We are IntechOpen, the world’s leading publisher of Open Access books
Built by scientists, for scientists

3,900
Open access books available

116,000
International authors and editors

120M
Downloads

154
Countries delivered to

TOP 1%
Our authors are among the most cited scientists

12.2%
Contributors from top 500 universities

WEB OF SCIENCE™
Selection of our books indexed in the Book Citation Index in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?
Contact book.department@intechopen.com

Numbers displayed above are based on latest data collected. For more information visit www.intechopen.com
Hydrolysis of Biomass Mediated by Cellulases for the Production of Sugars

Rosa Estela Quiroz-Castañeda and Jorge Luis Folch-Mallol

Additional information is available at the end of the chapter

http://dx.doi.org/10.5772/53719

1. Introduction

Cellulose, the most abundant organic molecule on Earth is found mainly as a structural component of plant and algal cell walls, is also produced by some animals, such as tunicates, and several bacteria [1]. Natural cellulose is a crystalline and linear polymer of thousands of D-glucose residues linked by β-1,4-glycosidic bonds, considered the most abundant and renewable biomass resource and a formidable reserve of raw material.

It does not accumulate in the environment due to the existence of cellulytic fungi and bacteria, which slowly degrade some of the components of plant cell walls. Both fungi and bacteria possess enzymes such as laccases, hemicellulases and cellulases, which efficiently degrade lignin, hemicellulose and cellulose, respectively [2-3].

In plant cell walls the cellulose microfibrils are encrusted in lignin and hemicellulose in a complex architecture that, together with the crystallinity of cellulose, makes untreated cellulosic biomass recalcitrant to hydrolysis to fermentable sugars [4]. However, a group of proteins with cellulose disrupting activity (expansins, expansin-like proteins, swollenins and loosenins) have the capacity of relaxing cell wall tension by disrupting the hydrogen bonds binding together cellulose fibrils and cellulose and other polysaccharides through a non-enzymatic process, improving subsequent sugar releasing [5-8].

An efficient degradation of this polysaccharide content into fermentable sugars could improve the production of biofuels. Rising energy consumption, depletion of fossil fuels and increased environmental concerns have shifted the focus of energy generation towards biofuel use [3].
In this chapter, we focus on cellulose degradation by cellulases in order to enhance sugars release from biomass. Cellulose structure, allomorphs and its hydrolysis by cellulytic organisms such as fungi and bacteria, is also reviewed, as well as cellulases structure, CAZY classification, their synergistic activity and the recently cellulases identified by metagenomic analysis, an excellent tool in this search of better cellulytic activity.

Another theme analyzed in this chapter is related to crystalline structure of cellulose, the main impediment to achieve full cellulose hydrolysis, and the role of proteins recently reported with cellulose disrupting activity that have improved saccharification processes. These proteins represent good candidates as an additive to enhance sugar production from plant biomass.

2. Structure and composition of the cell wall

It has been estimated that the net CO$_2$ fixation by land plants per year is approximately 56 X 10$^9$ tons and that the worldwide biomass production by land plants is 170–200 X 10$^9$ tons (Table 1). Of this amount, 70% is estimated to represent plant cell walls (revised in [9]).

Lignocellulose is a renewable organic material and is the major structural component of all cell plants. Lignocellulose plant biomass consists of three major components: cellulose (40–50 %), hemicellulose (20–40 %) and lignin (20–30 %) (Figure 1).

<table>
<thead>
<tr>
<th>Production</th>
<th>Tons</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Assimilated CO$_2$</td>
<td>56 X 10$^9$</td>
<td>[10]</td>
</tr>
<tr>
<td>Cell walls</td>
<td>150-170 X10$^9$</td>
<td>[9]</td>
</tr>
<tr>
<td>Lignocellulose</td>
<td>200 X10$^9$</td>
<td>[12]</td>
</tr>
<tr>
<td>Cellulose</td>
<td>100 X 10$^9$</td>
<td>[13-14]</td>
</tr>
<tr>
<td>Wheat straw</td>
<td>540 X 10$^9$</td>
<td>[15]</td>
</tr>
<tr>
<td>Soybean straw</td>
<td>200 X 10$^9$</td>
<td>[16]</td>
</tr>
<tr>
<td>Sugar cane bagasse</td>
<td>54 X 10$^9$</td>
<td>[17]</td>
</tr>
</tbody>
</table>

Table 1. Worldwide annual production of biomass

Minor components are proteins, lipids, pectin, soluble sugars and minerals (Table 2) [9]. It has a thickness of ~0.1 a 10 μm contrasting with <0.01 μm of cell membrane formed by proteins and phospholipids [18].

Examples of such biomass are angiosperms (hardwoods), gymnosperms (softwoods) and graminaceous plants (grasses such as wheat, giant reed and Miscanthus).
Cell walls should play a wide array of disparate and sometimes opposing roles: the resistance to mechanical stress is necessary as well as the shape of the cell and protection against pathogens; at the same time, besides it must be reasonably flexible to withstand shear forces, and permeable enough to allow the passage of signalling molecules into the cell [19].

3. Cellulose structure

Cellulose is the main component in the plant cell walls, and is made of parallel unbranched D-glucopyranose units linked by β-1,4-glycosidic bonds that form crystalline and highly organized microfibrils through extensive inter and intramolecular hydrogen bonds and Van der Waals forces, amorphous cellulose correspond to regions where this bonds are broken and the ordered arrangement is lost (Figure 2). The cellulose chains aggregated into microfibrils are reported to consist of 24 to 36 chains based on scattering data and information about the cellulose synthase [20-21].

Consecutive glucose molecules along chains in crystalline cellulose are rotated by 180°, meaning that the disaccharide (cellobiose) is the repeating unit [22].
Two different ending groups are found in each cellulose chain edge. At one end of each of the chains, a non-reducing group is present where a closed ring structure is found. A reducing group with both an aliphatic structure and a carbonyl group is found at the other end of the chains. The cellulose chain is thus a polarized molecule and the new glucose residues are added at the non-reducing end allowing chain elongation (Figure 2) [23].

A wide variety of Gram-positive and Gram-negative bacterial species are reported to produce cellulose, including *Clostridium thermocellum*, *Streptomyces spp.*, *Ruminococcus spp.*, *Pseudomonas spp.*, *Cellulomonas spp.*, *Bacillus spp.*, *Serratia*, *Proteus*, *Staphylococcus spp.*, and *Bacillus subtilis* [24].

**Figure 2.** Crystalline and amorphous structure of cellulose. The crystalline structure is conserved by hydrogen bonds and Van der Waals forces, in amorphous structure exists twists and torsions that alter the ordered arrangement. Reducing and non-reducing are shown.

### 4. Cellulose crystallinity

In plants, cellulose is synthesized by CESA proteins (Cellulose Synthase) embedded in plasmatic membrane arranged in hexameric groups called rosette particles [25].

