3

Plant Cell Walls: Basics of Structure, Chemistry, Accessibility and the Influence on Conversion

Brian H. Davison[1], Jerry Parks[1], Mark F. Davis[2] and Bryon S. Donohoe[2]

[1] *Oak Ridge National Laboratory and BioEnergy Science Center, Oak Ridge, USA*

[2] *National Renewable Energy Laboratory, Golden and BioEnergy Science Center, Oak Ridge, USA*

3.1 Introduction

This book is focused on the pretreatment of plant biomass, a necessary step for efficient conversion of plant cell-wall materials to liquid transportation fuels and other products. Pretreatment is required because it is difficult to access, separate, and hydrolyze monomeric sugars from biopolymers within biomass. Accessible sugars can be further upgraded to products through chemical processes such as aqueous phase reforming or biological routes such as fermentation of the sugars to ethanol. This resistance to degradation or difficulty to release the monomers (mostly sugars) is commonly referred to as recalcitrance [1]. Many methods can be employed to overcome recalcitrance, but the underlying cause of recalcitrance lies in the complex combination and diversity of chemical and structural features of the plant cell walls.

Recent studies by Perlack *et al.* [2,3] estimated that there is approximately 1.4 billion tons of biomass available annually in the United States, which could replace up to one-third of the petroleum-derived fuels currently used. This study determined that 1 billion tons from agricultural lands and an additional 368 million dry tons from forest lands could be sustainably harvested annually. The types and amounts of biomass available for conversion to biomass-derived fuels, whether agricultural residues, forest residues, or purpose-grown energy crops such as switchgrass, willow or poplar, is dependent on geography and climate, as shown in Figure 3.1.

Aqueous Pretreatment of Plant Biomass for Biological and Chemical Conversion to Fuels and Chemicals, First Edition.
Edited by Charles E. Wyman.

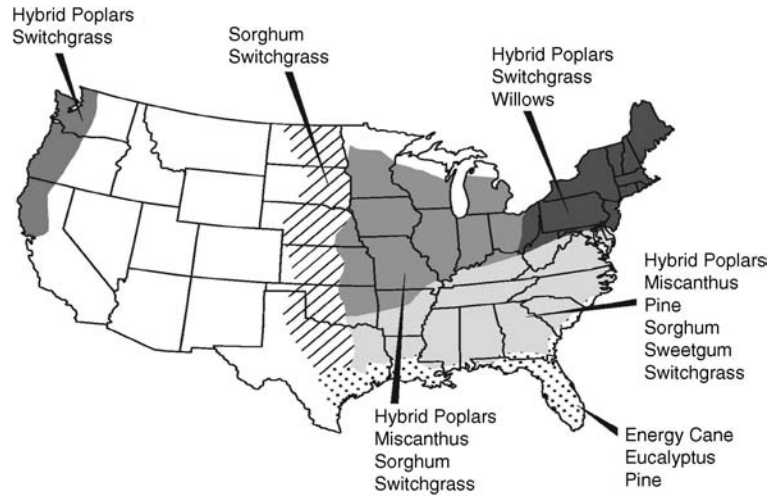

Figure 3.1 *Geographical distribution of non-agricultural residue feedstocks that can be plantation grown for biomass conversion. (Reproduced by permission of Ecological Society of America [4]).*

The wide diversity of biomass feedstocks available for conversion creates unique challenges for harvesting and conversion due to variability in both physical and chemical properties of the feedstocks. Many of the biological barriers to conversion of biomass to fuels were covered in a recent US Department of Energy Roadmap [5]. The three main classes of herbaceous, hardwood and softwood feedstocks have unique aspects of the cell-wall anatomy, macromolecular architecture, and polymer chemistry. These differences extend down to each species within these broader categories, and such differences in cell-wall composition and architecture contribute to the phenomena of biomass recalcitrance at multiple scales. Different pretreatment and bioconversion strategies have been developed to overcome biomass recalcitrance, and their success has been shown to vary with the biomass feedstock [6]. For example, dilute acid pretreatments have been shown to be very effective for the pretreatment of herbaceous materials but fail at the same thermochemical severity to be an effective pretreatment for woody feedstocks [7]. Severity is a function of temperature, duration time, and concentration (i.e., acid).

In this chapter, we will briefly survey the basics of plant cell-wall structure and composition as well as the variability of lignocellulose in plants that relate to pretreatment. We will then return to the chemical components of lignocellulose, first lignin and then cellulose. The discussion of hemicellulose will focus on its branching, decorations, and cross-linkages. We will also discuss the influences of the microscopic structure of biomass: the tissue, cellular, and macromolecular-scale architecture that contributes to recalcitrance. Finally, we will survey how molecular modeling and simulation is being used to explore biomass structure and inform the nature of recalcitrance in lignocellulosic biomass. Given the diversity of available feedstocks, the inherent variability of their plant cell walls, and the impact this variability can have on the recalcitrance of biomass, a primer on lignocellulose is appropriate to set the stage for the following chapters. The reader is referred to a modern textbook on the subject of plant cell walls for more detailed information [8].

3.2 Biomass Diversity Leads to Variability in Cell-wall Structure and Composition

The primary carbohydrate comprising the cell wall common to all plants is cellulose, a β-1,4 -linked glucose polysaccharide. Cellulose in the cell wall forms long, oriented microfibrils, which may coalesce into

larger and longer fibrils [9]. The cellulose microfibrils are hydrophobic and can be highly crystalline, features that contribute greatly to the recalcitrance of biomass. For example, the 100 crystal face of a microfibril is more hydrophobic than the other faces and selectively binds the CBD (cellulose binding domain) [10]. Hemicelluloses are a class of polysaccharides that have variable compositions and structures depending on the plant source. For example, hemicelluloses isolated from herbaceous grass species, such as switchgrass, are composed of glucuronoarabinoxylans which are complex, branched polysaccharides composed mainly of pentose (five carbon) sugars. On the other hand, hemicelluloses in softwood species, such as pines, are predominantly composed of galactoglucomannans, which have a backbone of β-1,4 linked D-mannopyranose and D-glucopyranose units. Hemicelluloses are amorphous, branched, single-chain polysaccharides and are not particularly recalcitrant to conversion. Hemicellulose chains are thought to interact with more than one cellulose fibril so that they form non-covalent cross-links between cellulose bundles.

