2 (at position 322 and 403 of EDGP sequence), which are conserved in many XEGIP homologs (Jones, 2012; Yoshizawa et al., 2012). The first Arg residue is not present in γ-conglutin because it is located in the deleted region, whereas the second is replaced with a His residue at position 406 (analogous to His_{374} in TAXI-I sequence). Furthermore, Leu_{202} and Pro_{203} of EDGP, which have been proved to be involved in stabilizing the interaction with the enzyme (Yoshizawa et al., 2012), are conserved in most XEGIPs, but not in TAXI-I (Sansen et al., 2004) and γ-conglutin.

On these bases we generated four γ-conglutin mutants (Fig. 1). The first two mutants were designed to confer XEGIP and TAXI-I features to the unmodified recombinant γ-conglutin (rγc), by adding the respective characteristic amino acid stretches to fill the deletion in IL1. The first mutant (rγcMuX), has been designed according to NEC4 sequence (Naqvi et al., 2005), whereas the second mutant (rγcMuT), has been designed according to TAXI-I sequence (Sansen et al., 2004). Two other mutants have been prepared from rγcMuX and rγcMuT sequences, in which His_{406} residue of IL2 was replaced with a Ser residue (rγcMuX-H/S and rγcMuT-H/S, respectively), to assess the role of this amino acid in rγc.

Unexpectedly, when rγc was tested, inhibitory activity against a GH11 glucanase, but not against GH12, was observed (Table 1). The level of inhibition was about 15% (pH 5.3) and 46% (pH 7.0). The two experimental conditions were adopted for the following reasons: the acid pH is optimal for GH activities and γ-conglutin assumes prevalently a monomeric form, whereas at neutral pH, γ-conglutin takes on an hexameric quaternary structure (Capraro et al., 2010; Czubinski et al., 2015) and GH is still active, even if at about half of its maximum activity (not shown).

(A)

```
                326
                 |
rγc       CYDTKKISG-----GVPSVDLIM
Nec4      CFDSRTIVSTRVGPAVPQIDLVL

γcMuX     CYDTKKIVSTRVGPAVPSVDLIM
```

(B)

```
                326
                 |
rγc       CYDTKKISG------GVPSVDLIM
TAXI-I    CYDTKT-LGNNLGGYAVPNVQLGL

γcMuT     CYDTKKILGNNLGGYAVPSVDLIM
```

(C)

```
              404
               | *
rγc       SLKSHGKSCSNLF
Nec4      SILFRQTTCAN-F

γcMuX-H/S SLKSSGKSCSNLF
```

(D)

```
              404
               | *
rγc       SLKSHGKSCSNLF
TAXI-I    SRLPHFTGCGGL-

γcMuT-H/S SLKSSGKSCSNLF
```

Fig. 1. Amino acid sequences of γ-conglutin mutants prepared for this work. The original γ-conglutin sequence (rγc) is indicated in the first row of each alignment. (A): mutation at IL1 of rγc by addition of the sequence of *N. benthamiana* XEGIP Nec4 (Naqvi et al., 2005), giving γcMuX mutant. (B): mutation at IL1 of rγc by addition of the TAXI-I sequence (Sansen et al., 2004), originating γcMuT mutant. (C): mutation at IL2 level in which the original His residue of γcMuX was changed to Ser. (D): mutation at IL2 in which the original His residue of γcMuT was changed to Ser. The resulting mutated sequences are indicated in bold.

Table 1
Inhibitory activities of wild-type recombinant γ-conglutin (rγc) and of its four mutants, tested against *Thricoderma longibrachiatum* xylanase (GH11) and *Aspergillus aculeatus* XEG (GH12), at two different pHs.

Protein	Activity inhibition (%)			
	GH11		GH12	
	pH 5.3	pH 7.0	pH 5.3	pH 7.0
rγc	15.2 ± 3.5	45.9 ± 10.1	n.d.[a]	n.d.[a]
γcMuX	49.6 ± 16.3	13.4 ± 6.8[a]	n.d.[a]	n.d.[a]
γcMuT	39,8 ± 4.7	50,3 ± 11.1	n.d.[a]	n.d.[a]
γcMuX-H/S	n.d.[a]	n.d.[a]	n.d.[a]	n.d.[a]
γcMuT-H/S	n.d.[a]	n.d.[a]	n.d.[a]	n.d.[a]

[a] measured enzyme activities were not significantly different ($P \geq 0.05$), respect to the control assay (without inhibitor). n.d.: not detectable.

The tests on rγcMuX and rγcMuT proteins revealed that the mutations allowed to increase the inhibitory capacity of rγc against GH11 to about 50% (rγcMuX) and 40% (rγcMuT), when enzyme activity was tested at pH5.3. At pH 7.0, rγcMuT showed essentially the same inhibition level of rγc, whereas rγcMuX was much less efficient, being the inhibition level limited to about 13% (Table 1). Activity against GH12 was again not observed. From this data, it is clear that the presence of the inserted sequences is not the main reason for the inhibitory activity appearance.

Both rγcMuX-H/S and rγcMuT-H/S mutant proteins are devoid of the inhibitory capacity of rγc (Table 1), meaning that His_{406} is required for a stable interaction between the inhibitor and the target enzyme.

Size exclusion chromatography was used to verify the stable interaction between rγc and the enzyme (Fig. 2). In this experimental setting the adopted inhibitor:enzyme molecular ratio was 1:1. Seed γ-conglutin was also tested as a control (chromatogram A). While the seed-extracted protein did not form any stable complex with the enzyme (chromatogram C), rγc interacted stoichiometrically with the GH11 (chromatogram E). However, a small fraction of the two proteins remained unbound.

The reasons of these findings are clearly structural and the post-translational modifications (PTMs) occurring to this kind of proteins may have a role. Although the variation in PTMs of GHIPs has been previously discussed, their functional role in GHIPs remains controversial (Scarafoni et al., 2010; Yoshizawa et al., 2011). As shown in Fig. 3, rγc was not proteolytically cleaved into subunits, since only a single polypeptide of 45 kDa is visible and no bands of 30 kDa and 17 kDa, corresponding to the processed subunits, are present. As expected, the four rγc mutants were not processed into subunits by proteolysis (Fig. 3). Therefore, γ-conglutin is able to inhibit a glucanase only in the unprocessed form.

