Lignocellulose constitutes a raw material with considerable potential for the production of fermentable sugars and the generation of biofuels. In nature, lignocellulosic waste from forestry, agriculture and gardening acts as the preferred carbon source of a number of bacteria and fungi endowed with the required ligninolytic machinery. These hydrolytic activities could potentially be complemented with those from other proteins that remodel the structure of cell walls, such as expansins, a group of proteins originally identified in plants that have the capacity of relaxing cell wall tension to allow cell growth. Expansins participate in processes where remodeling of the plant cell wall is required: organogenesis, fruit ripening, and growth of the pollen tube, among others. Expansins and expansin-like proteins have been proposed to act by disrupting the hydrogen bonds binding together cellulose fibrils and cellulose and other polysaccharides through a non-enzymatic process, enhancing subsequent degradation. In this manuscript, a review on plant cell wall composition and ligninolytic enzymes from cell wall-degrading bacteria and fungi is presented. Proteins with expansin-like activity, their properties and their potential application to enhance sugar release from lignocellulosic material are also reviewed.

Keywords: lignocellulose, cellulases, expansins, fungi, biofuels

Introduction

Lignocellulose, the principal and most abundant component of the renewable biomass produced by photosynthesis, is synthesized at an estimated rate of some 200 billion tons per year [1-2]. This biopolymer is formed principally by cellulose, hemicelluloses and lignin [1], in addition to small amounts of pectin, ash and proteins [2]. Cellulose is the most abundant polymer of our planet. In nature, it is found mainly as a structural component of plant and algal cell walls, although cellulose is also produced by some animals, such as tunicates, and several bacteria [3]. It does not accumulate in the environment, due to the existence of cellulytic fungi and bacteria degrading some of the components of plant cell walls; a process that, while extremely slow, plays a major role in carbon recycling within the biosphere [4]. Obviously, the existence of formidable reserves of lignocellulose that are continuously renewed represents an excellent opportunity for the development of biofuels, as a sustainable alternative to the fossil fuels currently in use [5].

Exploitation of cellulosic biomass is, however, currently limited by the absence of low-cost technologies for its processing. A promising strategy for this purpose is the use of cell wall remodeling and degrading enzymes from ligninolytic organisms [3].

Structure of the cell wall and composition of lignocellulose

The cell wall is a highly ordered structure formed, mainly, by cellulose, hemicelluloses and lignin, a phenolic polymer (Figure 1) [6]. Exact identity and relative abundances of each of these polymers vary even within the same plant, depending on age, tissue and growth stage [1]. Cell walls are structured so as to enable them to play a wide array of disparate, sometimes opposing roles. They provide resistance to mechani-

Corresponding author

Palabras clave: lignocelulosa, celulasas, expansinas, hongos, biocombustibles

cal stress, shape the cell and protect it against many pathogens; at the same time, they must be reasonably flexible to withstand shear forces, and permeable enough to allow the passage of signaling molecules into the cell [7].

**Cellulose**

Cellulose, the most abundant natural polymer, is highly stable and insoluble in water. It constitutes the principal component of plant cell walls, accounting for 50% of the dry weight of wood. Its degree of polymerization varies according to its origin, and may range from 2000 to 25,000 monomers [7-8]. Cellulose is formed by D-glucose monomers condensed together through β-1,4 glycosidic bonds, forming cellobiose molecules (β-1,4-linked glucose dimer) that are, in turn, linked together into straight, non-branched chains [9]. In nature cellulose is seldom found as single, isolated chains; forming filaments instead from the very moment it is synthesized. These filaments, denominated microfibrils, may contain from 36 to over 1200 cellulose chains, and have diameters of 5 to 15 nm [7]. Cellulose chains are held together in the microfibril through hydrogen bonds and Van der Waals forces, forming a crystalline, organized structure that is refractory to hydrolysis in certain areas of the microfibril (Figure 2A) [9]. Highly ordered, crystalline regions are interspersed with regions containing disorganized or amorphous cellulose, which constitute 5 to 20% of the microfibril. Amorphous regions are more susceptible to enzymatic degradation [10-11]. Cellulose is a highly resistant substrate that is, in turn, tightly associated with hemicelluloses and lignin, forming a structure that is very resistant to degradation. Degrading lignin, therefore, is a feat accomplished only by a few cellulolytic organisms [8].

**Hemicellulose**

Hemicellulose constitutes from 25 to 30% of wood by dry weight. It is a complex heteropolysaccharide composed mainly of pentoses (D-xylose and L-arabinose) and hexoses (D-glucose, D-mannose and D-galactose), usually acetylated and forming branched chains, in addition to 4-O-methylglucuronic, D-galacturonic and D-glucuronic acids, condensed through β-1,4 and, occasionally, β-1,3 glycosidic linkages [9]. These short lateral branches, formed by different sugars, make hemicellulose less refractory to a number of treatments (Figure 2B) [9, 12]. The components of hemicellulose are also classified as xylans, xyloglucans, mannans, glucomannans and glucans, bonded together through β-1,3 or β-1,4 linkages. Xylans are the most abundant component of hemicellulose (over 70% by weight). They are formed by D-xylose units condensed through β-1,4 linkages, and may carry different substitutions, originating arabinoxylans (if substituted with arabinose), such as those found in grasses, or glucuronoxylans and glucuronoarabinoxylans (if substituted with glucose or glucose-arabinose, respectively), which represent the main constituents of the secondary wall of dicotyledonous plants. In addition to xylose, xylans may contain arabinose, glucuronic acid or 4-O-methyl ether-glucuronic acid, and acetic, ferulic or p-coumaric acids. Exact composition and branching frequency depend on the origin.
of hemicellulose [13]. In hard woods from deciduous trees such as poplar, birch and elm, hemicellulose is formed mainly by xylans where 60 to 70% of residues are acetylated, whereas the soft woods of conifers such as pine and cedar sport hemicelluloses composed mainly of glucomannans [14].