Cellulose crystallites are thought to be imperfect, the traditional two-phase cellulose model describes cellulose chains as containing both crystalline (ordered) and amorphous (less ordered) regions. Crystalline structure of cellulose implies a structural arrangement in which all atoms are fixed in discrete position with respect to one another. An important feature of the crystalline array is that the component molecules of individual microfibrils are packed sufficiently tightly to prevent penetration not only by enzymes, but even by small molecules such as water. While its recalcitrance to enzymatic degradation may pose problems, one big advantage of cellulose is its homogeneity [1, 26-27].

Highly ordered, crystalline regions are interspersed with regions containing disorganized or amorphous cellulose, which constitute 5 to 20% of the microfibril. Many studies have shown that completely disordered or amorphous cellulose is hydrolysed at a much faster rate than partially crystalline cellulose; this fact supports the idea that the initial degree of crystallinity
is important in determining the enzymatic digestibility of a cellulose sample. Crystallinity, is
a measure of the weight fraction of the crystalline regions, is one of the most important
measurable properties of cellulose that influences its enzymatic digestibility [19, 28-30].

A parameter termed the crystallinity index (CI) has been used to describe the relative
amount of crystalline material in cellulose. Generally, in nature, crystallinity indexes range
from 40% to 95%, the rest is amorphous cellulose [31]. The degree of polymerization, (DP) is
the number of monomeric units in a polymer molecule, which in cellulose it ranges from 500
to 15,000 but varies depending the substrate (Table 2).

<table>
<thead>
<tr>
<th>Substrate</th>
<th>Crystallinity index</th>
<th>Degree of polymerization</th>
<th>Ref.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Carboxymethyl cellulose (CMC)&lt;a&gt;</td>
<td>NA</td>
<td>100-2000</td>
<td>[32]</td>
</tr>
<tr>
<td>Celldextrins &lt;a&gt;</td>
<td>NA</td>
<td>2-6</td>
<td>[32]</td>
</tr>
<tr>
<td>Avicelb</td>
<td>0.5-0.6</td>
<td>300</td>
<td>[13]</td>
</tr>
<tr>
<td>BCb</td>
<td>0.76-0.95</td>
<td>2000</td>
<td>[13]</td>
</tr>
<tr>
<td>PASCb</td>
<td>0-0.04</td>
<td>100</td>
<td>[13]</td>
</tr>
<tr>
<td>Cottonb</td>
<td>0.81-0.95</td>
<td>1000-3000</td>
<td>[13]</td>
</tr>
<tr>
<td>Filter paper &lt;b&gt;</td>
<td>0-0.45</td>
<td>750</td>
<td>[13]</td>
</tr>
<tr>
<td>Wood pulp &lt;b&gt;</td>
<td>0.5-0.7</td>
<td>500-1500</td>
<td>[13]</td>
</tr>
<tr>
<td>Fluka Avicel PH-101b</td>
<td>0.56-0.91</td>
<td>200-240*</td>
<td>[26]</td>
</tr>
<tr>
<td>Fluka celluloseb</td>
<td>0.48-0.82</td>
<td>280*</td>
<td>[26]</td>
</tr>
<tr>
<td>Sigma α-celluloseb</td>
<td>0.64</td>
<td>2140-2420*</td>
<td>[33]</td>
</tr>
</tbody>
</table>

*According to manufacturer’s data. <a>, Soluble; <b>, Insoluble.

Table 2. Some physical properties of cellulosic substrates

5. Cellulose allomorphs

The crystalline structure of cellulose has been studied since its discovery in the 19th century,
its structure was first established by Carl von Nageli in 1858, and the result was later veri‐
fied by X-ray crystallography [34-35].

In the past decades, many data on the polymorphism of cellulose were analysed, being the
most reliable data published after 1984, when the results of NMR spectroscopic studies of
cellulose were reported [36].

The repeating unit of the cellulose macromolecule includes six hydroxy groups and three
oxygen atoms. Therefore, the presence of six hydrogen bond donors and nine hydrogen
bond acceptors provides several possibilities for forming hydrogen bonds. Due to different
arrangements of the pyranose rings and the possible conformational changes of the hydroxymethyl groups, cellulose chains can exhibit different crystal packings [37].

Four different crystalline allomorphs of cellulose have been identified by their characteristic X-ray diffraction patterns and solid-state $^{13}$C nuclear magnetic resonance (NMR) spectra: celluloses I, II, III ($\beta$, III$\alpha$) and IV (IV$\alpha$, IV$\beta$). The most important allomorphs are cellulose I and II [22].

Some difference in symmetry and chain geometry have been found in unit cell dimensions of various allomorphs and some parameters have been established: $a$, interchain distance, $b$ unit chain length and $c$, intersheet distance, as well as the angles $A$, $\beta$ and $\gamma$ which are the angles between $b$ and $c$, $a$ and $c$, and $a$ and $b$, respectively, (Table 3) [38-40].

<table>
<thead>
<tr>
<th>Allo morph</th>
<th>Bond lengths (Å)</th>
<th>Angles (º)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$a$</td>
<td>$b$</td>
</tr>
<tr>
<td>I</td>
<td>6.717(7)</td>
<td>5.962(6)</td>
</tr>
<tr>
<td>$I_p$</td>
<td>7.784(8)</td>
<td>8.201(8)</td>
</tr>
<tr>
<td>II</td>
<td>8.10(1)</td>
<td>9.03(1)</td>
</tr>
<tr>
<td>II$\beta$</td>
<td>8.03(1)</td>
<td>9.04(1)</td>
</tr>
<tr>
<td>III$\alpha$</td>
<td>4.450(4)</td>
<td>7.850(8)</td>
</tr>
</tbody>
</table>

$I$, Fresh water algae Glaucocystis nostochinearum; $I_p$, Tunicate Halocynthia roretzi; II, Ramie cellulose (mercerized); $\beta$, Regenerated cellulose (Fortisan); III, Marine algae Cladophora. All crystal structures have been determined at 293ºK, except allomorph II (Fortisan) that was also determined at 100ºK (italics).

Table 3. Unit cell parameters of different cellulose allomorphs obtained by X-ray diffractions.

**Cellulose I** is the most abundant form found in nature, is a mixture of two distinct crystalline forms: cellulose $I_\alpha$, the predominant form isolated from bacteria (Acetobacter xylinum) and fresh water algae (Glaucocystis nostochinearum); and cellulose $I_p$, is the major form in higher plants such as cotton and wood celluloses, ramie and animal celluloses, for example in the edible ascidian Halocynthia roretzi [4]. Cellulose from the marine algae Cladophora sp. and Valonia ventricosa is a mixture of both forms, predominating $I_\beta$ [37]. Currently, cellulose I is receiving increased attention due to its potential use in bioenergy production.

**Cellulose II** has a triclinic one-chain unit cell where parallel cellulose chains stack through van der Waals interactions, with progressive shear parallel to the chain axis. **Cellulose II**$\beta$ has a monoclinic two-chain unit cell, which means parallel cellulose chains stacked with alternating shear (Figure 3) [46].