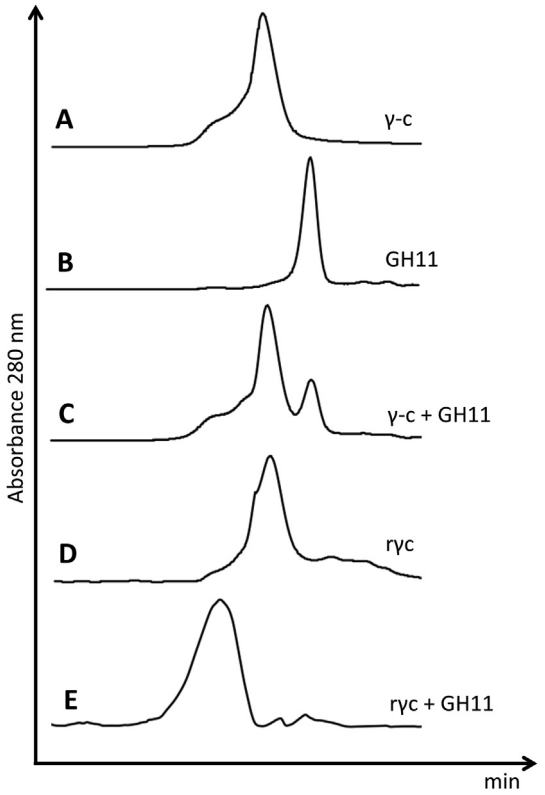


Fig. 2. Interaction of seed extracted (γ-c) and recombinant (rγc) γ-conglutins with GH11 xylanase as determined by SEC. Pure γ-c (A), pure rγc (D) and mixtures of the two with G11 xylanase in 1:1 molar ratio (C and E) were chromatographed. Pure GH11 were also analysed (B).

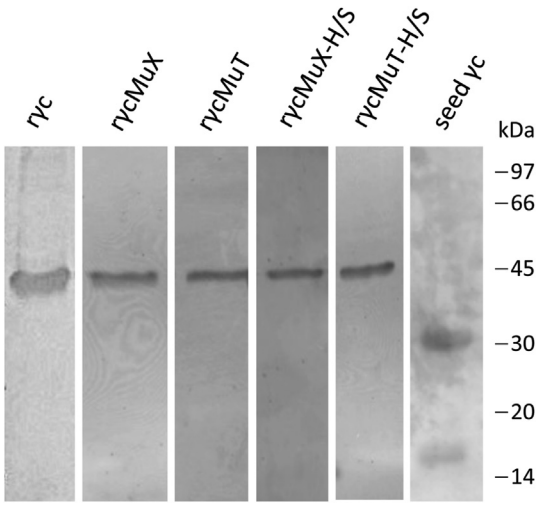


Fig. 3. Western blotting analysis of recombinant γ-conglutin (rγc) and of its mutants (rγcMuX, rγcMuT, rγcMuX-H/S and rγcMuT-H/S) prepared for this work. γ-Conglutin extracted from lupin seed has been also included. Mrs scaling is reported on the right.

All active XEGIPs consist of a single, not processed polypeptide chain (Qin et al., 2003; Shang et al., 2004; Naqvi et al., 2005), whereas wheat TAXI-I exists in two isoforms, namely form A (one polypeptide of 40 kDa) and form B (two subunits of 29 and 11 kDa linked by a disulphide bridge). Both isoforms have comparable inhibition properties and activities (Gebruers et al., 2004, 2010). In this case, the processing has not significant consequences on the structural characteristics, since it has no effect on the overall architecture of TAXI-I, insomuch as both molecular forms can coexist in TAXI-I crystals (Sansen et al., 2004). The 3D structure of *Lupinus albus* γ-conglutin has not been yet determined, but the model of the *L. angustifolius* homologous protein is available (Czubinski et al., 2015). The two lupin protein have more than 90% amino acid s.i. (Supplementary Fig. S3, A). *L. angustifolius* γ-conglutin structure refers to the seed-purified protein, which is also proteotically cleaved (Czubinski et al., 2015).

The 3D structures of *Lupinus augustifolius* γ-conglutin and Bg7S monomers, EDGP and TAXI-I are largely overlapping, a fact which reflects the general conserved structural conformation among this kind of GHIPs, regardless the PTM processing (Yoshizawa et al., 2012; Czubinski et al., 2015). IL1 and IL2 of EDGP and TAXI-I show a very similar spatial location. Conversely, superimposition of *L. angustifolius* γ-conglutin structure with EDGP and TAXI-I structures revealed that while IL2s overlap almost completely, IL1s lie with a different geometry in lupin protein (Czubinski et al., 2015), as direct consequence of the sequence deletion.

We aimed to evidence possible structural repercussions due to the lack of post translational proteolytic cleavage on the rγc structure and, in the attempt, an *in silico* analysis has been carried out by using an homology modelling approach. Rγc was modelled using both EDPG (PDB ID: 2VLA) and TAXI-I (PDB ID: 2B42) as templates. Representations of the two predictive models are reported in Fig. 4A and B, respectively (green-coloured). Both structures have been superimposed to that of *L. angustifolius* γ-conglutin (PDB ID: 4PPH, drawn in red). Even considering the intrinsic limitations of this approach, in view of the overall large overlapping, it stands out that in both cases IL1 was predicted to appear spatially closer to IL2 in the unprocessed rγc than in the proteolytically processed *L. angustifolius* γ-conglutin. This open the way to argue that the lack of PTM processing compel the two functional loops to lie in a position which may, at least in part, compensate the sequence deletion in IL2.

It is worth noting that the two wild-type legume proteins are, amongst all the homologues, not only inhibitory incompetent, but are the sole able to assembly to a quaternary structure, tetrameric in Bg7S and hexameric in γ-conglutin (Yoshizawa et al., 2011; Czubinski et al., 2015). Light scattering experiments (Fig. 5) indicated that rγc acquires a quaternary structure comparable to that of the natural protein (Capraro et al., 2010). A form of about Mr 240 kDa, compatible with a hexameric quaternary assembly, is prevalent at pH 7.0, whereas the protein is completely disassembled to monomers below pH 5.0 (approx. calculated Mr: 40 KDa). Most likely, the proteolytic processing influences the structural conformation of γ-conglutin and small local rearrangements may be the cause of the observed activity. Despite the fact that the general fold is preserved, the structure of γ-conglutin is particular, mainly because of the presence of a unique curved helix H4N (Czubinski et al., 2015), which suggests specific anchoring at target cellular structures, such as membranes (Czubinski et al., 2015) or at structures of other proteins. Moreover, the protomers of γ-conglutin undergo a completely different quaternary assembly than those of Bg7S protomers, because of some amino acid substitutions at interface regions (Czubinski et al., 2015).