The mannans and galactomannans of hemicellulose have a core structure of β-1,4-linked mannose residues, which is randomly branched with mannose and glucose residues in glucomannans. There are structural differences between hemicelluloses from different species and even different cell types in the same individuals [14-15]. The most important role of hemicellulose is to bond together lignin and cellulose fibers, thus providing rigidity to the cellulose-hemicellulose-lignin mesh. Lignin and hemicellulose are linked together mainly by ester bonds between arabinose residues in hemicellulose and hydroxyl groups in lignin residues, whereas cellulose binds to hemicellulose through hydrogen bonds [16].

**Lignin**

Lignin is one of the most abundant polymers in nature after cellulose and hemicellulose. It is highly resistant to chemical or biological degradation, providing structural support to the cell wall, decreasing its permeability and conferring resistance to the attack of microorganisms. Together, lignin and hemicellulose form an amorphous matrix imbibing cellulose fibers to protect them from degradation [17].

Structurally, lignin is a water-insoluble, irregular, branched heteropolymer formed through the polymerization of three phenylpropane-type aromatic alcohols (coumaryl, coniferyl and sinapyl alcohols) through C-C bonds and esters involving the aromatic rings. This polymer, with constitutes 20 to 30% of wood by weight, protects and confers rigidity to the structural polysaccharides (cellulose and hemicellulose) (Figure 2C) [4, 18-19]. The main constituent of lignin in soft woods is coniferyl alcohol; in hard woods, this place is occupied by coumaric and sinapinic acids instead [12]. Lignin is the component of lignocellulosic material exhibiting the highest resistance to degradation, which limits its application and that of the polysaccharides it protects. The number of microorganisms able to mineralize this substance is really small [18].

**Lignocellulose-degrading microorganisms**

There are natural organisms that can degrade lignocellulose. Some of these are aerobic cellulolytic bacteria of the *Actinomycetales* order (Phylum *Actinobacteria*) living in soils, water, humus, agricultural waste such as sugarcane bagasse and decaying leaves [20]. Enzyme systems composed of cellulases and xylanases capable of degrading cell wall components have been described in aerobic bacteria such as *Pseudomonas fluorescens subsp. cellulosa*, *Streptomyces lividans* and *Cellulomonas fimi* [21-23]. Anaerobic bacteria of the *Clostridiales* order (Phylum *Firmicutes*) generally found in soils, decaying plant waste, the rumen of ruminant animals, termite guts, compost, waste water and wood processing plants, also contain cellulolytic enzyme complexes denominated celluloloses [20]. Some cellulolytic anaerobic bacteria, such as *Butyrivibrio fibrisolvens*, *Fibrobacter succinogenes*, *Ruminococcus flavefaciens*, *Clostridium cellulovoran* and *C. cellulolyticum* are also endowed with cellulases and xylanases [24-26].

Extreme environments also host a number of cellulolytic microorganisms, such as the Antarctic bacterium *Pseudoalteromonas haloplanktis* [27].

Among fungi, the most efficient at using wood as substrate are the basidiomycetes, considered the principal taxonomic group involved in the degradation of wood with all its components [3, 17, 28]. However, the ability to utilize lignocellulosic material is widely distributed among fungi, from chytridiomycetes to basidiomycetes. The chytridiomycetes include a number of anaerobic species living in the gastrointestinal tract of ruminants [29]. *Anaeromyces*, *Caecomyces*, *Neocallimastix*, *Orpinomyces* and *Piroromyces* constitute the five most studied genera of anaerobic fungi [30]. Unlike aerobic fungi, their anaerobic counterparts are often endowed with large multienzyme complexes of cellulases and hemicellulases similar to those of bacterial cellulosomes [31-32] (Table 1). Examining the taxonomic composition of cellulolytic fungi inhabiting the decaying leaves and rotting woods of forest soils, zygomycetes are represented by a single genus, *Mucor*, while ascomycetes and basidiomycetes are represented by genera such as *Chaeottomium*, *Trichoderma*, *Aspergillus*, *Fusarium*, *Coriolus*, *Phanerochaete*, *Schizophyllum*, *Volvariella*, *Pseudoruporpus* and *Bjerkandera* [3, 33-35]. Two of the most studied fungi, due to their industrial relevance, are *Trichoderma reesei* and *Phanerochaete chrysosporium* [20].

### Table 1. Some aerobic and anaerobic cellulolytic fungi, enzymes they produce and substrates they degrade. Compiled from Dashtian M, Scharf H, Qin W [28] and Quiroz-Castañeda RE, et al. [33]

<table>
<thead>
<tr>
<th>Group</th>
<th>Fungus</th>
<th>Enzymes produced</th>
<th>Substrate</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Ascomycetes</strong></td>
<td><strong>Trichoderma reesei</strong></td>
<td>Cellulases (CMCase, CBH, BGL), hemicellulases (xylanases)</td>
<td>Wheat straw</td>
</tr>
<tr>
<td><strong>Basidiomycetes</strong></td>
<td><strong>Trichoderma harzianum</strong></td>
<td>Cellulases (CMCase, CBH)</td>
<td>Wheat straw</td>
</tr>
<tr>
<td><strong>Basidiomycetes</strong></td>
<td><strong>Aspergillus niger</strong></td>
<td>Cellulases, xylanases</td>
<td>Sugarcane bagasse</td>
</tr>
<tr>
<td><strong>Basidiomycetes</strong></td>
<td><strong>Phanerochaete chrysosporium</strong></td>
<td>Cellulases (CMCase, CBH, BGL, peroxidases, hemicellulases (xylanases)</td>
<td>Cedar sawdust, grape seeds, rye husk</td>
</tr>
<tr>
<td><strong>Fomitopsis palustris</strong></td>
<td></td>
<td>Cellulases (CMCase, CBH, BGL)</td>
<td>Microcrystalline cellulose</td>
</tr>
<tr>
<td><strong>Pycnoporus sanguineus</strong></td>
<td>Cellulases (CMCase, CBH, BGL, hemicellulases (xylanases)</td>
<td>Wheat straw, cedar sawdust, rice husk, corn stubbles, <em>Jatropha</em> husk</td>
<td></td>
</tr>
<tr>
<td><strong>Bjerkandera adusta</strong></td>
<td>Cellulases (CMCase, CBH, BGL, hemicellulases (xylanases)</td>
<td>Wheat straw, cedar sawdust, wheat straw, corn stubble, <em>Jatropha</em> husk</td>
<td></td>
</tr>
<tr>
<td><strong>Anaeromyces</strong></td>
<td><strong>Anaeromyces mucronatus</strong></td>
<td>Cellulases (CMCase), hemicellulases</td>
<td>Hay</td>
</tr>
<tr>
<td><strong>Caecomyces</strong></td>
<td><strong>Caecomyces comatus</strong></td>
<td>Cellulases and hemicellulases</td>
<td>Corn cobs</td>
</tr>
<tr>
<td><strong>Neocallimastix</strong></td>
<td><strong>Neocallimastix frontalis</strong></td>
<td>Cellulases and hemicellulases</td>
<td>Cotton fiber, wheat straw</td>
</tr>
<tr>
<td><strong>Piroromyces</strong></td>
<td><strong>Piroromyces sp.</strong></td>
<td>Cellulases (CMCase, CBH, BGL, hemicellulases (xylanases and mannanases)</td>
<td>Corn cobs</td>
</tr>
<tr>
<td><strong>Orpinomyces</strong></td>
<td><strong>Orpinomyces sp.</strong></td>
<td>Cellulases (CMCase, CBH, BGL, hemicellulases (xylanases and mannanases)</td>
<td>Straw</td>
</tr>
</tbody>
</table>