**Cellulose II** is the most crystalline thermodynamic stable form, it can also be obtained from cellulose I by two distinct routes: mercerization (alkali treatment) and regeneration (solubili-
Cellulose II, like cellulose I_{β}, has the monoclinic unit cell (space group P2₁). The different arrangement of the chains (parallel in cellulose I_{β} and antiparallel in cellulose II) is the most substantial difference between these two polymorphs. The cellulose is a highly rigid macromolecule due to the presence of a three-dimensional hydrogen bond network in addition to the C-O-C bonds between the glucopyranose rings. In the absence of such hydrogen bond networks the chains are much more flexible. These hydrogen bonds are responsible for both the poor solubility of cellulose and the difference in the reactivity of the hydroxy groups in esterification reactions (Figure 4) [37].

**Figure 3.** Differences between the monoclinic and triclinic forms of cellulose I. a) In the monoclinic form, cellobiose units stagger with a shift of a quarter of the c-axis period (0.26 nm), whereas the triclinic form exhibits a diagonal shift of the same amount. The angles shown depend on which crystallographic face is being viewed. A glucose unit is represented by rectangles (cellobiose, a dimer of glucose); image reproduced with publisher’s permission [23]. b) Mode of packing in the unit cell of cellulose I: mono and triclinic unit cell. Notice that the monoclinic angle γ is obtuse. Image reproduced with permission from PNAS Copyright (2012).

Cellulose III_{I} and III_{II} can be formed from cellulose I and II, respectively, by treatment with ammonia; in a reversible reaction. Besides producing the different allomorphs of cellulose, this chemical treatment can also alter other physical properties of cellulose, such as the degree of crystallinity and therefore enhanced cellulase accessibility and chemical reactivity. The degree of conversion of cellulose I to cellulose III depends on the reaction period and the temperature used in the final stage of the treatment [47-48].

In [45] solved the crystal structure of cellulose III by synchrotron X-ray and neutron fiber diffraction analyses, and showed that it has a lower packing density than cellulose I_{α} or I_{β} (Figure 4).

Cellulose IV can be most easily prepared by heating cellulose III, and therefore, two polymorphs of it also exist -celluloses IV_{I} and IV_{II} obtained respectively, from celluloses III_{I} and III_{II}. In general, cellulose IV could be prepared by treatment in glycerol at 260 °C after transformation into cellulose II or III. Cellulose I cannot be transformed directly into cellulose IV [46, 49].

Fibrillation makes cellulose IV_{I} less suitable for crystallographic analysis: that is, it makes it more difficult to interpret cellulose IV_{I} as a crystal. For these reasons, it is unclear whether is a crystal with an orthogonal unit cell or a less crystalline form of cellulose I [49]. A thorough review of cellulose crystalline allomorphs can be found elsewhere [46-47].
Although considerable progress has been made in elucidating the crystal structures of cellulose in microfibrils, they are still not well understood, and a deeper understanding of cellulose structure is required [50-51].

Figure 4. Projections of the crystal structures of cellulose I (α and β) II and III down the chain axes directions. C, O, and H atoms are represented as gray, red, and white balls, respectively. Covalent and hydrogen bonds are represented as full and dashed sticks, respectively. H atoms involved in hydrogen bonding are explicitly represented for only cellulose III. Only the major components of hydrogen bonds are represented. Adapted with permission from [45]. Copyright (2012) American Chemical Society.

6. Cellulose-degrading microorganisms

Since cellulose is very difficult to degrade as a component of plant cell walls, only a few microorganisms specialized for plant cell wall degradation can hydrolyse cellulose. Among these, anaerobic and aerobic genera of Domain Bacteria and fungi of Domain Eukarya are included.

Generally speaking, two types of systems occur in regards to plant cell wall degradation by microorganisms. In one type, the organism produces a set of free enzymes that act synergistically to degrade plant cell walls. In the second type, the degradative enzymes are organized into an enzyme complex located in cellular surface called the cellulosome. This complex is very effective in degrading plant cell walls [52].

Anaerobic and aerobic bacteria have different strategies to degrade cellulytic substrates; whereas anaerobic bacteria degrade cellulose using cellulosomes, aerobic bacteria secrete enzymes capable of degrading cellulose that freely diffuse to reach the substrate.

Anaerobic bacteria of the order Clostridiales (Phylum Firmicutes) are generally found in soils, decaying plant waste, the rumen of ruminant animals, compost, waste water, and wood processing plants; these bacteria have also been found in insects like termites (Isoptera), bookworm (Lepidoptera), and so, in a symbiotic relationship in their guts responsible for cellulosic feed digestion. Anaerobic hydrolysis represents 5% to 10% of global cellulose degradation [53-55].
Aerobic bacteria with cellulolytic activities of the order *Actinomycetales* (phylum *Actinobacteria*) have been found on soils, water, humus, agricultural waste (sugar cane) and decaying leaves, these bacteria excretes enzymes capable of degrading cellulose (cellulases) [52]. In aerobic bacteria *Pseudomonas fluorescens* subsp. cellulosa, *Streptomyces lividans* and *Cellulomonas fimi* cellulolytic systems of degradation have been reported [56-58].

Some anaerobic bacteria with cellulolytic activity are *Butyrivibrio fibrisolvens*, *Fibrobacter succinogenes*, *Ruminococcus flavefaciens*, *Clostridium cellulosivorans*, C. cellulolyticum and C. thermocellum [59-61].

Due to the significant diversity in the physiology of cellulolytic bacteria, sometimes is difficult to classify bacteria as mentioned above, therefore, on this basis, they can be placed into three diverse physiological groups: (1) fermentative anaerobes, typically Gram-positive, *(Clostridium and Ruminococcus)*, but with a few Gram-negative species *(Butyribrio and Acetivibrio)* that are phylogenetically related to the *Clostridium* assemblage *(Fibrobacter)*; (2) aerobic Gram-positive bacteria *(Cellulomonas* and *Thermobifida)* and (3) aerobic gliding bacteria, *(Cytophaga* and *Sporocytophaga)* [1, 53].

The ability to utilize lignocellulosic material is widely distributed among fungi, from chytridiomycetes to basidiomycetes. Among fungi, the most efficient at using wood as substrate are the basidiomycetes, considered the principal taxonomic group involved in the aerobic degradation of wood with all its components, they are the main organic material decomposition agents. These aerobic fungi produce extracellular enzymes allowing lignocellulose degradation (laccases, hemicellulases, and cellulases), although some Ascomycetes are able to degrade cellulose compounds as well. Unlike aerobic fungi, some of the Chytridiomycetes anaerobic fungi, have multienzymatic complexes similar to cellulosomes of bacteria [1, 3, 62-63] some members are anaerobic species living in the gastrointestinal tract of ruminants such as *Anaeromyces*, *Caecomyces*, *Neocallimastix*, *Orpinomyces* and *Piromyces*.

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 *Chaetomium*, *Trichoderma*, *Aspergillus*, *Penicillium*, *Fusarium*, *Coriolus*, *Phanerochaete*, *Schizophyllum*, *Volvariella*, *Pycnoporus* and *Bjerkandera*. Two of the most studied fungi, due to their industrial relevance, are *Trichoderma reesei* and *Phanerochaete chrysosporium*.