4. Conclusions

Despite the body of information available, including the 3D structures of some GH:GHIP complexes many structural aspects of the inhibition mechanisms and specificity of the XEGIP inhibitors still remain unclear and deserve further investigation. Due to its peculiar molecular features, γ-conglutin may be used as a model to deepen the understanding of the structure-function relationships of this kind of proteins. Authors suggested that lupin γ-conglutin and soybean Bg7S might represent a separate group of endo-β-glucanases inhibitor (Yoshizawa et al., 2011; Czubinski et al., 2015). XEGIPs and TAXIs have been found widely distributed across the plant kingdom (Qin et al., 2003), but homologous proteins proteolytically processed have been found only in legume seed (Kagawa et al., 1987).

Our results indicated that, despite the overall higher s.i. with active XEGIPs proteins, the specificity of action of rγc resembles that of TAXI-I because the presence of a conserved key His residue. The presence of IL1 is not strictly required to manifest inhibition, even if the specifically inserted amino acid stretches enhanced the activity. GH11 and GH12 enzymes have a similar fold pattern, with the substrate-binding region shaped as a groove formed by a concave β-sheet, suggesting that elements of inhibitor proteins involved in the inhibition mechanism should have common characteristics (Sansen et al., 2004; Czubinski et al., 2015). On the other hand, H_{406} in IL2 is confirmed to be necessary and sufficient to manifest the inhibitory competence of rγc, although the acquired activity is limited. Interestingly, a His residue is also present in the same position of BG7S sequence (Supplementary Fig. S3, B). On the contrary, in *L. angustifolius* γ-conglutin the His residue is substituted by a Tyr residue. Preliminary results indicate that γ-conglutin extracted from *L. angustifolius* seed lack of inhibitory activity against GH11 (not shown). If this might be attributable to the amino acid substitution or to proteolytic processing it remains to be investigated.

On the whole, the results provide new insights about structural characteristics at the basis of the lack of inhibitory activity of wild-type γ-conglutin and on the mechanisms of family-specific inhibition of GH11 and GH12 by GHIPs. At the same time, they raise intriguing evolutionistic questions about the significance of the proteolytic processing which these kind of seed legume proteins undergo and about the possible origin of the their target enzymes.

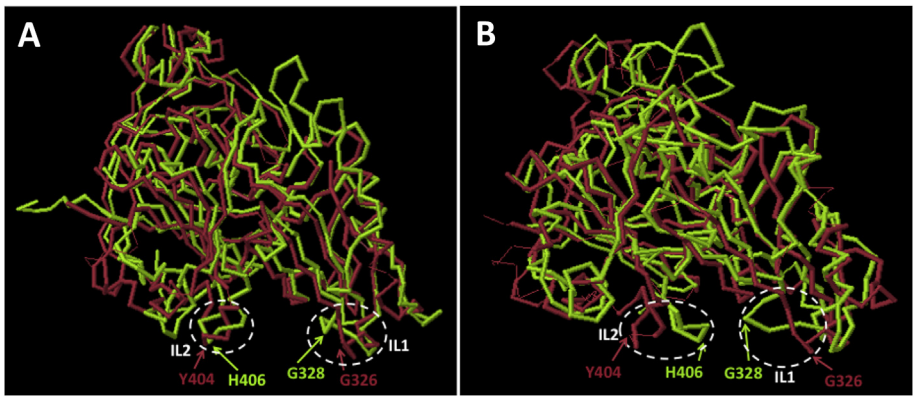


Fig. 4. Predictive 3D structures of rγc (reported as green backbone), obtained by homology modelling using carrot EDGP (Panel A) or and wheat TAXI-I (Panel B) proteins as templates (PDB ID: 2VLA and 2B42, respectively). Both models have been superimposed with *L. angustifolius* γ-conglutin (PDB ID: 4PPH) structure (in red). Functional region IL1 and IL2 are encircled by a white dashed line. In loop IL1 the arrows indicate the position of the amino acid stretch deletion. In loop IL2 the green arrow indicate His_{406} residue of rγc. In the same position of *L. angustifolius* γ-conglutin Tyr_{404} is instead present (red line). See text for details. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

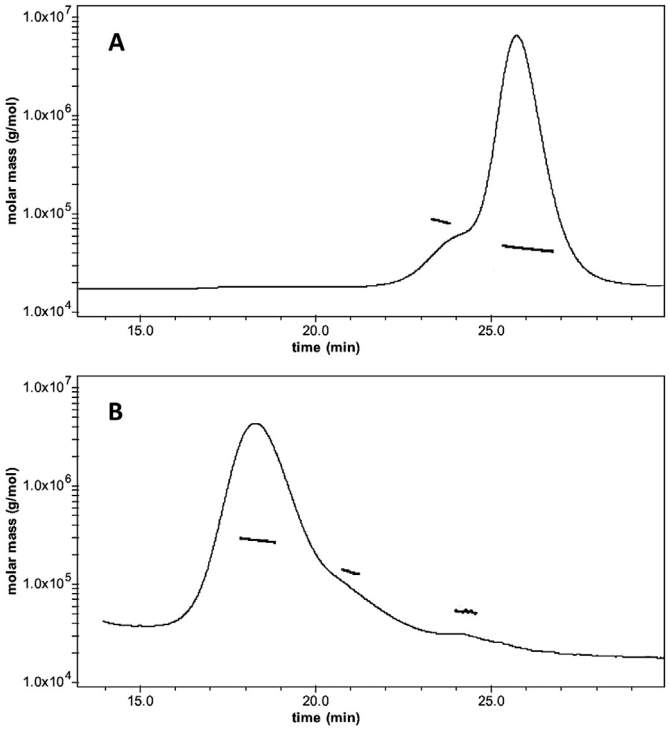


Fig. 5. Chromatograms of rγc SEC-separated at pH 4.5 (A) and at pH 7.5 (B), and analysed by DLS. The calculated molecular masses are reported as bold line segments, with the scaling values on the left axes. Protein elutions were recorded at 280 nm.

Contributions

Conceived and designed the experiments: AS, MD.
Performed the experiments: AC, SP, SB, JC.
Analysed data: AS, AC, SP, SB, JC, MD.
Bioinformatic and modelling analysis: AS, EG.
Contributed reagents/materials/analysis tools: AS, MD.
Wrote the paper: AS.