Enzymes involved in the hydrolysis of lignocellulose

Cellulases are O-glucoside hydrolases (GH) hydrolyzing the β-1,4 linkages of cellulose. They are predominantly found among prokaryotes and fungi [8]. More than a dozen fungal species producing cellulases have been described, including Trichoderma viride, T. harzianum, T. atroviride, T. reesei, Fusarium solani, Aspergillus niger, A. terreus, and P. chrysosporium, well known for their cellolytic abilities [18]. Cellulase genes have also been identified in the marine yeast Aureobasidium pullulans [36]. Curiously, a cellulase gene family (GH45), perhaps acquired through horizontal gene transfer, has been found in Bursaphelenchus xylophilus, a nematode infecting pine wood, [37]. GH are classified into cellulase families on the basis of amino acid sequence similarity [38]. Out of the currently existing 122 families, 14 correspond to cellulases. Most cellulases, together with other glucoside hydrolases, have a structure comprised of a catalytic module, a highly O-glycosylated linker, and a cellulose-binding module (CBM) [3]. This last domain facilitates cellulose hydrolysis by holding the catalytic module in close proximity to its substrate [39]. Cellulases are classified, depending on their enzymatic activity, into three major groups: exoglucanases, endoglucanases and β-glucosidases; some of which have been crystallized, enabling the determination of their tri-dimensional structure [40].

Exoglucanases or celllobiohydrolases (CBH) (EC 3.2.1.74; 1,4-β-D-glucan-glucanohydrolase and EC 3.2.1.91; 1,4-β-D-glucan-celllobiohydrolase) catalyze the successive hydrolysis of residues from the reducing and non-reducing ends of the cellulose polysaccharide, releasing celllobiose molecules as main product of the reaction [4]. These enzymes account for 40 to 70% of the total component of the cellulase system, and are able to hydrolyze crystalline cellulose (Figure 3A). They are monomeric proteins with a molecular weight ranging from 50 to 65 kDa, although there are smaller variants (41.5 kDa) in some fungi, such as Sclerotium rolfsii [41]. Exoglucanases have low levels of glycosylation (from 12% to none at all), their optimum pH is 4 to 5, and their optimum temperature varies from 37 to 60 °C, depending on the specific enzyme-substrate combination [42-43]. Exoglucanases form part of the cellulytic arsenal of the fungi causing white and soft rot and the plant pathogen S. rolfsii, but are found only in some of the basidiomycetes causing the brown rot, such as Fomitopsis palustris [44].

Endoglucanases (EG) (EC 3.2.1.4; 1,4-β-D-glucan-4-glucanohydrolase) randomly cleave internal linkages in amorphous cellulose filaments, generating randomly sized oligosaccharides and creating new chain ends that can in turn be attacked by exoglucanases [4]. Available evidence indicates that these are the enzymes that initiate the cellulytic process, randomly cleaving internal linkages at amorphous regions of cellulose [4].

Figure 3. Enzymes participating in the degradation of cell wall components. A) Degradation of cellulose. Three types of cellulytic activities synergize to degrade this material efficiently. Swollenin activity might enhance the efficiency of the process by relaxing the structure of cellulose in a non-hydrolytic manner. B) Degradation of hemicellulose. Endoxylanases and xylosidases are the main enzymes involved in this process. Additional accessory enzymes are required to degrade the side branches of the heteropolysaccharide. C) Degradation of lignin. The process involves laccases and peroxidases catalyzing oxidative reactions and generating highly unstable free radicals that favor further depolymerization. Images taken from Aro et al. [4].
the cellulose fiber and creating new reducing and non-reducing ends that are susceptible to the action of celllobiohydrolases [45]. Endoglucanases are monomeric enzymes with a molecular weight that ranges from 22 to 45 kDa, although some fungi such as S. rolfsii and Gloeophyllum sepiarium have endoglucanases twice this size [46]. In general, endoglucanases are not glycosylated; however, they sometimes may have relatively low amounts of carbohydrate (from 1 to 12%) [42]. Optimum pH is usually 4 to 5; the only known endoglucanase with a neutral pH optimum is that from the basidiomycete Volvariella volvacea, expressed in recombinant yeast [47]. Their optimum temperature ranges from 50 to 70 °C [48]. Endoglucanases have been successfully isolated from the basidiomycetes causing white and brown rot, from the plant pathogen S. rolfsii, from the yeast Rhodotorula glutinis and from the termite symbiont Termotermomyces sp. [42].