The final main structural polymer, lignin, is a complex three-dimensional polyphenolic polymer that partially encases the plant cell-wall polysaccharides and cellulose microfibrils in lignified (i.e., secondary) plant cell walls. Lignin is generally not found in the primary wall of newly formed cells. Lignin provides mechanical and elastic support, facilitates water and nutrient transport, provides a chemical barrier to microbial pathogens, and is also understood to be a key contributor to recalcitrance.

In addition to these three main polymers of lignocellulose, there are other non-structural components within the plant cell wall. These components, such as extractives, protein, ash, and pectin, vary greatly with species, tissue, plant maturity, harvest times, and storage, and are greatly influenced by environmental factors and stress. Extractives are a complex mixture of compounds which can include sugars, terpenoid compounds, and monolignols. A major theme of biomass recalcitrance that adds to the complexity is that many of the cell-wall components such as lignin, hemicellulose, and proteins can cross-link with each other to create a complex matrix that is resistant to chemical or biological attack.

Depending on the plant species, there is considerable variation in the relative amounts of each of the structural components, cellulose, hemicellulose, and lignin, within the cell walls (Table 3.1), as well as physical properties such as cell-wall thickness and porosity (Figure 3.2). The composition, structure, and interactions of the biopolymers composing the lignocellulosic matrix serve many interrelated functions for the plant, including the primary function of providing structural features that create mechanical support, allowing for internal transport of water, nutrients, and photosynthate throughout the plant. In physical terms, lignocellulose is a molecular-level structured composite material. An imperfect analogy to the macroscopic world is reinforced concrete. Here, cellulose fulfills the role of the steel rods (i.e., rebar) that provide strength over long distances. Hemicellulose represents the wire mesh or cable that wraps around the cellulose rods, providing extra strength and linkages. Lignin acts as the concrete that fills the remaining gaps and sets, holding everything in place while excluding water from the polysaccharide environment. It is important to note that terrestrial plants have co-evolved with herbivores and cell-wall degrading microbes; millions of years of evolution have therefore created living plants that are very resistant to both mechanical and biological decay.

Table 3.1 *Typical biomass composition for a variety of feedstocks, highlighting the diversity in chemical make-up (Wiselogel* et al.*, 1996) [11].*

Feedstock	Cellulose	Hemicellulose	Lignin
Corn stover	36.4	22.6	16.6
Wheat straw	38.2	24.7	23.4
Rice straw	34.2	24.5[a]	11.9
Switchgrass	31.0	24.4	17.6
Poplar	49.9	25.1	18.1

[a] Xylan value only was reported.

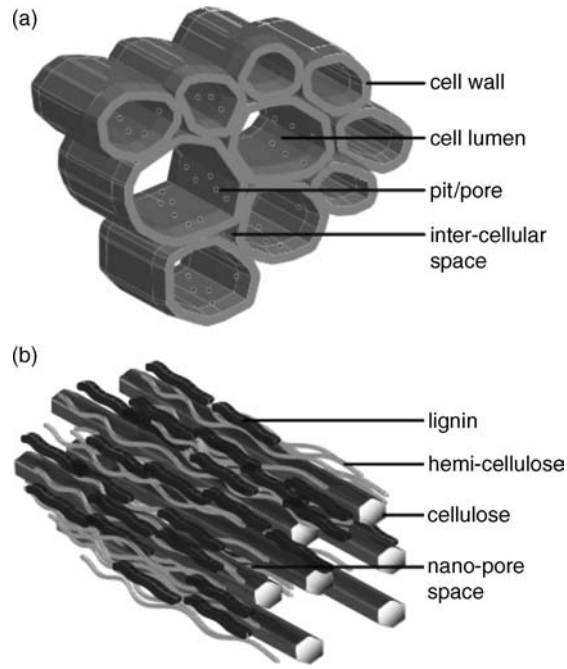

Figure 3.2 *Cartoon depiction of (a) cellular-scale and (b) macromolecular-scale structure and composition of lignocellulosic biomass. Several scales of porosity from the cell lumen to the nanopores between coated microfibrils are also indicated. (Reprinted with permission from Viamajala* et al. *[12] © 2010 Springer Science+Business Media B.V.).*

Previous research has contributed to a significant understanding of the biosynthesis of the various monomers from which the biopolymers found within the cell wall are built, and our knowledge of how the monomers are transported and assembled to create the plant cell wall is rapidly developing but is still incomplete. However, while monomer synthesis pathways are largely known (Figure 3.3a and c), the understanding of formation and polymerization into a specific part of the cell wall is incomplete. A plant cell wall consists of multiple layers formed during genesis and development of the cell. The composition and structure of these wall layers change depending on cell type, tissue, and location. The mature secondary cell wall contains most of the lignocellulosic biomass, but its structure and organization begins in the primary cell wall. Albersheim *et al.* [8] provide an excellent review and summary of the synthesis and nature of primary and secondary cell walls. Some of the complexity of the primary and secondary cell walls is shown in Figure 3.3.