Acknowledgements

J.C. was supported by a grant "Dote Ricerca" of Regione Lombardia and European Social Fund. We are indebted with Fabio Sessa for the invaluable technical assistance in chromatographic separations.

Appendix A. Supplementary data

Supplementary data related to this article can be found at http://dx.doi.org/10.1016/j.plaphy.2015.11.008.

References

Arnold, K., Bordoli, L., Kopp, J., Schwede, T., 2006. The SWISS-MODEL Workspace: a web-based environment for protein structure homology modelling. Bioinformatics 22, 195–201.
Biasini, M., Bienert, S., Waterhouse, A., Arnold, K., Studer, G., Schmidt, T., Kiefer, F., Gallo Cassarino, T., Bertoni, M., Bordoli, L., Schwede, T., 2014. SWISS-MODEL: modelling protein tertiary and quaternary structure using evolutionary information. Nucl. Ac. Res. 42, W252–W258.
Bradford, M.M., 1976. A rapid and sensitive method for the quantitation of micrograms quantities of proteins utilising the principle of protein-dye binding. Anal. Biochem. 72, 248–254.
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. Nucl. Acids Res. 37, D233–D238.
Capraro, J., Spotti, P., Magni, C., Scarafoni, A., Duranti, M., 2010. Spectroscopic studies on the pH-dependent structural dynamics of γ-conglutin, the blood glucose-lowering protein of lupin seeds. Int. J. Biol. Macromol. 47, 502–507.
Choi, H.W., Kim, N.H., Lee, Y.K., Hwang, B.K., 2013. The pepper extracellular xyloglucan-specific endo-β-1,4-glucanase inhibitor protein gene, CaXEGIP1, is required for plant cell death and defense responses. Plant Physiol. 161, 384–396.
Consonni, A., Lovati, M.R., Manzoni, C., Pizzagalli, A., Morazzoni, P., Duranti, M., 2010. Cloning, yeast expression, purification and biological activity of a truncated form of the soybean 7S globulin alpha' subunit involved in Hep G2 cell cholesterol homeostasis. J. Nutr. Biochem. 21, 887–891.
Consonni, A., Lovati, M.R., Parolari, A., Manzoni, C., Morazzoni, P., Magni, C., Duranti, M., 2011. Heterologous expression and purification of the soybean 7S globulin α' subunit extension region: *in vitro* evidence of its involvement in cell cholesterol homeostasis. Protein Expr. Purif. 80, 125–129.
Czubinski, J., Barciszewski, J., Gilski, M., Szpotkowski, K., Debski, J., Lampart-Szczapa, E., Jaskolski, M., 2015. Structure of γ-conglutin: insight into the quaternary structure of 7S basic globulins from legumes. Acta Crystallogr. D. Biol. Crystallogr. 71, 224–238.
Duranti, M., Consonni, A., Magni, C., Sessa, F., Scarafoni, A., 2008. The major proteins of lupin seed: characterisation and molecular properties for use as functional and nutraceutical ingredients. Trends Food Sci. Technol. 19, 624–633.
Fierens, K., Geudens, N., Brijs, K., Courtin, C.M., Gebruers, K., Robben, J., Van Campenhout, S., Volckaert, G., Delcour, J.A., 2004. High-level expression, purification, and characterization of recombinant wheat xylanase inhibitor TAXI-I secreted by the yeast *Pichia pastoris*. Protein Expr. Purif. 37, 39–46.
Fierens, K., Gils, A., Sansen, S., Brijs, K., Courtin, C.M., Declerck, P.J., De Ranter, C.J., Gebruers, K., Rabijns, A., Robben, J., Van Campenhout, S., Volckaert, G., Delcour, J.A., 2005. His374 of wheat endoxylanase inhibitor TAXI-I stabilizes complex formation with glycoside hydrolase family 11 endoxylanases. FEBS J.

Original Study **Open Access**

Jessica Capraro*, Elisabetta Galanti, Mauro Marengo, Marcello Duranti and Alessio Scarafoni

Cysteine-containing peptides are produced by sequential clipping, but not released, from lupin 11S storage globulin during early germination

DOI 10.1515/ped-2015-0005
Received August 23, 2015; accepted October 30, 2015

Abstract: The digestion of the seed storage proteins is a finely regulated process operated by several proteases whose action is influenced by the exposure of specific regions, which became progressively available upon their action. We focused our study on the initial stages of germination, where more subtle modifications to the storage proteins are expected. Small-size peptides containing cysteine residues and other possible metal-binding regions are *de facto* produced but are not released from the "parental" protein since they remain bound trough disulphide bridges. The meaning of these findings is discussed.

Keywords: Seed storage proteins; proteolysis; germination.

1 Introduction

In mature legume seeds the majority of proteins belong to the so-called "Storage Proteins" (SP). They have no enzyme activities and act as amino acid reserves destined to support the nascent seedling. In legume seed SPs are globulins usually classified according to their sedimentation coefficients (S) as 7S (also named as vicilins) and 11S (known with the trivial name of legumins). The 11S globulins are hexameric proteins of about 300 kDa. Each monomer is made by two subunits, which origin by a single precursor [1]. As all SPs, they are synthesized during seed developing and deposited inside the cells in specific membrane-bound organelles, called Protein Bodies (PBs), in which a variety of proteins and enzyme are also stored [1,2].

Lupin 11S globulin accounts up to 35% of the total SPs and is formed by polypeptides that originated from Post-Translational Proteolysis (PTP) at a single cleavage site of few precursors polypeptides of about 60 kDa, occurring during seed development [3]. This PTP thus forms two polypeptide chains, also referred to as acidic (α) and basic (β) chains, linked by disulphide bonds, which represent the legumin monomer. In mature *Lupinus albus* seeds are present α-polypeptides with M*r*s ranging from 42–52 kDa and β polypeptides of 20–22 kDa [4]. Such heterogeneity arises from the action of proteases acting during seed filling. The extent of this processing has no observed parallel in other plant species, including pea and soybean, where the endogenous cleavage of legumins during seed developing has been described previously [5].