Exhaustively hydrolyzing cellulose also requires the action of β-glucosidases (BGL) (EC 3.2.1.21), which hydrolyze cellobiose, releasing two molecules of glucose and thereby provide a carbon source that is easy to metabolize [4]. Fungi causing white and brown rot, mycorrhizal fungi, plant pathogens and yeast all produce these enzymes [42], β-glucosidases have molecular weights ranging from 35 to 640 kDa; they can be monomeric, reaching molecular weights of approximately 100 kDa, or exist as homo-oligomers, as is the case in the yeast Rhodotorula minuta [49]. Most β-glucosidases are glycosylated; in some cases, as that of the 300 kDa BGL from Trametes versicolor, glycosylation may be superior to 90%. Their optimum pH ranges from 3.5 to 5.5, and their optimum temperature ranges from 45 to 75 °C [28]. The activity of cellulase enzyme systems is much higher than the sum of the activity of its individual subunits; a phenomenon known as synergism [3]. Cellulase systems are not just simply a conglomerate of enzymes with components from all three cellulase types, but act coordinately to efficiently hydrolyze cellulose fibers [3].

Hemicellulose is degraded into monosaccharides and acetic acid. Xylans, the main carbohydrate of hemicellulose, require the coordinated action of several hydrolytic enzymes, such as xylanases and accessory proteins, for their degradation [9]. Most hemicelluloses are glycoside hydrolases, although some of them are carbohydrate esterases hydrolyzing the ester bonds linking acetate or ferulic acid with branched sugars.

Xylanases are the main enzymes participating in the degradation of hemicellulose. This group includes the endoxylanases (EC 3.2.1.8; endo-1,4-β-D-xylanase) which act on the main carbohydrate chain, hydrolyzing the linkages between xyran units and releasing oligosaccharides.

β-xylanases (EC 3.2.1.37; xylan 1,4-β-xylanase) release xylose by cleaving the bonds of xylan oligosaccharides [4, 9]. Degrading hemicellulose also requires accessory enzymes such as xylan esterases, ferulic and coumaric esterases, α-arabinofuranosidases and α-4-methyl glucuronosidases, among others, which act in a synergic fashion to hydrolyze hemicellulose efficiently (Figure 3B) [17]. Like cellulases, xylanases also have a modular structure with catalytic and substrate-binding domains, where the first determines specificity and reactivity towards the substrate, while the second facilitates binding of the enzyme to the substrate [9]. Genes coding for endoxylanases and β-xylanosidases have been cloned from different Aspergillus species, as well as from Penicillium, Agaricus bisporus and Magnaporthe grisea [50-51].

Hydrolases also comprise carbohydrate esterases, which catalyze the O- and N-deacetylation of substrates such as xylan, chitin and some peptidoglycans. There is a carbohydrate esterase in Aspergillus sp., denominated feruloyl esterase, which increases the release of sugars from lignocellulose by removing the ferulic acid residues cross-linking hemicellulose fibers, thus destabilizing the structure and making it more susceptible to the action of hydrolytic enzymes [52-53]. It has also been described that the levels of acetyl xylan esterase of the fungus Volvariella volvacea increase when it is grown in the presence of oat xylan, chitin, cellulose, cellulose, lactose or galactose as carbon source [54]. Selig et al. [55] demonstrated that both esterase and xylanase activities are capable of improving the efficacy of a celllobiohydrolase, acting in synergism to degrade lignocellulosic material.

Lignin depolymerization involves extracellular oxidative enzymes that release highly unstable products that later undergo oxidation reactions [56]. White rot-causing fungi are the most efficient organisms regarding lignin degradation. They are endowed with peroxidases and laccases that participate in ligninolyis [9]. Two groups of peroxidases have been characterized: lignin peroxidase (LiP; EC 1.11.1.14) and manganese-dependent peroxidase (MnP; EC 1.11.1.13); both are oxidoreductases catalyzing hydrogen peroxide-dependent oxidative reactions that involve phenolic and non-phenolic compounds and are essential for lignin degradation [9].

Lignin peroxidase is a glycoprotein with a heme group on its active center. It is the most effective peroxidase, and can oxidize phenolic and non-phenolic lignin compounds, amines, aromatic ethers and aromatic polycyclic compounds [17].

Manganese-dependent peroxidase is also a glycoprotein. It uses manganese as substrate, oxidizing it from Mn²⁺ to Mn³⁺. The latter is in turn a strong oxidant that reacts with phenolic lignin compounds [57]. Biochemical and molecular studies have found a third type of peroxidase, originally described for the first time in the basidiomycete Pleurotus eryngii [58]. This enzyme has received the name of versatile peroxidase (VP), as it combines the activities of manganese-dependent peroxidases and lignin peroxidases. A PV was found in the fungus Bjerkandera adusta strain UAMH 8258, where its synthesis is stimulated by exogenously administered organic acids such as glycolate, glyxololate and oxalate [59]. This enzyme efficiently oxidizes Mn²⁺ to Mn³⁺ and can oxidize a number of substrates in the absence of Mn, such as 2,6-dimethoxyphenol, guaiacol, ABTS, 3- hydroxyanthrancinonic acid, α-amisin and p-anisidine. Its affinity for veratryl alcohol is comparable to that of LiP from other fungi [60].

Laccases (EC 1.10.3.1; p-diphenol dioxygen oxidoreductase) are polyphenol oxidases bearing four copper ions in their active center that catalyze the oxidation of phenolic compounds to quinones [61]. The presence of manganese peroxidase. Anttonie 1995;305(Pt 3):1005-10.

Murty MV, Chandra TS. Purification and properties of an extra cellular xylanase enzyme of Clostridium strain SAV. Antonie van Leeuwenhoek. 1992;61(1):35-41.


oxidation of many phenolic and non-phenolic compounds in the presence of mediators, coupling the reduction of molecular oxygen to water [4]. These enzymes oxidize lignin and generate highly unstable aromatic radicals that favor its depolymerization through breakage of C4 ether bonds, breakage of aromatic rings and demethoxylation [17] (Figure 3C). They have been found in basidiomycetes such as P. chrysosporium, Pleurotus ostreatus, T. versicolor and Pycnoporus sanguineus [61].