3.3 Processing Options for Accessing the Energy in the Lignocellulosic Matrix

The challenge in pretreatment, based on the model of lignocellulose as a molecular-level structured composite material, is to disassemble this composite into its valuable constituent monomers (i.e., sugars and aromatics) without loss or damage to the sugar monomers released by pretreatment. The sugars, or polysaccharides, must be preserved in order to be a feedstock for further specific bioconversion processes into fuels or chemicals. The lignin might be used as an energy source for power but also has significant potential for higher-value uses. This is one of the great benefits of the so-called biological approaches to lignocellulose conversion: the ability to make specific products and have potentially valuable co-products.

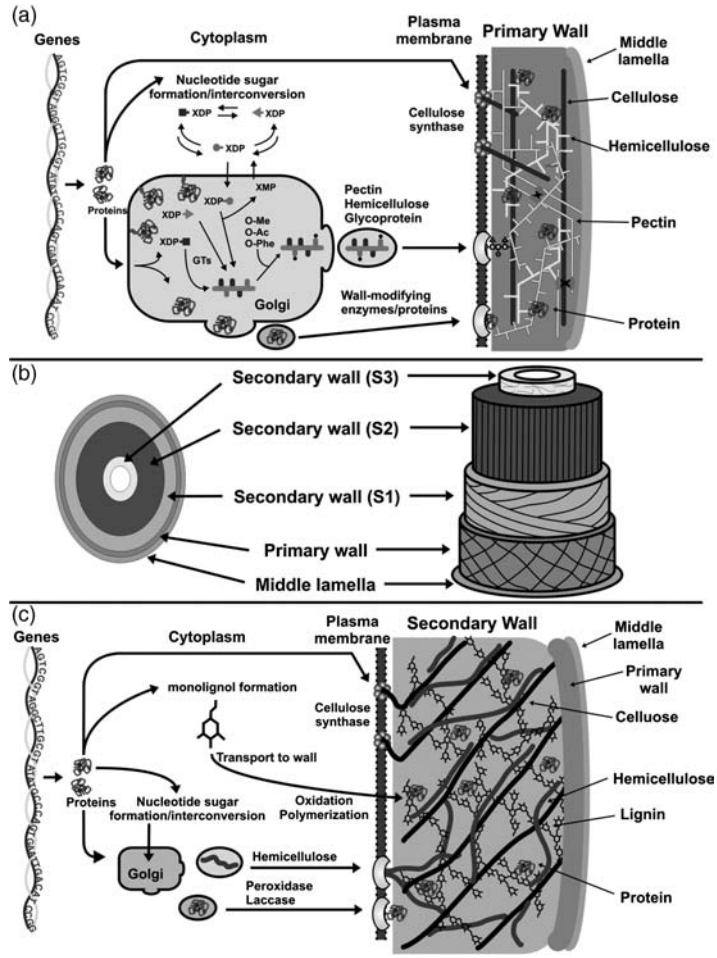

Figure 3.3 *The plant cell wall. (a) Biosynthesis of primary cell walls. (b) Plant cell-wall layers. (c) Biosynthesis of secondary plant cell walls. (Reprinted with permission from Mohnen, Bar-Peled, Sommerville, Chpt 5, page 95, in Biomass Recalcitrance: Deconstructing the Plant Cell Wall for Bioenergy. Michael E. Himmel, ed., Wiley-Blackwell publishing [1]).*

Hemicellulose is easier to remove from the cell-wall matrix than the other components during pretreatment under acidic conditions. The downside of this ease of removal from the cell wall is that the hydrolyzed free sugars and oligosaccharides will continue to react under these pretreatment conditions and degrade into undesirable compounds for biological conversion. These include furans which, in addition to being a loss of sugars and the overall process yield, are also inhibitory to subsequent biological conversion processes. While most of the hemicellulose can be easily removed by pretreatment, the covalent cross-linkages appear to create a residue of hemicellulose or lignin carbohydrate complexes, which may continue to shield or partially protect cellulose microfibrils from attack. Hemicellulose can be substituted at hydroxyl groups with different side-chain chemical groups, primarily acetyl groups, but ferulic acid and other side groups are significant as they may be more resistant to hydrolysis. The acetyl groups easily hydrolyze under thermal or acidic conditions to form dilute acetic acid, which serves to lower the pH of the pretreatment liquor and increase the rate of hydrolysis. Thus, even a hot water pretreatment has the benefit of actually being a

mild autocatalyzed dilute acid pretreatment. The negative aspect is that the acetate formed must be neutralized before further processing steps such as enzymatic hydrolysis and also can be inhibitory to many fermentative microorganisms.

We note that there is another family of approaches to dealing with the difficult problem of biomass recalcitrance. These are the thermochemical approaches such as pyrolysis and gasification. These processes deal with the intimate associations of lignocellulose as a composite material by foregoing preservation of the cell-wall monomers. In pyrolysis, the solid biomass is "cracked" at temperatures of 450–650 °C in the absence of oxygen into a "biocrude," a very complex mixture of organics and water that is highly acidic and oxygenated. Considerable upgrading of pyrolysis oil to reduce the acidity and oxygen content is necessary before the material is suitable either as a blend stock for a petroleum refinery or as a finished fuel. In gasification, the goal is for all organic carbon to be converted into carbon monoxide and hydrogen, a mixture called syngas, which can then be further upgraded by catalysis or fermentation to fuels, chemicals, or other products. These approaches have the benefits of speed and higher theoretical product yields, but with the added challenges and costs of building the desired specific products from either "biocrude" or single-carbon feedstocks. They also sacrifice some of the capacity to make specific value-added by-products from lignin, sugar, or fiber that can be preserved with biological approaches. Regardless of the approach used, all biomass conversion technologies have to overcome the multiscale complexity of biomass particles. The organization of plant tissues, cells, and cell-wall architecture creates barriers to heat and mass transport and inhibits accessibility to enzymes.