Nowadays, more than 14,000 fungi, which are active against cellulose and other insoluble fibres, are known [1, 24, 64-66]. A more detailed list of cellulose degrading bacteria and fungi is listed in Table 5.

### 7. Cellulose degradation mediated by cellulosome

Selective pressure of evolution is the force driving microorganisms to adapt a new environment, in anaerobic conditions is necessary a machinery for the extracellular degradation of substrates, such as the recalcitrant crystalline components of the plant cell wall. Due to this, the anaerobes tend to adopt different strategies for degrading plant components, being the cellulosomes the most remarkable feature.
<table>
<thead>
<tr>
<th>Group</th>
<th>Fungi</th>
<th>Enzymes**</th>
<th>Substrate</th>
<th>Ref</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aerobic fungi (extracellular cellulolytic enzymes) Ascomycetes</td>
<td><em>T. reesei</em></td>
<td>Cel, Xyl</td>
<td>Wheat straw</td>
<td>[68-69]</td>
</tr>
<tr>
<td></td>
<td><em>T. harzianum</em></td>
<td>Cel</td>
<td>Cellulose</td>
<td>[70]</td>
</tr>
<tr>
<td></td>
<td><em>A. niger</em></td>
<td>Cel</td>
<td>Sugar cane bagasse</td>
<td>[71]</td>
</tr>
<tr>
<td></td>
<td><em>Schizophyllum commune</em></td>
<td>Cel, Xyl</td>
<td>Microcrystalline cellulose, rice xylan</td>
<td>[72]</td>
</tr>
<tr>
<td></td>
<td><em>P. chrysosporium</em></td>
<td>Cel, Xyl</td>
<td>Red oak, grape seed, barley bran, sorghum</td>
<td>[2, 63, 73]</td>
</tr>
<tr>
<td></td>
<td><em>B. adusta</em></td>
<td>Cel, Xyl</td>
<td>Oak and cedar sawdust, rice husk, corn stubble, wheat straw and <em>Jatropha</em> seed husk</td>
<td>[74]</td>
</tr>
<tr>
<td></td>
<td><em>Fomitopsis palustris</em></td>
<td>Cel</td>
<td>Microcrystalline cellulose</td>
<td>[75]</td>
</tr>
<tr>
<td>Anaerobic rumen fungi (cellulosomes) Chytridiomycetes</td>
<td><em>Anaeromyces mucronatus</em></td>
<td>Cel, Xyl</td>
<td>Orchard grass cellulose</td>
<td>[76]</td>
</tr>
<tr>
<td></td>
<td><em>Caecomyces communis</em></td>
<td>Cel</td>
<td>Microcrystalline cellulose, alfalfa hay</td>
<td>[77]</td>
</tr>
<tr>
<td></td>
<td><em>Neocallimastix frontalis</em></td>
<td>Cel, Xyl</td>
<td>wheat straw</td>
<td>[78]</td>
</tr>
<tr>
<td></td>
<td><em>Orpinomyces sp.</em></td>
<td>Cel</td>
<td>Avicel</td>
<td>[79]</td>
</tr>
<tr>
<td></td>
<td><em>N. patriciarum</em></td>
<td>Cel</td>
<td>CMC</td>
<td>[80]</td>
</tr>
<tr>
<td></td>
<td>Acidothermus cellulolyticus</td>
<td>Cel</td>
<td>Whatman paper No1, Microcrystalline cellulose, Azurine crosslinked hydroxyethylcellulose (AZCL-HEC)</td>
<td>[81]</td>
</tr>
<tr>
<td></td>
<td>Actinospora robiniae</td>
<td>Cel</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Actinosynnema mirum</td>
<td>Cel</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Catenulispora acidiphila</td>
<td>Cel</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Cellulomonas flavigena</td>
<td>Cel</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Thermobispora bispora</td>
<td>Cel</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Xylanimonas cellulosilytica</td>
<td>Cel</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Aerobic bacteria (cellulosomes) Actinobacteria</td>
<td>Clostridium thermocellum*</td>
<td>Cel</td>
<td>Crystalline cellulose</td>
<td>[82]</td>
</tr>
<tr>
<td></td>
<td><em>Thermomonospora fusca</em></td>
<td>Cel, Xyl</td>
<td>wheat straw, oat spelt xylan</td>
<td>[83]</td>
</tr>
<tr>
<td></td>
<td><em>Caldicellulosiruptor kristjanssonii</em></td>
<td>Cel</td>
<td>Microcrystalline cellulose</td>
<td>[84]</td>
</tr>
<tr>
<td></td>
<td><em>Anaerocellum thermophilum</em></td>
<td>Cel, Xyl</td>
<td>Microcrystalline cellulose, xylan</td>
<td>[85-86]</td>
</tr>
</tbody>
</table>

*Termophytic bacteria. **Cel: cellulases; Xyl: xylanases.

Table 4. Fungi and bacteria with cellulolytic activity.
The occurrence of a cellulosome was first observed in the thermophilic bacterium, *C. thermo‐cellum* and has now been described in a number of mesophilic anaerobic bacteria and with some anaerobic fungi particularly *Piromyces* sp.[52, 67].

Cellulosomes are large extracellular enzyme complexes capable of degrading cellulose, hemicelluloses, and pectin; they may be the largest extracellular enzyme complexes found in nature, although the individual cellulosomes size range from 0.65 MDa to 2.5 MDa, some polycellulosomes have been reported to be as large as 100 MDa, [87].

The cellulosome structure is characterized by two components: (a) the non-enzymatic scaffolding proteins with enzyme binding sites called cohesins, (b) enzymes with dockerins protein interacting with cohesins in the scaffolding protein (Figure 5).

Depending of the bacterial species, the scaffolding protein varies in the number of cohesins and cellulose binding modules (CBM) that binds the cellulosome tightly to the substrate and concentrates the enzymes to a particular site of the substrate. Recently, a more complex cellulosome structure with multiple interacting scaffolding proteins that allows the binding of as much as enzymes has been revealed [52].

The cohesin-dockerin interconnect the different scaffoldin components, whereby the specificities among the individual cohesin-dockerin complexes dictate the overall supramolecular architecture of the participating components [88].

In short, the enzymatic cellulosome system may exceed the potential of non-cellulosomal degradative system due to its structural organization, efficient binding to the substrate, the variety of hydrolytic enzymes acting synergistically [52]. Cellulosomes have not been identified in bacteria (or eukarya) that grow above 65 ºC, and have not been identified in the Archaea [89].