### Expansins: plant cell wall-remodeling proteins

In addition to lignocellulose-degrading enzymes, there are also enzymes involved in remodeling the cell wall, denominated expansins, which facilitate its later degradation. Expansins increase the extensibility and relax the tension of plant cell walls. They were first identified through studies of the mechanism involved in pH-dependent extension of plant cell walls (‘acid growth’) [62]. Plant cell wall pH is usually determined by the activity of an H⁺ ATPase localized to the plasma membrane, which pumps protons onto the cell wall; it ranges from 5.5 to 4.5. Low extra-cellular pH (lower than 5.5) causes the cell wall to relax, due mainly to the action of expansins, whose optimum pH is acid [63]. McQueen-Mason et al. [62] isolated two proteins from the cell walls of cucumber hypocotyls that induced the extension of previously heat-inactivated cell walls. Cloning and sequencing of expansin genes [64] revealed, after sequence homology searches in genomic and expressed sequence tags (EST) databases, that expansins are coded by large multigene families present from bryophytes to angiosperms [65]. Expansins are also present in monocotyledonous plants (Arabidopsis), ferns and mosses [66].

After the discovery of expansins, Cosgrove et al. [67] found that group I allergens from grass pollen have regions of significant similarity with the aminoc acid sequence of expansins, and demonstrated that maize pollen extracts exhibited expansin activity when applied in vitro to plant cell walls; a finding that was later corroborated for group I allergens from pollen of other grasses. These proteins with expansin activity secreted by pollen have been proposed to participate in softening of the stigma and tissues of the style to facilitate the penetration of pollen through the pollen tube [68]. Expansins have no hydrolytic activity (glucosidase) and therefore, the bonds they break, if any, are most likely non-covalent [69]. In fact, expansins have been suggested to work by breaking hydroxyl bonds between cellulose fibers or between cellulose and other polysaccharides (xyloglucans), using a non-enzymatic mechanism [70-71] (Figure 4B). Expansins have molecular weights ranging from 25 to 28 kDa and, like cellulases, have a two-domain modular structure in which a C-terminal end, is homologous to group II pollen allergens from grasses. Some authors have speculated that this might be a polysaccharide-binding domain, due to the presence of aromatic and polar aminoacids on the protein surface, where two tryptophans and one tyrosine would form a planar platform of aromatic residues favoring such binding [63, 73].

Domain II of group II pollens folds as a β-sandwich formed by two sheets of four antiparallel β strands each (Figure 4A). In fact, a β-sandwich formed by 3 to 6 β strands per sheet is the most common fold in carbohydrate-binding modules of proteins binding substrates such as crystalline cellulose or chitin [74]. Whitney et al. [75] incubated a cellulose chain bound to xylolygucans of bacterial origin with a cucumber expansin and detected a rapid relaxation of the structure of said compound. This result, obtained through the use of “artificial cell walls”, suggests that expansins modulate the binding between cellulose fibers and xylolygucans, relaxing or breaking the bonds keeping them together. Recently, Wei et al. [76] reported that a cucumber α-expansin synergizes with a pectin lyase by breaking the hydrogen bonds between pectin and xylolygucans.

A total of four families are included in the expansin superfamily: α expansins (EXPA), β expansins (EXPB), α-expansin like-proteins (EXLA) and β-expansin like-proteins (EXLB) [72]. The EXPA members of glucosidase hydrolase family 45 (GH45), which includes mainly β-1,4-endoglucanases of fungal origin. The DPBB domain of members of this family adopts a six-stranded beta barrel structure forming a substrate-binding groove [72]. There are a number of cysteines residues in this domain that are conserved among members of the GH45 family, and are involved in disulfide bonding in the case of fungal expansins. Despite the presence of the GH45 catalytic domain in expansins, no hydrolytic activity has been detected for the latter [72]. Domain II, at the C-terminal end, is homologous to group II pollen allergens from grasses. Some authors have speculated that this might be a polysaccharide-binding domain, due to the presence of aromatic and polar aminoacids on the protein surface, where two tryptophans and one tyrosine would form a planar platform of aromatic residues favoring such binding [63, 73]. Domain II folds as a β-sandwich formed by two sheets of four anti-parallel β strands each (Figure 4A). In fact, a β-sandwich formed by 3 to 6 β strands per sheet is the most common fold in carbohydrate-binding modules of proteins binding substrates such as crystalline cellulose or chitin [74]. Whitney et al. [75] incubated a cellulose chain bound to xylolygucans of bacterial origin with a cucumber expansin and detected a rapid relaxation of the structure of said compound. This result, obtained through the use of “artificial cell walls”, suggests that expansins modulate the binding between cellulose fibers and xylolygucans, relaxing or breaking the bonds keeping them together. Recently, Wei et al. [76] reported that a cucumber α-expansin synergizes with a pectin lyase by breaking the hydrogen bonds between pectin and xylolygucans.

A total of four families are included in the expansin superfamily: α expansins (EXPA), β expansins (EXPB), α-expansin like-proteins (EXLA) and β-expansin like-proteins (EXLB) [72]. The EXPA members of glucosidase hydrolase family 45 (GH45), which includes mainly β-1,4-endoglucanases of fungal origin. The DPBB domain of members of this family adopts a six-stranded beta barrel structure forming a substrate-binding groove [72]. There are a number of cysteines residues in this domain that are conserved among members of the GH45 family, and are involved in disulfide bonding in the case of fungal expansins. Despite the presence of the GH45 catalytic domain in expansins, no hydrolytic activity has been detected for the latter [72]. Domain II, at the C-terminal end, is homologous to group II pollen allergens from grasses. Some authors have speculated that this might be a polysaccharide-binding domain, due to the presence of aromatic and polar aminoacids on the protein surface, where two tryptophans and one tyrosine would form a planar platform of aromatic residues favoring such binding [63, 73]. Domain II folds as a β-sandwich formed by two sheets of four anti-parallel β strands each (Figure 4A). In fact, a β-sandwich formed by 3 to 6 β strands per sheet is the most common fold in carbohydrate-binding modules of proteins binding substrates such as crystalline cellulose or chitin [74]. Whitney et al. [75] incubated a cellulose chain bound to xylolygucans of bacterial origin with a cucumber expansin and detected a rapid relaxation of the structure of said compound. This result, obtained through the use of “artificial cell walls”, suggests that expansins modulate the binding between cellulose fibers and xylolygucans, relaxing or breaking the bonds keeping them together. Recently, Wei et al. [76] reported that a cucumber α-expansin synergizes with a pectin lyase by breaking the hydrogen bonds between pectin and xylolygucans.