3.4 Plant Tissue and Cell Types Respond Differently to Biomass Conversion

Plants are complex multicellular structures that have evolved multiple mechanisms to resist degradation by microbial enzymes. The harvested and milled chunks of biomass that arrive at a biorefinery comprise a heterogeneous mixture of tissue types including leaves, stems, cobs, and chipped wood due to the variability discussed in Section 3.2. Each of these tissues contains its own mix of cell types that respond differently to pretreatment and saccharification. If we consider a typical biomass feedstock grass species such as switchgrass, the tissue types can be simplified into the categories of vascular tissues and ground tissues [13]. Vascular tissues contain the elongated cells that transport water and nutrients through the living plant and those cells are themselves surrounded by layers of thick-walled supporting cells. These supporting cells (termed sclerenchyma or fiber cells) with their thickened, lignified secondary cell walls comprise most of the mass and therefore contain most of the carbohydrates in grass biomass particles. The other thick-walled cell types in vascular bundles of grass species are xylem cells. While the function of xylem cells in the living plant is distinct from the function of the supporting fiber cells, both cell types share the characteristic of a thickened, lignified cell wall that is especially recalcitrant during pretreatment and conversion.

In contrast to the xylem and supporting fiber cells in grass species such as switchgrass, the space between the vascular bundles and the epidermis is filled with ground tissue. Ground tissue makes up most of the stem's volume and is composed of thin-walled parenchyma cells that typically lack a secondary wall, are not lignified, and are relatively amenable to pretreatment and saccharification processes. It is not uncommon to see pretreated grass stems where the ground tissue is severely impacted and appears to have largely solubilized, while the vascular bundles and epidermis appear nearly unchanged [14]. Wood parenchymas consist of cells whose primary function is the conduction and storage of food materials. There are three types of wood parenchyma: ray parenchyma which make up the bulk of ray tissue; epithelial parenchyma which surround resin canals and therefore are only present in coniferous woods; and axial parenchyma which extend along the grain in the form of strands.

Different plant species have long been known to have different responses to pretreatment. However, evidence is increasing that there is significant variability even within a species. Studer *et al.* [15] recently

presented analyses of a number of *Populus* samples for hot water pretreatment and enzymatic sugar release. These screening results showed variability from about 0.2 to almost 0.7 g carbohydrate released per gram of biomass. Surprisingly, some lower lignin content samples showed appreciable hexose release after only hot water treatment.

3.5 The Basics of Plant Cell-wall Structure

The thickness of the walls in xylem and fiber tissues is due to the formation of an extensive secondary cell wall as the cell matures. The sequence of events in the development of these cells is that after the cells have ceased growing by elongation, they begin to deposit a multilamellar secondary cell wall toward the cell lumen side of the primary cell wall [16]. At some point during and following secondary wall formation, the entire wall is infused with lignin monolignols [17]. The lignin polymerizes into the spaces in the existing wall and, consequently, significantly lowers its porosity [18]. Finally, the cell dies and its contents are absorbed, leaving a rigid water-impermeable cell-wall barrier. The biomass conversion perspective of the plant cell wall is, by necessity, heavily skewed toward the thick, lignified, secondary cell walls. These are the walls that harbor most of the mass and therefore most of the structural sugars in biomass. Unfortunately, they are also the most recalcitrant.

Chemically, the plant cell wall is understood to be composed of cellulose, hemicelluloses, pectins, proteins, and often lignin. However, it is the complex intermingling of cross-linked layers brought together largely by self-assembly and template-assembly processes that create the complexity and resilience of the plant cell wall. The complex macromolecular architecture of the cell wall is a result of the self-ordering properties of cell-wall polymers generating a high degree of structural organization and complexity [19]. Three themes that govern the architectural plan of the plant cell wall are: (1) fibrous structural units embedded in an amorphous matrix; (2) covalent and non-covalent cross-linking; and (3) a polylamellate construction [20].

The plant cell wall is sometimes envisioned as three intermingled networks that are extensively restructured during development. One is a protein network formed by the various classes of glycoproteins that contribute a scaffold to initial cell plate [21]. This network organizes cell plate membranes and the incorporation of initial cell-wall components as they are delivered to the cell plate by Golgi-derived vesicles. This original scaffold gets extensively remodeled and recycled during cell growth and maturation. While the protein network is not usually considered a significant contributor to recalcitrance, its role in establishing a template for cell-wall construction is still important. The second network is the pectic polysaccharide network. In the primary cell wall, the pectin network is credited with dictating key structural and mechanical properties of the wall including water content and porosity [22]. Again, this network is remodeled during growth and development and is usually not of critical concern for biomass conversion because the pectic polysaccharides are not especially recalcitrant to hydrolysis and extraction by pretreatment.

Finally, and arguably most importantly from the perspective of bioconversion, is the network of cellulose and hemicellulose. The main structural unit of the wall, the cellulose microfibrils thought to be composed of *c.* 36 chains of cellulose, are 5–10 nm in diameter, many micrometers in length, and spaced 20–40 nm apart [23]. They are embedded in, and non-covalently cross-linked by, a matrix created by the other major wall polysaccharides: the hemicelluloses. The cellulose/hemicellulose network is intermeshed with, but not cross-linked to, the pectic polysaccharide network. The final porosity of the cell wall is 5–10 nm [24]. This porosity is sufficient to allow some diffusion of very small proteins but too small to allow significant accessibility to cellulolytic enzymes. Mass transfer considerations suggest that pore size should be in the range of 50–100 nm to allow sufficient penetration of enzymes into cell walls. One of the primary goals of pretreatment is to increase cell-wall porosity for effective enzyme transport and penetration to the cellulose surface.