![Figure 5. Cellulosome structure. A dockerin is appended to catalytic (enzyme) and noncatalytic carbohydrate-binding modules (CBMs). Dockerins bind the cohesins of a noncatalytic scaffoldin, providing a mechanism for cellulosome assembly. Image reproduced with publisher’s permission [90].](http://dx.doi.org/10.5772/53719)
8. Cellulose degradation mediated by non-cellulosomal enzymes

Aerobic cellulolytic bacteria and fungi use a system for cellulose degradation consisting of sets of soluble cellulases. Cellulases are inducible enzymes by cellulosic substrates, which are synthesized by a large diversity of microorganisms including both fungi and bacteria during their growth on cellulosic materials. These microorganisms can be aerobic, anaerobic, mesophilic or thermophilic. Among them, the genera of *Clostridium*, *Cellulomonas*, *Thermomonospora*, *Trichoderma*, and *Aspergillus* are the most extensively studied cellulases producers [91].

The aerobic cellulase mechanism evolved in terrestrial microorganisms that colonise solid substrates and therefore secrete cellulases to enable degradation of the substrate. Because of the recalcitrance of plant cell walls some cellulolytic microorganisms secrete up to 50% of their total protein during growth on biomass or cellulose [53, 90].

Cellulases are composed of independently folding and structurally and functionally discrete units called domains, making cellulases modular enzymes. Structurally fungal cellulases are simpler as compared to bacterial cellulosomes [32, 88, 92].

Fungal cellulases have two independent domains: a catalytic domain (CD) and a cellulose-binding domain (CBD), which is joined by a short poly linker region to the catalytic domain at the N-terminal. The CBD is comprised of approximately 35 amino acids, and the linker region is rich in serine and threonine [93].

It is clear that the role of the CBD is to bind the enzyme to the cellulose so that the CD keep closer to the substrate and it also gives the CD time to move the chain into its active site before the enzyme diffuses away from the cellulose. It is still not clear whether the CBD also can modify cellulose or otherwise assist cellulose hydrolysis by the catalytic domain [94].

The mixture of free cellulases act synergistically to degrade crystalline cellulose increasing the specific activity up to fifteen fold higher than that of any individual cellulase [95].

9. Cellulolytic organisms from extreme environments

Novel enzymes with application in industry require improved features to tolerate extreme conditions of temperature, pH and salinity. Some microorganisms live in these environments, so called extremophiles and are considered a source of enzymes with potential biotechnological applications.

Extreme environments host a number of cellulolytic microorganisms, such as the Gram-negative Antarctic bacterium *Pseudoalteromonas haloplanktis*, collected from seawater, which secretes a psychrophilic cellulase, Cel5G, this cold adapted enzyme displays a high specific activity at low and moderate temperatures and a rather high thermosensitivity induced by a decrease of the intramolecular interactions [96-97].
Extremely thermophilic cellulose-degrading microorganisms are of particular and biotechnological interest owing to the presence of highly thermostable enzymes. A deeper analysis of these organisms is reported in [89, 98-99].

The group of thermophilic cellulolytic prokaryotes includes two aerobic species, *Rhodothermus marinus* and *Acidothermus cellulolyticus*, and numerous anaerobes of the genera *Caldicellulosiruptor*, *Clostridium*, *Spirochaeta*, *Fervidobacterium* and *Thermotoga* (reviewed by [100]).

All members of the genus *Caldicellulosiruptor* are extremely thermophilic, cellulolytic, and non-spore-forming anaerobes with Gram-positive type cell wall, capable of fermenting different types of carbohydrates and have been isolated mostly from neutral or slightly alkaline geothermal springs in New Zealand, Iceland and California [100].

Recently, thermostable cellulases have also been reported in the thermophilic *Geobacillus* sp. R7 that produces a cellulase with a high hydrolytic potential when grown on pretreated agricultural residues (corn stover and prairie cord grass). In fact, it was demonstrated that *Geobacillus* sp. R7 can ferment the lignocellulosic substrates to ethanol in a single step, improving bioethanol production with important potential for cost reductions. Cellulases genes were also identified in several *Sulfolobales* strains, however, their physiological function is not well understood [101-102].

Another thermophilic bacterium *Anaerocellum thermophilum* degrade lignocellulosic biomass untreated as well as crystalline cellulose and xylan [86].

While cellulases are widespread in Fungi and Bacteria, only one archaeal cellulase, an endoglucanase from *Pyrococcus furiosus*, has been reported. This enzyme exhibits a significant hydrolyzing activity toward crystalline cellulose even tough it lacks a CBD, the role of this intracellular enzyme in Archaea is unclear, given that Archaea are apparently unable to grow on cellulose [103-104].

In the alkali tolerant fungus *Penicillium citrinum* an alkali tolerant and thermostable cellulases were found which may have potential effectiveness as additives to laundry detergents [105].

In this search to improve cellulases activity, hybrids of hyperthermostable glycoside hydrolases have been constructed as reported by [106], for example, using the structural compatibility of two hyperthermostable family 1 glycoside hydrolases, *P. furiosus* CelB and *Sulfolobus solfataricus* LacS a library of hybrids using DNA family shuffling was created.

This study demonstrates that extremely thermostable enzymes with limited homology and different mechanisms of stabilization can be efficiently shuffled to form stable hybrids with improved catalytic features.

Alkaliphilic, thermophilic and halophilic microbial species have the potential to yield valuable new products for biotechnological industry. Alkaliphilic polymer-degrading enzymes such as proteases, lipases and cellulases are most frequently isolated from *Bacillus* or related species. Cellulases and lipases are important not only as components of washing detergents, but they are also applied in the paper and pulp, pharmaceutical, food, leather, chemical or waste treatment industries [107-108].
10. Cellulases structure

Proteins with hydrolytic activity such as cellulases and hemicellulases comprises a complex molecular architecture of discrete modules (a catalytic domain (CD) and one or more CBDs), which are joined by unstructured linker sequences [109].

The catalytic domain spans more than 70% of protein sequence. A sequence analysis of these domains in different cellulases shows a significant variability between them, in fact, active site of the enzyme has distinct three dimensional arrangements: in tunnel shape for a processive exo degradation or in a cleft shape for an endo degradation. This domain is N-glycosylated and is responsible of the cleavage of the glycosidic bond, which occurs through an acid hydrolysis mechanism, using a donor of protons and a nucleophyle or base such as glutamic and aspartic acid [1, 110-111].

The cellulose binding domain facilitates hydrolysis by keeping the catalytic domain nearby the substrate, therefore the presence of CBD is important for cellulases starting and processivity [112]. The CBDs, which is usually O-glycosylated, contain from 30 to about 200 amino acids, and exist as a single, double, or triple domain in a protein. Their location in the protein can be both, C or N terminal and occasionally is centrally positioned.

The CBDs bring the enzyme into a closer and prolonged association with the substrate, increasing the rate of catalysis, this domain was found to function more efficiently in substrate degradation, and removing the CBM from the enzyme or from the scaffolding in cellulosomes dramatically decrease its enzymatic activity (revised in [109]).

In the union of CBD and cellulose, some non polar residues left exposed, mostly tyrosines and tryptophans, showing the flat face of their aromatic ring towards the pyranose ring, this interaction is stabilized by polar residues that form hydrogen bonds [61].

Besides cellulases, CBDs have also been found in other polysaccharides degrading enzymes: hemicellulases, endomannanases and acetilxylanesterases [113].