Seed germination is a complex physiological process where many biochemical processes are initiated or resumed following water imbibition [6]. SPs are hydrolysed up to free amino acids to support the seedling growth. The complete degradation of storage proteins takes place inside the storage organelle. This process is usually triggered by pre-existing and by newly formed proteases [3,7]. Limited proteolysis that triggers unlimited degradation of storage globulins are catalysed by low specificity endopeptidases [3]. The final products of the action of the proteases are free amino acids and small peptides, that in dicots, are transported outside of the PBs into the cytosol, where the peptides will be degraded further to amino acids by the action of aminopeptidases [1]. Structural features of SPs prevent cleavage by proteinases that are simultaneously present in the same compartment. It is generally accepted that the structural compactness of SPs is an important feature for their controlled proteolysis. Disulphide bridges, which are formed between the domains of legumin precursor, stabilize the conformation of mature legumin subunits [1,3]. Cleavage of susceptible

***Corresponding author Jessica Capraro,** Department of Food, Environmental and Nutritional Sciences (DeFENS), Università degli Studi di Milano, via G. Celoria 2, 20133 Milano, Italy,
E-mail: jessica.capraro@unimi.it
Elisabetta Galanti, Mauro Marengo, Marcello Duranti and Alessio Scarafoni, Department of Food, Environmental and Nutritional Sciences (DeFENS), Università degli Studi di Milano, Milano, Italy

sites causes the destabilization of the tertiary structure of storage globulins and bestows susceptibility to unlimited proteolysis [2].

Here we show the results about the effects of the early proteolytic events during germination occurring to lupin legumin, that lead to the production of peptides involved in disulphide bridges. Consequences on the metal-binding capacities of the protein are also discussed.

2 Material and Methods

L. albus. seeds (cv. Multitalia) were generously supplied by Dr. A. Conocchiari (Agroservice S.p.A., Italy). All chemicals were from Sigma-Aldrich (Milano, Italy) unless otherwise indicated. Protein concentrations were determined according to the Bradford assay [8], using bovine serum albumin as a standard. Before germination, seeds were surface-sterilized with 0.2% sodium hypochloride for 20 min, rinsed thoroughly with distilled water, vernalized overnight at 4°C, and then germinated at 22°C for 48, 72 and 96 hours. At the end of the incubation time, the seeds were immediately frozen with liquid nitrogen and kept at -80°C.

Purification of lupin 11S globulin (legumin) from mature and germinated seeds was performed using a previously described procedure based on anion exchange chromatography separation [7]. To prevent hydrolysis during the procedure, proteins were extracted in the presence of a protease inhibitors cocktail. At the end of the purification, the protein was extensively dialyzed against water, lyophilized and kept at -20°C in a sealed jar until used.

Reactive cysteine thiols were measured essentially according to Iametti et al. [9]. Protein samples were dissolved (1 mg mL^{-1}) in solution containing 8 M urea and, when indicated, 10 mM DTT. After 10 min incubation at room temperature, excess DTT was removed by rapid gel filtration on syringe-type Sephadex G-25 columns. DTNB (0.9 mL, 2 mM in 50 mM sodium phosphate buffer, pH 7.5) was added to 0.1 mL of the eluted protein. After 15 min incubation at room temperature, solutions were centrifuged for 5 min at 12000 × *g*, and the absorbance of the supernatants was red at 412 nm. Each sample has been analysed in triplicate.

SDS-PAGE was carried out on 12% or 16% polyacrylamide gels according to Laemmli [10], under reducing conditions, using a MiniProtean III electrophoresis apparatus (Bio-Rad). Runs were carried out at constant 16 mA for each gel. Coomassie Brilliant Blue or silver staining [11] was used for gel staining as indicated. Molecular weight markers were LMW-SDS Marker Kit and Low-range Amersham Rainbow Marker from GE Healthcare (Milan, Italy).

Analysis of the oligomeric structures of legumin were carried out by size exclusion chromatography using a Waters HPLC (model 625, Milford, MA) and a Superose 12 HR 10/30 column (GE Healthcare) equilibrated in 50 mM sodium phosphate buffer, pH 7.5, containing 0.1 M NaCl. The calibration curve for M*r* assessment was set using thyroglobulin (670 kDa), amylase (200 kDa), glucose oxidase (180 kDa), alcohol dehydrogenase (150 kDa), transferrin (76 kDa), bovine serum albumin (67 kDa) and egg albumin (45 kDa) as standard protein markers. Protein elution was monitored at 280 nm.

3 Results

Lupin legumin was purified from germinating seeds at 48, 72 and 96 hrs after the onset of germination. Dry seeds were used as control. The SDS-PAGE patterns are shown in Figure 1A. Their most striking was the appearance of a band of about 60 kDa, that rapidly appeared and started to fade between 72 and 96 hrs. This polypeptide is most likely the legumin precursor, not yet processed, which origin from the translation of stored mRNA, namely pre-existing messenger trapped in the cotyledonary tissue during the desiccation of the mature seed, as previously reported for lupin [5]. The group of polypeptides with M*r* between 40 and 50 kDa, namely the legumin α chains, varied quantitatively, indicating that proteolysis took place. The appearing of apparently new polypeptides in this M*r* range represent intermediate products rather than newly expressed protein [1-3]. The β subunit of about 20 kDa remained unvaried in all samples.

Analysis of the oligomeric structures of legumin was performed by gel filtration chromatography under native conditions. The results are shown in Figure 1B. The global native arrangement of the legumin monomers was essentially maintained during the time course at the early stages of germination, since no shifts of M*r* could be detected. Only the sample collected at 96 hrs showed a modest decrease in size, indicating that a significant proteolytic degradation occurred and that some regions close to the protein surface or otherwise on it began to be trimmed off.

To assess any change of the disulphide bonds pattern during germination, we measured the reacting cysteine thiols by using the Elmann reagent [12]. The results are summarized in Table 1. Analysis of 8M-denatured, non-reduced proteins indicated the absence of free cysteines.