3.6 Cell-wall Surfaces and Multilamellar Architecture

Most of the biomass that is processed for bioconversion is dead plant tissue composed of chambers enclosed by cell walls referred to as lumen (see Figure 3.2). The lumen are the chambers that housed the cellular organelles and machinery that originally constructed the cell walls while the plant was alive and actively growing. Cell-wall components are typically synthesized within the living cell (by the secretory organelles or by transmembrane proteins) and then transported into the cell wall where the final steps in assembly and maturation occur. In the case of xylem and fiber cells, the cells themselves undergo programmed cell death and are recycled by the plant. The cell lumen represents the largest scale of porosity that is relevant to biomass recalcitrance. The cell lumen is in the size range of tens of micrometers in diameter and can be many hundreds of micrometers in length. In their longitudinal dimension they are often designed to transport materials; transport is more limited in the radial direction, however. This is the scale of porosity that many people imagine when they think about a biomass particle because it is the scale that is nearly visible to the naked eye. These pores do have some impact on recalcitrance because the cell lumen and intercellular spaces can trap air that may impede the bulk flow of pretreatment chemistries throughout a biomass particle.

However, the macroporosity is not the critical barrier for cellulolytic enzymes. For these enzymes, entering the cell lumen is relatively easy but only provides the enzyme access to one surface of the multilayered cell wall. Somewhat surprising to the casual observer used to viewing only 2D slices of biomass, the intercellular spaces present an additional and important access point to the cell-wall surface. The intercellular space, formed at the juncture among adjacent cells forming cell corners, is a continuous 3D space throughout the plant tissue, which provides critical access to the "back side" of the cell walls (Figure 3.4). By exploiting this route, enzymes gain access to the middle lamella and primary cell walls that have a different composition from the secondary cell wall on the cell lumen side. One of the architectural motifs of the plant cell wall is that it is constructed as multiple, concentric layers or lamellae [25]. Single lamellae are one cellulose microfibril thick but, in transmission electron microscopy (TEM) micrographs, the lamellar structure is often evident in the staining pattern of layers of varying thickness.

The structural impact of thermochemical pretreatment and enzymatic saccharification on biomass cell-wall surfaces has been studied by multiple groups [14,26–29]. These investigations have revealed the

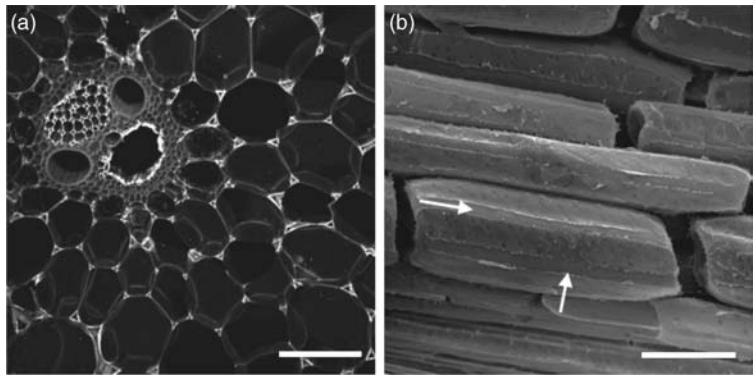

Figure 3.4 *(a) Confocal scanning laser microscopy and (b) scanning electron microscopy micrographs revealing the intercellular spaces within a biomass particle. (a) A section of untreated corn stover stem labeled with the JIM5 pectin antibody highlights the cell corners. (b) In a pretreated corn stover sample the cells have become disjoined, revealing the cell corner surfaces as regions of coalesced lignin accumulation (arrows). Scale bars = 50 μm. (Images courtesy of NRELs BSCL, unpublished).*

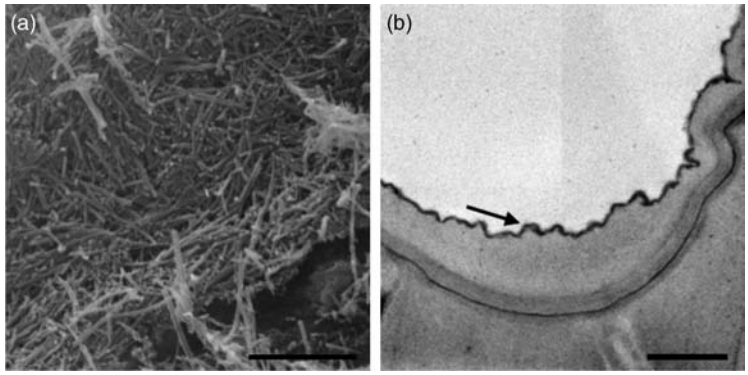

Figure 3.5 *(a) Scanning electron microscopy and (b) transmission electron microscopy micrographs revealing the impact of pretreatment on eroding the surface of switchgrass stem cell walls. (a) A lime pretreatment showing partially unsheathed layers of microfibrils. (b) An aqueous ammonia pretreated sample displays an irregularly eroded surface (arrow). Scale bars: (a) 5 μm; and (b) 1 μm. (Adapted from Donohoe* et al. *[29] © 2011, Royal Society of Chemistry).*

different modes of cell-wall deconstruction that directly affect the enzyme-accessible cell-wall surface area. One mode of increasing the accessible surface area of secondary cell walls is to etch away the lignin-containing matrix at the lumen surface. Etching is seen most commonly in pretreatments that employ base chemistry. This etching treatment typically does not have an impact on the wall beyond 10–20 nm into the surface, but can still provide a substantial increase in initial enzyme binding (Figure 3.5).

Another mode of increasing surface area during pretreatment is by delamination. Delamination is often seen in samples treated with acidic pretreatments, is enhanced by flowthrough reactors, and can increase the accessible surface area throughout the 3D volume of the cell wall. A final mode of increasing accessibility in pretreated biomass is through rapid pressure drops at the end of pretreatment, such as steam gun or AFEX protocols. These systems can cause delamination of the wall and also generate additional new micropore structure within the walls [29].