The linker peptide is a sequence of amino acids connecting the cellulose binding domain and the catalytic domain. This linker contains from 6 to 59 amino acids and functions as a flexible hinge that allows the independent function of each domain [114]. The sequence of the linker varies between enzymes, however, the composition is typically rich in proline, treonine and serine, like in the sequence PTPTPTPTT(PT)₇ of the endoglucanase of C. fimii and NPSGNPPGGNPPGGTTRRPAATTGSPG of the cellobiohydrolase CBHI of T. reesei.

Treonine and serine residues of the peptide linker are highly O-glycosylated to be protected from proteolysis; if the linker is completely absent or is too short then both domains, CBD and CD, obstruct each other and the affinity reduces. Based on the similarities of the linker between cellulases it has been suggested that it could be acting as a flexible hinge facilitating independent function of the domains (Figure 6) [115-116].
11. Mechanisms of cellulose biodegradation

Once the cellulase has recognized a free chain end, it threads the chain into the tunnel to form a catalytically active complex (CAC). Because cellulose decrystallization in water is free-energetically unfavourable, the tunnels or clefts of cellulase CDs contain hydrophobic and polar residues that form favourable contacts with the chain.

Several studies have mutated hydrophobic residues in the CD tunnels of cellulases and chitinases (structurally similar to cellulases), and have demonstrated that hydrophobic residues need to be present in the CD tunnels for digestion of crystalline cellulose to occur [117]. Additionally, in [27] have shown that removal of hydrophobic residues in cellulase and chitinase tunnels can increase processivity rates on more accessible polymers.

Once a cellulase forms a CAC with a cellodextrin chain, the hydrolysis reaction occurs usually via a retaining or inverting mechanism, depending on the directionality of the enzyme. After the reaction occurs, the product must be expelled and another CAC formed by threading another cellobiose unit into the CD (Figure 6) [117].

In most cases, the hydrolysis of the glycosidic bond is catalysed by two amino acid residues of the enzyme: a general acid (proton donor) and a nucleophile/base [111]. Depending on the spatial position of these catalytic residues, hydrolysis occurs via overall retention or overall inversion of the anomeric carbon. Recently, a completely unrelated mechanism has been demonstrated for two families of glycosidases utilizing NAD$^+$ as a cofactor [118-119].

The retaining glycoside hydrolase mechanism leads to a net retention of the configuration at the anomeric carbon (C1) of the substrate after cleavage, since the hydrolysis of a glycosidic bond creates a product with the same configuration at the anomeric carbon as the substrate had before hydrolysis.

The inverting glycoside hydrolase mechanism leads to a net inversion of the configuration at the anomeric carbon (C1) of the substrate after cleavage. This is performed via a single nucleophilic displacement mechanism, where the hydrolysis of a $\alpha$-glycosidic bond creates a product with the $\beta$-configuration, and vice-versa [120].

12. Cellulose biodegradation

Although more than a dozen fungal species considered as cellulose degraders have been reported (including T. viride, T. reesei, F. solani, A. niger, A. terreus, P. chrysosporium, B. adusta and P. sanguineus) [3, 74]; and even with cellulases identified in nematodes (Bursaphelenchus xylophilus, a nematode infecting pine wood), yeast (Aureobasidium pullulans) and marine bacteria (Saccharophagus degradans), the search of new cellulases genes continues. This have led to the construction of metagenomic libraries and bioprospecting analysis from several environments: buffalo rumen, higher termite guts, bovine ruminal protozoan, decomposing pop-
lar wood chips and hardwood forest leading to the identification of new genes and organisms with cellulolytic activities [107, 121-130].

To have a better impression of the latest developments regarding fungal carbohydrate-active enzymes, the following sections will discuss the enzymes needed for cellulose degradation.

(a) Cel7A binding to cellulose, (b) recognition of a reducing end of a cellulose chain, (c) initial threading of the cellulose chain into the catalytic tunnel, (d) threading and formation of a catalytically active complex, (e) hydrolysis in a processive cycle, and (f) product expulsion and threading of another cellobiose (shown in yellow in e and f). Image reproduced with publisher’s permission [131].

Figure 6. Activity on substrate of cellulase (exoglucanase, Cel7A) of T. reesei. The enzyme has a small carbohydrate-binding domain (CBD) of 36-amino acid, a long flexible linker with O-glycan (dark blue), and a large catalytic domain (CD) with N-linked glycan (pink) that can thread a single chain of cellulose into the catalytic tunnel of 50 Å.

13. Cellulases

Multiple types of modular cellulases formed by catalytic and carbohydrate binding domains have been discovered, including at least two exo-β-glucanases, or cellobiohydrolases (CBHs, CBH I and CBH II), four endoglucanases (EG; EG I, EG II, EG III, EG V), and one β-glucosidase (BG) [1].

Cellulases are O-glucoside hydrolases (GH, EC 3.2.1.), a widespread group of enzymes which hydrolyse the β-1,4 linkages or glycosidic bond between two or more carbohydrates or between a carbohydrate and a non-carbohydrate moiety. GH are classified into cellulases families on the basis of amino acid sequence similarity [31, 132].

A classification of glycoside hydrolases in families based on amino acid sequence similarities has been proposed a few years ago. Because there is a direct relationship between sequence and folding similarities, this classification reflects the structural features of these enzymes better than their sole substrate specificity, and helps to reveal the evolutionary relationships between these enzymes, which represent a convenient tool to deduce information of the mechanism [132-133].
Out of the currently existing 125 families, 15 correspond to cellulases (GHF 1,3, 5, 6, 7, 8, 9, 12, 44, 45, 48, 51, 74, 116, and 124), and 64 families group the cellulose binding domains (see http://www.cazy.org/). In [134] an excellent review of and classification system for many CBD families is provided.

The widely accepted mechanism for enzymatic cellulose hydrolysis involves synergistic actions by endoglucanases (EGL, EC 3.2.1.4), exoglucanases or cellobiohydrolases (CBH, EC 3.2.1.74; 1,4-β-D-glucan-glucanhydrolase and EC 3.2.1.91; 1,4-β-D-glucan cellobiohydrolase), and β-glucosidases (BGL, EC 3.2.1.21).

Endoglucanases hydrolyse accessible intramolecular β-1,4-glucosidic bonds of cellulose chains randomly to produce new chain ends; exoglucanases processively cleave cellulose chains at the reducing and non-reducing ends to release soluble cellobiose or glucose; and β-glucosidases hydrolyse cellobiose to glucose in order to eliminate cellobiose inhibition (13). These three hydrolysis processes occur simultaneously as shown in Figure 7.

The activity of cellulase enzyme systems is much higher than the sum of the activity of its individual subunits; a phenomenon known as synergism, so they have to be considered not just simply a conglomerate of enzymes with components from all three cellulase types, but as a mixture that efficiently hydrolyse cellulose fibres.