A total of four families are included in the expansin superfamily: α expansins (EXPA), β expansins (EXPB), α-expansin like-proteins (EXLA) and β-expansin like-proteins (EXLB) [72]. The EXPA
family includes proteins participating in the relaxation and extension of plant cell walls through a pH-dependent mechanism. These proteins would participate in developmental processes such as organogenesis [76], the degradation of cell walls during the ripening of fruits [77-79] and other processes where extending the cell wall is crucial [66, 80-81]. The EXPB family includes group I pollen allergens from grasses. These proteins are secreted by pollen and have been suggested to soften the tissues of the stigma and style to facilitate the penetration of the pollen tube [67]. EXPB proteins, unlike EXPA members, relax specifically the cell walls of grass cells, probably reflecting differences regarding the organization of cell walls between grasses and dicotyledonous plants [67]. An HFD motif has been found in domain I of EXPA and EXPB family members that is known to form part of the active site of endoglucanases. EXLA and EXLB do not have this sequence motif, which suggests that their mode of action differs to that of the other expansins [72].

The EXLA and EXLB families are comprised of proteins identified by sequence analysis which, despite possessing the two-domain organization typical of expansins, have a number of divergent sequence features that separate them from the EXPA and EXPB families [82]. EXLA family members have a conserved CDCR motif towards the N-terminal end of domain I, and an approximately 17 aminoacid long extension towards the C-terminal end of domain II that is not found among the remaining expansin families [72]. A recent report by Dermasev et al. [83] ascribed an important role to a tomato EXLB protein during early stages of the interaction with the mycorrhizal fungus Glomus intraradices, based on the fact that transiently silencing the transcription of the EXLB protein caused a reduction in spore formation and arbuscular expansion. Another group included in the superfAMILY is the X-like expansin family (EXXL), comprised of proteins exhibiting weak sequence homology with the domains of EXPA and EXP members, and identified in organisms other than plants [82], such as the mucilaginous fungus Dictyostelium [84] and the bacteria Bacillus subtilis, Clavibacter michiganensis and Haella chejuensis [85-87].

The denomination of expansin or expansin-like is reserved for proteins exhibiting both domain I and domain II. Proteins with only one of these domains are not classified as expansins [82].

**Biological role of expansins**

The role of expansins has been studied using a diverse array of experimental approaches, ranging from immunohistochemistry, gene expression analysis and ectopic expression of expansin genes to gene silencing with antisense technology and transgenic plants.

**Expansin immunohistochemistry**

Immunohistochemistry has been used to locate expansins to meristems and growth zones of plant roots and stems, as well as to forming leaf primordia in apical meristems and epidermal cell walls of forming roots [76, 88]. Expansins have been found to be distributed evenly along the cell wall, and are not restricted to specific sites or the cell wall-cytosplasmic membrane interface. In some occasions, Golgi vesicles are also labeled with expansin antibodies, indicating that expansins travel to the cell wall via the secretory route [72].

Immunofluorescence studies in fine roots from maize have shown expansins to accumulate at the cytoplasm and cell wall of emerging primordia [89]. Expansins of the EXPB family are quite abundant in pollen from prairie plants, and have been found both on pollen surface and intracellularly [90].

**Gene expression analysis**

Expansin gene expression studies by Northern blot and in situ hybridization have shown these genes to be subject to differential regulation depending on organ, tissue and cell type under study, responding differently to plant hormone treatments, light and pollination. These studies have revealed that expansins are involved in a number of events, going from germination, fruit ripening and pollination to growth responses under flood conditions. For instance, Reinhart et al. [88] found that gene LeExp18, coding for a tomato α-expansin, was expressed during the formation of visible leaf primordia in the apical meristems of tomato plants [88]. Another α-expansin is differentially up-regulated during fruit ripening in tomato and strawberry [78-79]. Similarly, transcripts from an α-expansin accumulate in the endosperm during the germination of tomato seeds, possibly to remodel the cell wall and facilitate the appearance of the radicle [91]. In maize, five genes coding for α and β expansins exhibited differential regulation patterns between seedlings and adult plants for different organs. In rice internodes, gibberellin (GA) induces the expression of five β-expansin genes, whose levels correlate positively with growth rate [92-93].

**Ectopic expression of expansin genes**

Pien et al. [94] induced the local expression of the cucumber expansin gene Cex29 in incipient leaf primordia of the apical meristems of tobacco leaves. The results showed that the expansin induced early growth of leaf primordia with a change in phytoaxis (arrangement of leaves along the shoot) for the apical meristem.

Applying expansins to tomato leaf primordium forced their growth, resulting in deformed leaves, as described by Fleming et al. [95]. The exogenous delivery of expansins to Arabidopsis hypocotyls stimulates their expansion to a degree comparable to that achieved when applying auxin at 1 μM [81].

**Use of antisense sequences and expansin transgenic plants**

Although the multigene nature of expansin families has hindered the task of determining the biological role of individual genes, there are some results worth discussing in this area. Cho and Cosgrove [96], for instance, used antisense sequences for an Arabidopsis α expansin, observing that reductions in its steady state levels were accompanied by significant reductions in growth rate. Also, suppressing the expression of the tomato gene LeEXP1 during ripening results in firmer fruits that can be stored for longer periods [97]. In most cases, silencing expansin genes leads to growth


Rosa E Quiroz-Castañeda and Jorge L Folch-Mallol

Cell wall remodeling proteins
inhibition, whereas excessive ectopic expression leads to abnormal growth.

Expansins have been identified as an important player in developmental processes requiring a decrease of cell wall tension, such as fruit ripening [77, 97], the formation of xylem [98], abscission during the development of parasite plants [99], seed germination [91], the penetration of the pollen tube through the stigma [67, 100], association with mycorrhizal fungi [83], the development of nitrogen-fixing nodules in legumes [101], the development of parasite plants [102], and rehydration of the resurrection plant, Craterostigma plantagineum, which coils when dry and extends when hydrated [103].