3.7 Cell-wall Ultrastructure and Nanoporosity

The intercellular spaces formed at the junction between cells create an interconnected network throughout plant tissues, but the cell lumen are in fact also a continuous interconnected space. In the living plant, cell-to-cell communication is mediated through plasmodesmata, the *c.* 50-nm-diameter plasma-membrane-lined tubules that connect adjacent cells through the cell-wall barrier. In senesced dried biomass, the remains of the plasmodesma are seen as pores or pits in the cell wall. Pits are regions in the cell wall where no secondary cell wall was deposited and an open pore is maintained between adjacent cell lumen (Figure 3.6). Pits may provide an escape route for the air trapped in the cell lumen of dry biomass; however, these small openings can close up due to cell-wall drying and can become occluded by relocalized lignin globules during pretreatment. Pits are an ultrastructural feature of the cell wall that can be considered part of the nanoscale porosity of a biomass particle. However, even though pits and pit fields are only 20–100 nm in diameter, they still do not represent a fundamental barrier to cellulolytic enzymes. In fact, between the interconnected intercellular space created at the cell corners and the interconnected cell lumen space created by the cell-wall pits, enzymes have good initial access to the two surfaces of the cell wall.

It is clear from multiple lines of research that the most fundamental barrier to effective enzymatic conversion of lignocellulosic biomass to sugars is enzyme accessibility to a reactive surface. At one level,

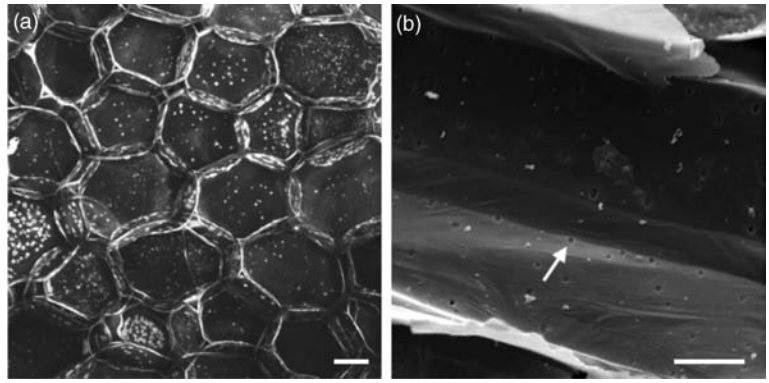

Figure 3.6 *(a) Confocal scanning laser microscopy and (b) SEM micrographs showing pits in corn stover stalk cell walls. (a) Section of dilute acid pretreated corn stover labeled with a fluorescently tagged CBM3 probe highlights the cell pits. (b) At higher resolution, an SEM micrograph reveals pits that have a pit membrane (arrow) with smaller micropores. Scale bars = 20 μm. (Unpublished images courtesy of NRELs BSCL).*

accessibility is a simple concept of physical dimensions and barriers. The architecture of the plant cell wall is a complex biomatrix with porosity on the scale of 5–10 nm. While pectins may have been the controlling factor in primary cell-wall porosity, lignin becomes the porosity gatekeeper in cells with lignified secondary cell walls. Cel7a, the main exoglucanase in a *Trichoderma reesei* enzyme mixture, is a *c.* $4 \times 5 \times 13$ nm enzyme [30]. It would seem that the native porosity of plant cell walls would be sufficient to let these enzymes pass. In fact, the native cell wall is not nearly porous enough. Part of the reason is that the enzyme brings along with it a hydration shell that increases its effective size, requiring a much larger pore space for effective penetration. This restricted nanoporosity of the cell-wall matrix is a key aspect of biomass recalcitrance.

The primary role of pretreatment is to increase the porosity and accessibility of biomass cell walls to cellulolytic enzymes. Although enzymes can thoroughly penetrate cell walls after high severity pretreatments, incomplete cellulose conversion by cellulases suggests that additional barriers exist at the macromolecular level. One potential barrier is obstruction of the active face of cellulose microfibrils by residual lignin or hemicellulose that prevent cellulases from binding to cellulose. The critical concept therefore becomes not just physical accessibility, but accessibility to the right substrate surface.

3.8 Computer Simulation in Understanding Biomass Recalcitrance

In this section we provide an overview of selected recent applications of computer simulation applied to understand plant cell-wall chemistry and structure at the molecular level and inform the origins of recalcitrance. Various publications are highlighted to provide examples from the recent literature that show how simulation is being used to help overcome recalcitrance. Beckham *et al.* discuss the use of computation for understanding enzymatic degradation of biomass; in addition, we discuss the role of simulation in understanding lignocellulosic structure [31].

3.8.1 What Can We Learn from Molecular Simulation?

Many important biomolecular phenomena cannot be studied directly using currently available experimental techniques. Computer simulation is particularly well suited to such systems. Simulation techniques have

advanced to the point that they can provide atomic-resolution insight into the structure and dynamics of macromolecular assemblies such as lignocellulose. Other techniques can be used to calculate highly accurate molecular structures and properties of lignin. For example, accurate strengths (energies and enthalpies) of key linkages in lignin are among the many useful quantities available through computational means [32].

The standard tools of molecular simulation are classical molecular dynamics (MD) simulation and quantum chemistry. In MD, a molecular system is propagated in time according to Newton's second law ($F = ma$), usually under temperature and pressure control. The potential energy of the system is governed by a force field, which includes all the energetic terms needed to describe the interactions among all atoms in the system. A typical system, consisting of solute(s), solvent, and ions, might contain tens to hundreds of thousands of atoms, but simulations of multimillion-atom systems are now possible. Timescales on the order of nanoseconds to microseconds are achievable for systems of this size. A challenge is that many of the systems of interest here are at the upper ranges of these atom numbers and timescales.