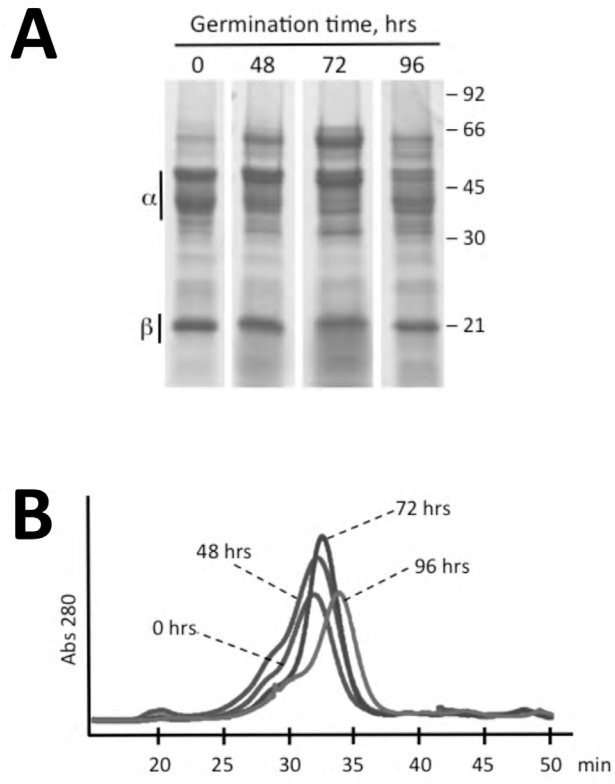


Fig. 1. Lupin 11S globulin analysis during germination time course. **A:** SDS-PAGE under reducing condition. α indicates the polypeptide family forming the major subunit. β indicate the position of the small subunit. Weight molecular markers are reported as kDa. **B:** Gel-filtration under native conditions.

Thus, all –SH groups are involved in disulphides bonds even at a late stage off germination, when hydrolysis operated by endogenous proteases is evident. Instead, when denaturation was carried out in the presence of the reducing agent, some differences among the sample became evident. Although experimental data are impaired by a non negligible standard error, it appears fairly clear that thiol groups are determined in even number in all cases.

Table 1. Number of cysteine residues in legumin protein during germination time course, assesssed under not-reducing and reducing conditions.

Germination time (hrs)	# Cys	
	- DTT	+ DTT
0	0,5 ± 0.2	6.3 ± 0.6
48	0.3 ± 0.2	4.1 ± 0.3
72	0.3 ± 0.1	3.7 ± 0.2
96	0.6 ± 0.3	2.2 ± 0.3

Figure 2 reports the electrophoretic patterns of the fractions retained by the gel-filtration spun columns. The left gel concerns the unreduced proteins, whereas the right gel shows the separation of the reduced samples. In this latter case some small-sized polypeptides are also appearing, indicating that cysteine-containing peptide may be released from 11S globulin only after reduction.

4 Discussion

Dramatic changes occur within the seed's tissues during germination, when anabolic and catabolic events inherently overlap and interplay. The digestion of the seed storage proteins is a key process aimed to provide free amino acids essential to support the seedlings growth. This complex process is operated by pre-existing and *de novo* synthesized proteases [1-3]. We focused our study on the initial stages of germination, where more subtle modifications to the storage proteins are expected, to contribute in shedding some light on the possible structural repercussions occurring to legumin during proteolysis. As a matter of facts, small structural modifications are at the basis of the controlled degradation of the storage reserves of the germinating seed. It has been previously shown that the first step of the lupin legumin degradation is the selective removal of a peptide of about 3 kDa located at the surface of the protein, due to the action of a highly specific protease cutting at a double arginine sequence. The core of this peptide is formed by two consecutive stretches, the first made of six histidine residues and the second of nine glutamic acids [7]. The loss of this peptide apparently did not cause major structural variations in the compactness of the legumin, as evidenced by the chromatographic determination shown in Figure 1B. Due to its peculiar amino acid composition, this peptide is likely located at the surface of the native protein, being charged negatively at physiological pH. It has been postulated that it may be involved in metal binding, acting as a carrier for divalent ions [7], once it is released from the "parental" protein.

In the present work we followed the fate of the peptides originating by the proteolytic processing located around the cysteine residue, all involved in disulphide bridges in the native legumin. According to the deduced primary structure of Lupin legumin (Figure 3), all cysteine of the β polypeptide must be involved in inter-chain bonds. Thus, intra-chain bonds may be found only in the major subunits of the monomers. It has been previously showed, and here we confirm (Figure 1A), that the β subunit is hydrolysed only in the final stages of germination [2,4]. It follows that

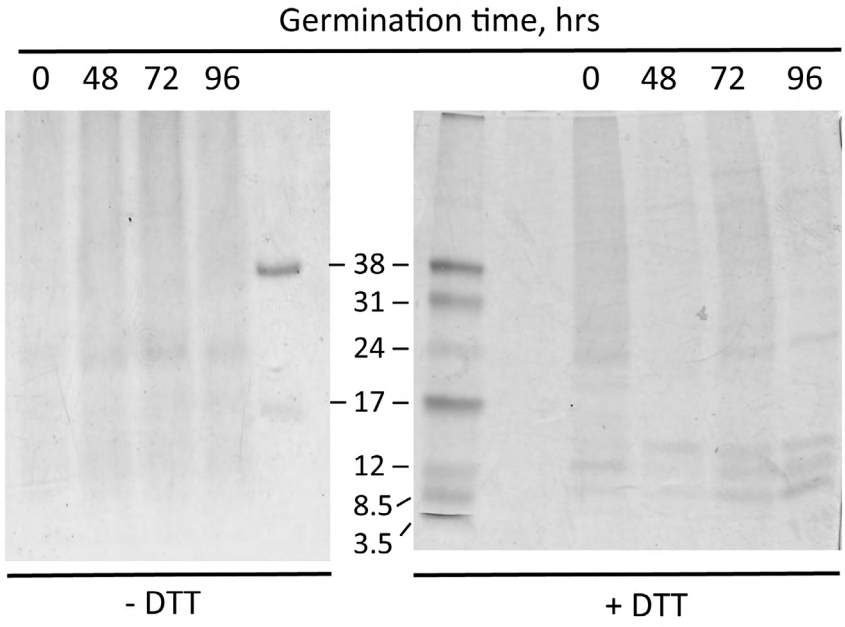


Fig. 2. SDS-PAGE analysis of the retained fraction obtained by gel filtration from the purified 11S globulin during germination time course. On the left gel the proteins were denatured with urea before the chromatography, whereas the proteins on the second gel were treated with urea and DTT.

```
STFRQQPQENECQFQRLNALEPDNTVQSEAGTIETWNPKNDELRCAGVALSRCTIQRNGL   60
RRPFYTNAPQEIYIQQGRGIFGMIFPGCGETYEEPQESEKGQGPRPQDRHQKVEHFKEGD  120
IIAVPTGIPFWMYNDGQTPVVAITLIDTTNLDNQLDQIPRRFYLSGNQEQEFLQYQEKEG  180
GQGQQQEGGNVLSGFDDEFLEEALSVNKEIVRNIKGKNDDREGGIVEVKGGLKVIIPPTM  240
RPRHGREEEEEEEEEDERRGDRRRRHPHHHHHEEEEEEEEEWSHQVRRVRRPHRHHHHRK  300
DRNGLEETLCTMKLRHNIGESTSPDAYNPQAGRFKTLTSIDFPILGWLGLAAEHGSIYKN  360
ALFVPYYNVNANSILYVLNGSAWFQVVDCSGNAVFNGELNEGQVLTIPQNYAAAIKSLSD  420
NFRYVAFKTNDIPQIATLAGANSEISALPLEVVAHAFNLNRDQARQLKNNNPYKFLVPPP  480
QSQLRAVA                                                       488
```