Some plants adapted to aquatic environments elongate markedly when flooded, which correlates with the activation of expansin genes [69, 104]. Rice, for instance, when flooded and subjected to hypoxia, increases the expression of an α-expansin, and increases growth of the coleoptile. Similarly, there was a positive correlation between growth of cotton fibers and the expression of α expansins during the early stage of elongation [105-106].

Expansin genes are also induced under drought; for instance, leaf shoots of the temperature-tolerant grass Agrostis scabra induce the AaEXP1 gene, coding for an expansin-like protein, after a 1 hour exposure to heat stress [107]. In the resurrection plant, C. plantagineum, there is a correlation between the expression of leaves during dehydration and increased levels of α-expansin transcripts in the cell walls of leaves, suggesting a role for these proteins in the regulation of leaf growth during dehydration. The involvement of expansins in drought and dehydration-related processes must, however, be investigated further [103].

Taxonomic diversity of expansins

Expansins and expansin-like proteins have been detected in angiosperms such as Arabidopsis thaliana, Oryza sativa, Zea mays and Triticum aestivum [65, 93, 108-109], gymnosperms such as pine and poplar, ferns such as Postia placenta, and the moss Physcomitrella patens. Some members of the expansin superfamily have been found even in a potato-infecting nematode, Globodera rostochiensis, where they are hypothesized to favor the infection process [108, 110-112] (Table 2). Proteins with expansin-like activity have been found even in a potato-infecting nematode, Globodera rostochiensis, where they are hypothesized to favor the infection process [108, 110-112] (Table 2). Some plants adapted to aquatic environments elongate markedly when flooded, which correlates with the activation of expansin genes [69, 104]. Rice, for instance, when flooded and subjected to hypoxia, increases the expression of an α-expansin, and increases growth of the coleoptile. Similarly, there was a positive correlation between growth of cotton fibers and the expression of α expansins during the early stage of elongation [105-106].

Expansin genes are also induced under drought; for instance, leaf shoots of the temperature-tolerant grass Agrostis scabra induce the AaEXP1 gene, coding for an expansin-like protein, after a 1 hour exposure to heat stress [107]. In the resurrection plant, C. plantagineum, there is a correlation between the expression of leaves during dehydration and increased levels of α-expansin transcripts in the cell walls of leaves, suggesting a role for these proteins in the regulation of leaf growth during dehydration. The involvement of expansins in drought and dehydration-related processes must, however, be investigated further [103].

Other proteins with expansin-like activity

The expansin-like proteins described for fungi such as T. reesei, which utilizes plant material, may obviously participate in the degradation of cellulose. In other words, fungal expansins and expansin-like proteins may be involved in plant pathogenesis or in the degrada-

tion of cell walls, aimed at using its components as carbon source [108]. Proteins with expansin-like activity denned swollesins have been identified in ascomycete fungi, such as Trichoderma and Aspergillus (113-115) (Table 2). Saloheimo et al. [114] cloned and expressed in S. cerevisiae a swollesin gene from T. reesei, swollesin coding for a protein that modifies the structure of cellulose in swollen regions of cotton fibers (hence the name) without releasing reducing sugars. Swol is a fungal expansin-like protein, containing a pollen allergen domain and a cellulose-binding domain.

Other fungal swollesins include the one described for Trichoderma asperellum [115]. This protein has been an ascribed an important role in the process of colonization of cucumber roots, since its overexpression increases the colonization success rate. Proteins with expansin activity could be used to improve the efficiency of cellulose bioconversion processes. For example, a swollesin purified from Aspergillus fumigatus has been used in combination with cellulases to facilitate the saccharification of microcrystalline cellulose (Avicel) [113]. Kim et al. [85] also described the synergism of an EXL from B. subtilis in the enzymatic hydrolysis of cellulose. Recently, Quiroz-Castañeda et al. [116] cloned and characterized a new protein with expansin activity from the basidiomycete fungus Bjerkandera adusta, denominated loosenin (LOOS1). This protein only contains a DPBB domain, and is able to relax the structure of cotton, enhancing up to 7.5-fold the rate of release of reducing sugars from agave (Agave tequilana) fiber. Given the optimum pH of LOOS1 (pH 5), it could be applied to processes of saccharification from natural substrates, facilitating the release of reducing sugars together with cellulases and expansins.

Table 2. Organisms containing members of the expansin superfamily and other expansin-like proteins

<table>
<thead>
<tr>
<th>Organism</th>
<th>EXP A</th>
<th>EXP B</th>
<th>EXL A</th>
<th>EXL B</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>Arabidopsis</td>
<td>26</td>
<td>6</td>
<td>3</td>
<td>1</td>
<td>[63]</td>
</tr>
<tr>
<td>Rice</td>
<td>33</td>
<td>18</td>
<td>4</td>
<td>1</td>
<td>[85]</td>
</tr>
<tr>
<td>Poplar</td>
<td>27</td>
<td>3</td>
<td>2</td>
<td></td>
<td>[87]</td>
</tr>
<tr>
<td>Papaya</td>
<td>15</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>[86]</td>
</tr>
<tr>
<td>Physcomitrella patens</td>
<td>27</td>
<td>7</td>
<td>0</td>
<td>0</td>
<td>[94]</td>
</tr>
<tr>
<td>Zea mays</td>
<td>5</td>
<td>8</td>
<td>0</td>
<td>0</td>
<td>[82]</td>
</tr>
<tr>
<td>Triticum aestivum</td>
<td>30</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>[97]</td>
</tr>
<tr>
<td>Pinus taeda</td>
<td>7</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>[85]</td>
</tr>
</tbody>
</table>

Table 2. Organisms containing members of the expansin superfamily and other expansin-like proteins