Unlike classical MD, quantum chemical calculations consider electrons explicitly. Quantum chemistry is aimed at calculating the molecular wave function (or electron density) and properties deriving from it. Such calculations can yield very accurate structures and energies of molecular systems and are particularly useful for studying chemical reactions. Numerous quantum chemical methods and approximations of varying accuracy and computational cost exist. Most commonly, density functional theory (DFT) methods are used because they can achieve high accuracy at reasonable computational costs. Computationally more demanding *ab initio* methods are also used, often as benchmarks to assess the accuracy of more approximate methods such as DFT. The following studies exemplify how the combination of theory and experiment can yield key insights into lignin structure and reactivity that might not be obtainable otherwise.

3.8.2 Simulations of Lignin

Classical and quantum chemical approaches have been used to study lignin biosynthesis and degradation. Durbeej used DFT calculations to investigate the spin density distributions of coniferyl radical and dilignol radicals [33]. Spin distributions show where an unpaired electron prefers to reside in a radical species. For coniferyl radical, the author found that the largest fraction of the unpaired electron spin was localized on the phenolic oxygen (O4), although there was also significant spin localization on the β carbon atom of the allyl alcohol group and the aromatic ring carbons (C3 and C5) at the *meta* positions relative to the allyl group [34]. As sites with the highest spin density are expected to be the most reactive, these results provide a possible explanation for the prevalence of β-O-4 linkages in natural lignin.

Durbeej and Eriksson [35] performed DFT calculations to investigate the thermodynamics and kinetics of dehydrogenative coupling of coniferyl species (radical-radical and radical-alcohol) to form guaiacyl β-O-4 dimers. They obtained reaction energies of more than *c.* 20 kcal/mol for the radical-radical coupling process, quantifying the favorability of monolignol radical coupling reactions.

Computational studies were carried out on lignin pyrolysis reactions of phenethyl phenyl ether (PPE), which is the simplest model of a β-O-4 linkage in lignin. Jarvis *et al.* [36] used accurate *ab initio* calculations and Transition State Theory to calculate rate constants for several PPE pyrolysis reactions. They observed very good agreement between computation and photo-ionization mass spectrometry and matrix-infrared (IR) experiments. Beste *et al.* [37] used DFT calculations to predict the selectivity of PPE pyrolysis reactions for the β-O-4 linkages. They calculated structures and energies of transition states for these processes, and the computed values were found to be in good agreement with experimental data.

In a recent computational study [38], bond dissociation enthalpies were computed using DFT for 69 lignin model compounds containing β-O-4, α-O-4, β-5, and biphenyl bonds. A key finding of that work was the prediction that oxidation of primary and secondary alcohol groups on the alkyl substituents of lignin macromolecules will result in lower bond dissociation energies. Exploiting this property may lead to more

facile deconstruction or conversion of lignin to useful bioproducts. In a similar study, Parthasarathi *et al.* [39] quantified the homolytic bond dissociation energies of various C—C and C—O (ether) bonds in a series of 65 lignin-like model compounds. They observed a strong inverse relationship between the calculated length of a particular C—C bond and its bond strength: the longer the bond, the easier it is to break. However, no similar correlation was found for ether bonds. Instead, electronic and steric effects arising from the presence of methoxy substituents were found to have significant effects on ether bond strengths. Ether linkages were found to be weakened significantly when methoxy substituents are present at both *ortho* positions on the phenyl rings.

Sangha *et al.* [40] used DFT calculations to carry out a comprehensive study of radical-radical coupling reactions involving *p*-coumaryl, coniferyl, and sinapyl alcohol radicals. Coupling reactions with β-O-4, β-β, and β-5 linkages were computed to have the most favorable reaction enthalpies, whereas dimers with 5-O-4, 5-5, and β-1 linkages were less favorable by *c.* 5–20 kcal/mol. The authors found that the presence of *p*-coumaryl radicals has a significant effect on product regioselectivity, particularly in reactions involving coniferyl radical. In the presence of p-coumaryl radicals, coniferyl radicals enhance the favorability of C—C bond formation relative to C—O interunit bonds. The study therefore reinforced strategies involving modification of the composition of lignin in plants to improve sugar release from lignocellulosic biomass.

3.8.3 Simulations of Cellulose

Bergenstråhle *et al.* [41] used MD to study, at the molecular level, why cellulose is insoluble. They calculated the free energy required to separate solvated oligomers of glucose, cellobiose, cellotriose, and cellotetraose in different orientations. The important finding from their work was that inter-oligomer hydrophobic association and hydrogen bonding are responsible for the insolubility of cellulose. Specifically, they found a high entropic cost for hydrating glucose hydroxyl groups and no increase in configurational entropy of a short cellulose chain in going from the crystalline to the dissolved state. Taken together, these results clearly show why the crystalline state is strongly favored over the solvated state.

Matthews *et al.* [42] used MD to study models of microcrystalline cellulose Iβ. During their simulations, they observed significant conformational changes from the initial state of their models including a twist along the axis of the microcrystal, a significant tilt in the sugar rings, and significant interlayer hydrogen bonding. The most striking result of the study was the structuring of water molecules near the cellulose surface. The local water density at the interface was 1.3 times that of the bulk solvent, due to strong hydrogen bonding with hydroxyl groups of the glucose units. Based on the simulations and experimentally observed slow rates of enzymatic hydrolysis, the authors hypothesized that the local water density near the cellulose surface might create a barrier that inhibits the approach of cellulase enzymes. The detailed insight obtained from this study is difficult to obtain from experiments.

In a recent study, Matthews *et al.* [43] used available X-ray crystal structures to construct a model of a hydrated cellulose Iβ microfibril to study structural changes resulting from thermal treatment. They performed MD simulations at 500 K for 100 ns and observed a shift from a 2D to a 3D hydrogen bonding network and concomitant loss of twist in the structure. The study showed how changes in local conformations and non-bonded interactions can translate into large-scale structural metamorphoses. Their results provided atomic-level insight and helped explain several previous observations from X-ray diffraction, spectroscopy, calorimetry and thermogravimetry; they also suggested follow-up experiments based on their findings and simulations.