14. Endoglucanases

These enzymes cleave internal linkages in amorphous cellulose filaments, generating oligosaccharides with different sizes and creating new chain ends that can in turn be attacked by exoglucanases (135). The cellulolytic process is initiated by endoglucanases that randomly cleave internal linkages at amorphous regions of the cellulose fibre and creating new reducing and non-reducing ends that are susceptible to the action of cellobiohydrolases [136].

Endoglucanases are monomeric enzymes with a molecular weight that ranges from 22 to 45 kDa, although some fungi such as Sclerotium rolfsii and Gloeophyllum sepiarium have endoglucanases twice this size [137]. In general, endoglucanases are not glycosylated; however, they sometimes may have relatively low amounts of carbohydrate (from 1 to 12%) [2]. Unlike other endoglucanases reported with optimum pH 4 to 5; the only known endoglucanase with a neutral pH optimum is that from the basidiomycete Volvariella volvacea, expressed in recombinant yeast. Basically, their optimum temperature ranges from 50 to 70 °C [138-139].

Exhaustively hydrolysing cellulose also requires the action of β-glucosidases (BGL) (EC 3.2.1.21), which hydrolyse cellobiose, releasing two molecules of glucose and thereby provide a carbon source that is easy to metabolize. Fungi causing white and brown rot, mycorrhizal fungi and plant pathogens produce these enzymes [2, 135].

According to [13], primary hydrolysis occurs on the surface of solid substrates and releases soluble sugars with a degree of polymerization (DP) up to 6 into the liquid phase upon hydrolysis by endoglucanases and exoglucanases. This depolymerisation step performed by
endoglucanases and exoglucanases is the rate-limiting step for the whole cellulose hydrolysis process. The second hydrolysis involves primarily the hydrolysis of cellobiose to glucose by β-glucosidases, although some β-glucosidases also hydrolyse longer cellodextrins. The combined actions of endoglucanases and exoglucanases modify the cellulose surface characteristics over time, resulting in rapid changes in hydrolysis rates [32].

To assay endoglucanase activity, there are substrates that are used, such as carboxymethylcellulose (CMC), a soluble amorphous cellulose form that is an excellent substrate for endocellulases and its hydrolysis does not require a CBD [110].

15. Exoglucanases

Also known as cellbiohydrolases, these enzymes catalyse the successive hydrolysis of residues from the reducing and non-reducing ends of the cellulose, releasing cellobiose molecules as main product, which are hydrolysed by β-glucosidases. They account for 40 to 70% of the total component of the cellulase system, and are able to hydrolyse crystalline cellulose.

Exoglucanases have shown specificity on the ends of cellulose, such as *T. reesei* cellbiohydrolase (CBH) I and II that act on the reducing and non-reducing cellulose chain ends, respectively [112].

These enzymes 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*. Low levels of glycosylation (around 12% to none at all) are found in these enzymes; and their optimum pH is 4 to 5, with an optimum temperature from 37 to 60 °C, depending on the specific enzyme-substrate combination [137, 140].

Exoglucanases form part of the cellulolytic machinery of the fungi causing white and soft rot and they are found only in some of the basidiomycetes causing the brown rot, such as *Fomitopsis palustris* [141].

Crystalline cellulose (Avicel, bacterial cellulose or filter paper), which is the main form of cellulose in most plant cell walls are good substrates for exoglucanase activity assay, because it has a low DP and relatively low accessibility; however, some endoglucanases can release considerable reducing sugars from Avicel [13].

16. β-glucosidases

β-D-glucosidases hydrolyse soluble cellobiose and other cellodextrins with a DP up to 6 to produce glucose in the aqueous phase in order to eliminate cellobiose inhibition [13].

These enzymes have molecular weights ranging from 35 to 640 kDa, and they can be monomeric or exist as homo-oligomers, as is the case β-glucosidase of the yeast *Rhodotorula minuta* [142]. 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 (3). β-D-glucosidase activities can be measured using cellobiose, which is not hydrolysed by endoglucanases and exoglucanases [13].

17. Synergy between cellulases

Synergistic cooperation between cellulases is a prerequisite for efficient degradation of cellulose, but its molecular mechanisms are not fully understood. Synergistic action has been observed between two different cellubiohydrolases and between endoglucanases. However, more synergistic mechanisms have been proposed [143-144]:

Synergy endo-exo, occurs between endo and exoglucanases, where the action of endoglucanases provide free ends of the cellulose chain to the exoglucanases.

Synergy exo-exo, exoglucanases progressively act on reducing and non-reducing ends of the cellulose chain.

Synergy between exoglucanases and β-glucosidases, the latter process cellobiose produced as final product of the action of the exoglucanases.

Intramolecular synergy between catalytic domain and cellulose binding domain of cellulases.

As a whole system, plant cell wall polysaccharides should be degraded efficiently not only by synergy between cellulases but with participation of the other degrading enzymes as xylanases.

In (145) a synergistic mechanism between cellulases and xylanases in order to saccharify wheat straw for bioethanol production is reported. More recently, a new type of synergism between enzymes that employ oxidative reactions to break glycosidic bonds and hydrolytic enzymes was reported in chitin degradation [28].

Although a significant amount of information has been generated related to the action of cellulases and their mechanisms to degrading cellulose, the biodegradation of crystalline cellulose...
lose is still a slow process because the substrate is insoluble and poorly accessible to enzymes.

To overcome this situation scientists have optimized ratio of cellulolytic enzymes, and it was found that the best saccharification of crystalline cellulose is achieved with the enzyme blend: 60:20:20 (CBHI:CBHII:EGI) wherein a saturated level of BG was included to eliminate cellobiose inhibition [146]. In a different report, the impact of the cellulase mixture composition on cellulose conversion was modelled, and the findings suggested different optimum ratios for substrates with different characteristics, specifically degrees of polymerization and surface area [147].

Also, researchers have pointed out the use of proteins that relax plant cell wall structure as a complementary activity before action of cellulases in order to improve saccharification.

18. Plant cell remodelling proteins

In addition to lignocellulose-degrading enzymes, there are also enzymes involved in remodelling the cell wall, which could facilitate its later degradation.

18.1. Expansins

Expansins are pH-dependent wall-loosening proteins required for cell enlargement and expansion in many developmental processes. Although to date their precise mechanism of action remains unclear, evidence point toward a role in dissociating the cell wall polysaccharide complex that links together wall components, thus promoting slippage between wall polymers and, eventually, expansion in cell wall [148-149].

These proteins are coded by large multigene families present from bryophytes to angiosperms and also present in monocotyledonous plants (rice, maize), dicotyledonous plants (Arabidopsis), ferns and mosses.

Expansins have no hydrolytic activity (glucosidase) and therefore, it has been suggested to work by breaking hydrogen bonds between cellulose fibres or between cellulose and other polysaccharides (xyloglucans), using a non-enzymatic mechanism (Figure 8) [150-153].

Expansins have molecular weights ranging from 25 to 28 kDa and, like cellulases, have a two-domain modular structure and an approximately 20 amino acids-long amino-terminal signal peptide [149].