<table>
<thead>
<tr>
<th>Organism</th>
<th>EXP A</th>
<th>EXP B</th>
<th>EXL A</th>
<th>EXL B</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>Trichoderma reesei</td>
<td>3</td>
<td></td>
<td></td>
<td></td>
<td>[98]</td>
</tr>
<tr>
<td>Trichoderma reesei</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td>[88]</td>
</tr>
<tr>
<td>Bjerkandera adusta</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td>[90]</td>
</tr>
<tr>
<td>Aspergillus nidulans</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td>[99]</td>
</tr>
<tr>
<td>Aspergillus fumigatus</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td>[87]</td>
</tr>
<tr>
<td>Trichoderma asperellum</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td>[89]</td>
</tr>
<tr>
<td>Pastia plaetosa</td>
<td>7</td>
<td></td>
<td></td>
<td></td>
<td>[100]</td>
</tr>
<tr>
<td>Dicystostele discoides</td>
<td>5</td>
<td></td>
<td></td>
<td></td>
<td>[81]</td>
</tr>
<tr>
<td>Bacillus subtilis</td>
<td>1</td>
<td></td>
<td>EXL</td>
<td></td>
<td>[78]</td>
</tr>
<tr>
<td>Hahella chejuensis</td>
<td>1</td>
<td></td>
<td>EXL</td>
<td></td>
<td>[80]</td>
</tr>
<tr>
<td>Rice</td>
<td>80</td>
<td></td>
<td></td>
<td></td>
<td>[57]</td>
</tr>
<tr>
<td>Laccaria bicolor</td>
<td>11</td>
<td></td>
<td></td>
<td></td>
<td>[89]</td>
</tr>
<tr>
<td>Pleurotus ostreatus</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td>[90]</td>
</tr>
<tr>
<td>Neurospora crassa</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td>[91]</td>
</tr>
</tbody>
</table>

*Sequences identified in the corresponding genome whose activity has not been evaluated.

like proteins. For example, it might be used as an additive to obtain fermentable sugars from lignocellulose. The idea of using plant expansins in saccharification processes has, in fact, been patented [117].

**Biotechnological potential of lignocellulose**

Using lignocellulose materials means, in the first place, gaining access to the hemicellulose and cellulose embedded within the lignin matrix. The timber and wood industries, as well as gardening activities and agriculture, generate huge amounts of lignocellulosic leftovers every year [28] that can be potentially put to use in the preparation of animal fodder or the generation of biofuels. In 2008 alone, two countries, the United States and Brazil, accounted for 90% approximately (out of a total of 12.3 billion gallons) of the global production of ethanol. Brazil produces ethanol from sugarcane juice, employing a total of 344 production plants, whereas the United States obtains ethanol from cornstarch, employing 217 production plants [17].

The use of sugarcane juice or cornstarch for producing ethanol has turned out to be highly controversial though, as it competes with the production of food and is relatively expensive, making ethanol less cost-efficient than currently used fossil fuels.

These problems have prompted a search for alternative raw materials to produce biofuels, among which lignocellulosic materials occupy a relevant position. One illustration of the potential of lignocellulosic materials for the production of biofuels is provided by a plant run in Canada by Iogen Corporation, which produces over 3 million liters of ethanol from 30 tons of wheat, oat and rye straw per day [118, 119]. As mentioned above, the degradation of lignocellulosic biomass depends on the accessibility of cellulose and hemicellulose contained therein and the subsequent hydrolysis of these polysaccharides into 5- and 6-carbon sugars, which are then turned into ethanol by fermentation. In-bicon, a Danish company, has built a proof-of-concept plant for producing ethanol from plant biomass with an annual capacity of 1.4 million gallons, and Japan-owned Nippon Oil has built plants for the production of biofuels from cellulose that are slated to produce some 67 billion gallons by 2014 [118]. In Europe, the production of bioethanol from lignocellulosic waste has increased with the participation of Finnish companies such as SEKAB, with expertise in the use of sugarcane bagasse and sawdust for this purpose. Recently, in India, Praj Industries has built two plants processing up to 2 tons per day of corn stover, cobs and bagasse as well as agricultural waste and sawdust, producing ethanol from processes based on pre-treatments that efficiently remove and separate lignin from its accompanying polysaccharides [118].

The current option for obtaining biofuels in Latin America is the use of non-cellulosic compounds. This production has seen its volume increase in the last few years, and biofuels are currently produced from sugarcane and wheat by countries such as Brazil, Colombia, Paraguay and Costa Rica. In Mexico, the production of ethanol as biofuels is limited to two plants with an annual output of 2 million liters, obtained from sugarcane. In Argentina ethanol is produced from molasses and sorghum. In Central American countries like El Salvador, Nicaragua and Guatemala, the production of bioethanol is not significant and the necessary infrastructure is still under construction; however, there is government support for developing this industry [120]. The main challenge posed by the conversion of biomass into ethanol is to increase yields to the point that it becomes competitive, from a cost perspective, with currently used fossil fuels [28]. Yields, however, are limited by the lignin matrix protecting cellulose and hemicellulose from degradation, in addition to the presence of crystalline regions in cellulose hindering its hydrolysis. These barriers are currently dealt with using chemical and physical pretreatments that constitute, therefore, a necessary prerequisite for its degradation. Such pretreatments employ high temperatures and extreme pH; a costly and often inefficient option. In addition, they often generate compounds that inhibit later fermentative stages, such as furans and phenolic derivatives [121].

Expansins constitute a potentially useful tool in the utilization of lignocellulosic biomass with biotechnological purposes, as they would optimize the degradation of cellulosic material by relaxing the tension of cell walls under slightly acid conditions (pH ≤ 5.0), similar to those optimal for the action of most cellulases. The pH normally employed in current saccharification and fermentation processes for the production of biofuels falls precisely within this range (4.8 to 5), and therefore expansins might be used as additives to enhance the efficiency of enzymatic saccharification [122].

**Conclusions**

There are advantages to using lignocellulosic material over sugar (sugarcane or starch) as starting material for the production of bioethanol, including, but not limited to, its availability and low cost. One of its fundamental limitations, however, is the need to degrade lignocellulose; a material that, owing to its composition and structure, is refractory to degradation by most organisms and is therefore seldom used as carbon source in nature. It is possible to release fermentable sugars from plant biomass with the use of cellulolytic and xylanolytic enzymes. Recently, it has been described that proteins with expansin activity from fungi, bacteria and plants can be used to remodel plant cell walls, making them more amenable to enzymatic degradation and increasing considerably the efficiency of their release. Applying these new proteins and increasing the efficiency of these processes would enable the production of larger amounts of sugars from lignocellulosic waste, which would then become a very attractive starting material for obtaining biofuels.


Received in June, 2011. Accepted for publication in September, 2011.