Fig. 3. Deduced amino acid sequence of lupin 11S globulin precursor (UniProtKB: Q53I54). The signal peptide has been removed. Black blocks indicate the cysteine residues. Residues in bold (at positions 110, 112 and 118) indicate a possible metal-binding active site typical of cupin proteins. The cleavage site originating acidic and basic subunits of the 11S globulin precursor is underlined. The basic subunit is reported in italics.

the peptides involved in disulphide bonds that we have shown to be trimmed derive from the α subunit.

Measurement of reactive thiols (Table 1 and Figure 2) indicated that no release of peptides from the parental protein occurs, since they remain bound to the remainder of the protein through disulphide bridges. It has been shown that in several cases reduction of seed proteins during germination is conducted by *ex novo* synthesized thioredoxin h [13]. It is likely that in our case a possible action of thioredoxin h became evident at later times. The reduction of disulphide groups of wheat and rice SPs has been shown to increase progressively during seed germination [14]. Thus, reduction of the disulphide bridges may also represent a key event in the regulation of the legumin breakdown. Our results indicate that cysteine-containing peptides are produced sequentially. The first peptide appears within 48 hrs from the onset of germination, and the second is appearing later, typically between 72 and 96 hrs. It remain to be determined if the first event is necessary to trigger the production of the second peptide or whether the two have no mutual relationship and thus other proteolytic events must take place in due time.

In homologous legumin proteins from different origin, the intra-chain bonds are positioned in slightly

buried regions compared with the inter-chain bond. In addition, the inter-chain disulphide bonds are located on the interface between protomers. These bonds could contribute to structural stability and folding of the protein.

Legumins belong to the cupin superfamily, which present a typical conserved β-barrel fold. Studies based on sequence alignment reveal that most plant cupins contain a single metal-binding site, characterized by two conserved histidines and a glutamic acid, able to bind manganese ions [15]. Although legumins are not the most important proteins with regard to metal-binding capacity, the lupin one presents some molecular features that indicate potential metal-binding sites other that the mentioned histidine/glutamate-rich peptide [7]. Lupin legumin (Figure 3) shows a conserved cupin metal-binding site located close to one cysteine in the major subunit, in which the second hystidine is substitute by a lysine. Such sequence variation is also present in few other legumin sequences found in peanut, soybean and fababean [16] and could alter the binding features of the site.

Acknowledgements: J.C. was supported by a grant "Dote Ricerca" of Regione Lombardia and European Social Fund.

Contributions: Conceived and designed the experiments: JC, AS. Performed the experiments: AS, JC, EG, MM. Analysed data: JC, AS, EG, MD. Contributed reagents/materials/analysis tools: AS, MD. Wrote the paper: AS, JC.

References

[1] Muentz K., Intracellular protein sorting and the formation of protein reserves in storage tissue cells of plant seeds. Bioch. Physiol. Pflanzen, 1989, 185, 315-335.
[2] Shutov A.D., Vaintraub I.A., Degradation of storage proteins in germinating seeds. Phytochemistry, 1987, 26, 1557-1566.
[3] Shutov A.D., Baumlein H., Blattner F.R., Muentz K., Storage and mobilization as antagonistic functional constraints on seed storage globulin evolution J. Exp. Bot., 2003, 54, 1645-1654.
[4] Duranti M., Consonni A., Magni C., Sessa F., Scarafoni A., The major proteins of lupin seed: characterisation and molecular properties for use as functional and nutraceutical ingredients. Trends Food Sci. Technol., 2008, 19, 624–633.
[5] Scarafoni A., Ronchi A., Prinsi B., Espen L., Assante G., Venturini G., Duranti M., The proteome of exudates from germinating *Lupinus albus* seeds is secreted through a selective dual-step process and contains proteins involved in plant defense, FEBS J., 2013, 280, 1443-1459.
[6] Bewley J.D., Seed germination and dormancy, Plant Cell, 1997, 9, 1055-1066.
[7] Capraro J., Sessa F., Magni C., Scarafoni A., Maffioli E., Tedeschi, G., Duranti, M., Proteolytic cleavage at twin arginine residues affects structural and functional transitions of lupin seed 11S storage globulin, PLoS ONE (2015), 10(2): e0117406. doi: 10.1371/journal.pone.0117406.
[8] Bradford M.M., A rapid and sensitive method for the quantitation of micrograms quantities of proteins utilising the principle of protein-dye binding, Anal. Biochem., 1976, 72, 248-254.
[9] Iametti S., De Gregori B., Vecchio G., Bonomi F., Modifications occur at different structural levels during the heat denaturation of β-lactoglobulin, Eur. J. Biochem., 1996, 237, 106-112.
[10] Laemmli U.K., Cleavage of structural proteins during the assembly of the head of bacteriophage T4, Nature, 1970, 227, 660-665.
[11] Jin L.-T., Hwang S.-Y., Yoo G.-S., Choi, J.-K., A mass spectrometry compatible silver staining method for protein incorporating a new silver sensitizer in sodium dodecyl sulfate-polyacrylamide electrophoresis gels., Proteomics, 2006, 6, 2334–2337.
[12] Ellman G.L., A colorimetric method for determining low concentrations of mercaptans, Arch. Biochem. Biophys., 1958, 74, 443-450.
[13] Renard M., Alkhalfioui F., Schmitt-Keichinger C., Ritzenthaler C., Montrichard F., Identification and characterization of thioredoxin h isoforms differentially expressed in germinating seeds of the model legume *Medicago truncatula*, Plant Physiol., 2011, 155, 1113-1126.
[14] Hiroyuki Yano H., Wong J.H., Cho M.J, Buchanan, B.B., Redox changes accompanying the degradation of seed storage proteins in germinating rice, Plant Cell Physiol., 2001, 42, 879-883.
[15] Dunwell J.M., Culham A., Carter C.E, Sosa-Aguirre C.R., Goodenough P.W., Evolution of functional diversity in the cupin superfamily, Trends in Biochem. Sci., 2001, 26, 740-746.
[16] Jain, A.K., Cloning and structural analysis of a cDNA clone encoding glycinin (Gly-1) seed storage protein of peanut, Electr. J. Biotech., 2004, 7. DOI: 10.2225/vol7-issue3-fulltext-13.