Domain 1 occupies the N-terminal part of the protein, and it has a DPBB (Double Psi Beta Barrel) structure. It is homologous to the catalytic domain of members of glycoside 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. Despite the presence of the GH45 catalytic domain in expansins, no hydrolytic activity has been detected for the latter [5].
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 amino acids on the protein surface, where two tryptophan and one tyrosine would form a planar platform of aromatic residues favouring this binding (149, 154). Domain II folds as a β-sandwich formed by two sheets of four antiparallel β strands each (Figure 8). 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 [155].

(a G2A protein from *Phleum pratense*; PDB 1WHO). In (a), the domain I forms a barrel; amino-acid residues that are conserved in expansins are indicated in the single-letter amino-acid code. Images reproduced with publisher BioMed permission [5].

Figure 8. a) Expansin proposed activity; b) Expansin domain I (the catalytic domain of a GH45 endoglucanase from *Humicolainsolens*; PBD, 2ENG); c) Expansin domain II

Expansins are classified in four families: α-expansins (EXPA), β-expansins (EXPB), α-expansin like-proteins (EXLA) and β-expansin like-proteins (EXLB) [5].

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, the degradation of cell walls during the ripening of fruits and other processes where relaxation of the cell wall is crucial [156-159].

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 [154].

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. Although an HFD motif, that is known to form part of the active site of endoglucanases, has been found in domain I of EXPA and EXPB family members, they do not have hydrolytic activity [5, 160].

The **EXLA and EXLB families** do not have this sequence motif, which suggests that their mode of action differs to that of the other expansins. The EXLA and EXLB families are comprised of proteins identified by sequence analysis which, despite possessing the two- organi-
zation typical of expansins, have a number of divergent sequence features that separate them from the EXPA and EXPB families [161].

Another group included in the expansin superfamily is the expansin-like X family (EXLX), comprising proteins that exhibit weak sequence homology with the domains of EXPA and EXPB members, and identified in organisms other than plants, such as the mucilaginous fungus Dictyostelium and the bacteria Bacillus subtilis, and Hahella chejuensis [161-164].

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 [161]. However other proteins with similar disrupting activity of the cell wall have been reported.

Expansins and expansin-like proteins have been detected in angiosperms such as Arabidopsis thaliana, Oryza sativa, Zea mays and Triticum aestivum, gymnosperms such as pine and poplar, ferns such as Regnellidium diphyllum and Marsilea quadrifolia 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 promote the infection process [165-169].

18.2. Novel proteins with expansin-like activity

Proteins with expansin-like activity called swollenins and loosenins have been identified in ascomycete and basidiomycete fungi such as T. reesei, A. fumigatus and B. adusta [6-8, 170].

In [7], a swollenin gene from T. reesei denominated swo1, was cloned and expressed in Saccharomyces cerevisiae, coding for a protein that modifies the structure of cellulose in swollen regions of cotton fibres (hence the name) without releasing reducing sugars. Swo1 is a fungal expansin-like protein, containing a pollen allergen domain and a cellulose-binding domain.

Proteins with expansin activity could be used to improve the efficiency of cellulose bioconversion processes. For example, a swollenin purified from A. fumigatus has been used in combination with cellulases to facilitate the saccharification of microcrystalline cellulose (Avicel) [8]. In [163] also is described the synergism of an EXLX from B. subtilis in the enzymatic hydrolysis of cellulose and recently, and a new protein with expansin activity from the basidiomycete fungus B. adusta, denominated loosenin (LOOS1] was cloned and characterized [6].

Not only expansins, but also swollenins and loosenin represent good candidate as pretreatment to enhance sugar production from plant biomass. For example, loosenin activity was efficient to release reducing sugars (after cellulase treatment) from Agave tequilana, a crop extensively grown in some areas of Mexico, which shredded fibrous waste is usually burnt or left to decompose. Indeed, A. tequilana fiber became a susceptible substrate for a cocktail of commercial cellulases and xylanases in the presence of LOOS1. Loosenin shows optimum activity at the same pH as most cellulolytic enzymes, opening the possibility to use them as a mixture. This protein is able to relax the structure of cotton, enhancing up to 7.5-fold the
rate of release of reducing sugars from agave fibre. Something similar was observed when a cucumber expansin was incubated with a compound of cellulose and xyloglucans of bacterial origin and occurred a rapid relaxation of the structure of this compound, suggesting that expansins modulate the binding between cellulose fibres and xyloglucans, relaxing or breaking the bonds keeping them together [171].

Given the optimum pH of LOOS1 (pH 5) and other expansin like proteins, they could be applied to processes of saccharification of natural substrates, facilitating the release of reducing sugars together with cellulases. For example, it might be used as an additive to obtain fermentable sugars from pretreated yellow poplar as reported in [172].

In [173], used swollenin as a pretreatment of cellulosic substrates and observed that even in non-saturating concentrations, a significant accelerated hydrolysis occurred. They also correlated particle size and crystallinity of the cellulosic substrates with initial hydrolysis rates, and it could be shown that the swollenin induced-reduction in particle size and crystallinity resulted in high cellulose hydrolysis rates.

It is not surprising that the idea of using plant expansins in saccharification processes has been patented [174-176].

The efficient enzymatic saccharification of cellulose has been a challenge over the past 50 years, mainly due to its crystallinity, which make it a recalcitrance substrate with a high potential to be used as a carbon source.

The bioconversion of cellulose to ethanol is the process where most interest has been focused. Fortunately, increasing of the loosened cellulose surface area by the use of non-hydrolytic proteins, a process known amorphogenesis, would allow access to hydrolytic enzymes making the saccharification process more efficient [177].

19. Conclusions

Cellulose biodegradation represents the major carbon flow from fixed carbon sinks to atmospheric CO₂, this process is very important in several agricultural and waste treatment processes. Also, cellulose contained in plant wastes could be used as a raw material to produce sustainable products and bioenergy to replace depleting fossil fuels. However, one of the most important and difficult technological challenges is to overcome the recalcitrance of natural cellulosic materials, which must be enzymatically hydrolysed to produce fermentable sugars. In order to achieve this goal, new enzymes with cellulolytic activities are being improved and organisms with novel properties have been found. Although the efforts are being directed to improve cellulolytic activity, proteins capable to relax plant cell structure (expansins, swollenins and loosenin) could be used as a biological pretreatment since they would be disrupting crystalline structure of cellulose making it more accessible to the enzymes and enhancing sugar releasing.
Author details

Rosa Estela Quiroz-Castañeda and Jorge Luis Folch-Mallol

*Address all correspondence to: rosa.quiroz@uaem.mx

Biotechnology Research Centre, Autonomous University of Morelos, Cuernavaca, Morelos, México

References


[38] Li Y, Lin M, Davenport JW. Ab initio studies of cellulose I: crystal structure, intermolecular forces, and interactions with water. The journal of physical chemistry. 2011;115:11533-9.


[99] Li DC, Li AN, Papageorgiou AC. Cellulases from thermophilic fungi: recent insights and biotechnological potential. Enzyme Res. 2011;2011:308730.