272, 5872–5882.
Gebruers, K., Brijs, K., Courtin, C.M., Fierens, K., Goesaert, H., Rabijns, A., Raedschelders, G., Robben, J., Sansen, S., Sørensen, J.F., 2004. Properties of TAXI-type endoxylanase inhibitors. Biochim. Biophys. Acta 1696, 213–221.
Gebruers, K., Debyser, W., Goesaert, H., Proost, P., Van Damme, J., Delcour, J.A., 2001. *Triticum aestivum* L. endoxylanase inhibitor (TAXI) consists of two inhibitors, TAXI I and TAXI II, with different specificities. Biochem. J. 353, 239–244.
Gebruers, K., Fierens, K., Dornez, E., Van Campenhout, S., Courtin, C.M., Delcour, J.A., 2010. Functional xylanase inhibition activity of two molecular forms of recombinant TAXI-IA. J. Cereal Sci. 52, 516–519.
Hasegawa, H., Holm, L., 2009. Advances and pitfalls of protein structural alignment. Curr. Opin. Struct. Biol. 19, 341–348.
Holm, L., Rosenström, P., 2010. Dali server: conservation mapping in 3D. Nucl. Acids Res. 38, W545–W549.
Jones, R.W., 2012. Multiple copies of genes encoding XEGIPs are harboured in an 85-kB region of the potato genome. Plant Mol. Biol. Rep. 30, 1040–1046.
Juge, N., 2006. Plant protein inhibitors of cell wall degrading enzymes. Trends Plant Sci. 11, 359–367.
Kagawa, H., Yamauchi, F., Hirano, H., 1987. Soybean basic 7 S globulin represents a protein widely distributed in legume species. FEBS Lett. 226, 145–149.
Laemmli, U.K., 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature 227, 660–665.
Lever, M., 1972. New reaction for colorimetric determination of carbohydrates. Anal. Biochem. 47, 273–279.
Macauley-Patrick, S., Fazenda, M.L., McNeil, B., Linda, M., Harvey, L.M., 2005. Heterologous protein production using the *Pichia pastoris* expression system. Yeast 22, 249–270.
Naqvi, S.M.S., Harper, A., Carter, C., Ren, G., Guirgis, A., York, W.S., Thornburg, R.W., 2005. Nectarin IV, a potent endoglucanase inhibitor secreted into the nectar of ornamental tobacco plants. Isolation, cloning, and characterization. Plant Physiol. 139, 1389–1400.
Pauly, M., Andersen, L.N., Kaupinen, S., Kofod, L.V., York, W.S., Albersheim, P., Darvill, A.G., 1999. A xyloglucan-specific endo-β-1,4-glucanase from *Aspergillus aculeatus*: expression cloning in yeast, purification, and characterization of the recombinant enzyme. Glycobiology 9, 93–100.
Pollet, A., Sansen, S., Raedschelders, G., Gebruers, K., Rabijns, A., Delcour, J.A., Courtin, C.M., 2009. Identification of structural determinants for inhibition strength and specificity of wheat xylanase inhibitors TAXI-IA and TAXI-IIA. FEBS J. 276, 3916–3927.
Qin, Q., Bergmann, C.W., Rose, J.K., Saladie, M., Kolli, V.S., Albersheim, P., Darvill, A.G., York, W.S., 2003. Characterization of a tomato protein that inhibits a xyloglucan-specific endoglucanase. Plant J. 34, 327–338.
Sansen, S., De Ranter, C.J., Gebruers, K., Brijs, K., Courtin, C.M., Delcour, J.A., Rabijns, A., 2004. Structural basis for inhibition of *Aspergillus niger* xylanase by *Triticum aestivum* xylanase inhibitor-I. J. Biol. Chem. 279, 36022–36028.
Scarafoni, A., Di Cataldo, A., Vassilevskaia, T.D., Bekman, E.P., Rodrigues-Pousada, C., Ceciliani, F., Duranti, M., 2001. Cloning, sequencing and expression in the seeds and radicles of two Lupinus albus conglutin γ genes. Biochem. Biophys. Acta 1519, 147–151.
Scarafoni, A., Ronchi, A., Duranti, M., 2010. γ-Conglutin, the *Lupinus albus* XEGIP-like protein, whose expression is elicited by chitosan, lacks of the typical inhibitory activity against GH12 endo-glucanases. Phytochemistry 71, 142–148.
Scarafoni, A., Ronchi, A., Prinsi, B., Espen, L., Assante, G., Venturini, G., Duranti, M., 2013. The proteome of exudates from germinating *Lupinus albus* seeds is secreted through a selective dual-step process and contains proteins involved in plant defense. FEBS J. 280, 1443–1459.
Shang, C., Shibahara, T., Hanada, K., Iwafune, Y., Hirano, H., 2004. Mass spectrometric analysis of posttranslational modifications of a carrot extracellular glycoprotein. Biochemistry 43, 6281–6292.
Toerroenen, A., Mach, R.L., Messner, R., Gonzalez, R., Kalkkinen, N., Harkki, A., Kubicek, C.P., 1992. The two major xylanases from *Trichoderma reesei*: characterization of both enzymes and genes. Biotechnology 10, 1461–1465.
Towbin, M., Staehelin, T., Gordon, J., 1979. Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets: procedure and some applications. Proc. Natl. Acad. Sci. U.S.A. 76, 4350–4354.
Yoshizawa, T., Shimizu, T., Hirano, H., Sato, M., Hashimoto, H., 2012. Structural basis for inhibition of xyloglucan-specific endo-β-1,4-glucanase (XEG) by XEG-protein inhibitor. J. Biol. Chem. 287, 18710–18716.
Yoshizawa, T., Shimizu, T., Yamabe, M., Taichi, M., Nishiuchi, Y., Shichijo, N., Unzai, S., Hirano, H., Sato, M., Hashimoto, H., 2011. Crystal structure of basic 7S globulin, a xyloglucan-specific endo-β-1,4-glucanase inhibitor protein-like protein from soybean lacking inhibitory activity against endo-β-glucanase. FEBS J. 278, 1944–1954.