Beckham *et al.* [44] used MD simulations to examine the morphology dependence on enzymatic cellulose deconstruction at atomic detail. They calculated the thermodynamic work required to decrystallize a single chain from the surface of four different cellulose polymorphs. It was found that the required work decreased in the order $I\beta > I\alpha > III_I > II$, consistent with experimental observations that the synthetic

polymorphs cellulose II and III_I are less recalcitrant to cellulolytic degradation by enzymes than cellulose I. Possible molecular-level explanations for this behavior in cellulose II and III_I include increased chain fluctuations, fewer inter-chain hydrogen bonds, and greater exploration of conformational space upon decrystalization. The authors pointed out the importance of future experiments that will reveal the shapes of cellulose microfibrils at nanometer length scales, which will inform subsequent simulation studies.

Chundawat *et al.* [45] used MD simulation to show how ammonia pretreatment of cellulose Iβ induces conversion to the less recalcitrant cellulose III_I polymorph. The structural rearrangement resulted in up to fivefold enhanced rates of saccharification. Analysis of MD trajectories revealed distinct differences between the two polymorphs in the torsional states of hydroxymethyl groups, as well as the total number and types of hydrogen bonds. In addition to revealing molecular-level insight into the physical origins of recalcitrance, this work may lead to improved pretreatment and enzymatic protocols for more efficient production of biofuels and bioproducts.

3.8.4 Simulation of Lignocellulosic Biomass

Besombes *et al.* [46] constructed a model of cellulose Iβ and studied the adsorption of a guaiacyl β-O-4 dimer onto various surfaces using MD. They dissected the interaction energies and found that van der Waal's interactions were the most important for adsorption of the phenolic rings of lignin onto the (200) face of the cellulose surface, whereas hydrogen bonding was also important for adsorption to other surface faces. These results were compared with Raman and IR spectroscopy and photoconductivity experiments, and the authors found them to be consistent. They concluded that lignin prefers to be oriented parallel to the surface of the secondary cell wall and that adsorption is largely dispersive in nature.

Schulz *et al.* [47] developed a method to perform accurate and efficient MD simulations of multimillion-atom biological systems on supercomputers. To test their implementation, they used a previously constructed model of cellulose Iβ, which consisted of 36 cellulose chains and was 80 glucose units in length (*c.* 40 nm). They were able to achieve 30 ns/day of MD on the 3.3 million-atom system using the Jaguar supercomputer. One of the major contributions of their work was the development of an atomistic MD method that helps bridge the gaps in size and timescales between theory and experiment.

3.8.5 Outlook for Biomass Simulations

Computer simulations enable studies of molecular detail that are inaccessible to currently available experimental techniques. Here, we have provided a brief summary of recent results from the literature demonstrating a few specific examples of how simulation can complement experimental studies. A challenge remains in how to simulate a "complete" MD model of cellulose, lignin, and hemicellulose, as well as their interactions with enzymes and chemical and thermal reactions, although progress continues to be made. Recently, models of secondary plant cell walls incorporating cellulose, xylan, water and lignin were constructed and MD simulations carried out [48]. A further challenge is how to relate these simulations to the different plant cell-wall types and structures. As simulation methods continue to improve and more studies are carried out, our understanding of the molecular basis for recalcitrance will lead to improved technologies for producing biofuels and bioproducts.

3.9 Summary

The diversity of plant anatomy, cellular structures, and cell-wall chemical composition present in herbaceous and woody crops gives rise to a broad distribution of biomass feedstocks containing unique features that can assist the conversion of the feedstocks into useful fuels, chemicals, and other products. Over time,

plants have evolved cellular structures and modified their cellular composition to resist insects and microbial attacks, and this recalcitrance to decay prevents easy conversion of plant materials to products. The goal of pretreatment is to open up the cellular structures at a molecular lever to allow access by enzymes that can selectively release sugars of the carbohydrates components, often primarily cellulose. Pretreatment conditions need to be optimized to remove the less recalcitrant plant components such as the hemicelluloses while minimizing degradation reactions that reduce sugar yields and create inhibitory compounds that can interfere with further downstream biological processing. There is a large body of knowledge that has been assembled on the broad details of plant cell-wall chemistry and structure. This information shows that variability exists at the species, individual, tissue, and molecular levels; it also shows that this variability is bounded by some common features. However, we still lack a systematic understanding of how the diverse plant cell-wall features are impacted during pretreatment, and this lack of knowledge impacts our ability to develop both improved feedstocks and advanced biological processes that can efficiently and economically produce biofuels and other products. Challenges remain to measuring and detecting the specific structure and chemical changes occurring to the cell wall during pretreatment due to both plant heterogeneity and the lack of analytical methods to detect changes at the molecular level. To bridge the gap until the next generation of analytical methods are developed that can measure changes at the molecular level, computer simulations can be used to provide insight into both cell-wall construction and deconstruction. Several of the chapters that follow (Chapters 6–11) will illustrate how specific pretreatments attack different cell-wall chemistries and structures.

Acknowledgements

This work was supported and performed as part of the BioEnergy Science Center, managed by Oak Ridge National Laboratory. The BioEnergy Science Center is a US Department of Energy Bioenergy Research Center, support by the Office of Biological and Environmental Research in the US DOE Office of Science.

References

1. Himmel, M. (ed.) (June 2008) *Biomass Recalcitrance: Deconstructing the Plant Cell Wall for Bioenergy*, Wiley-Blackwell.
2. Perlack, R.D., Wright, L.L., Turhollow, A., Graham, R.L., Stokes, B., and Erbach, D.C. (2005) Biomass as Feedstock for a Bioenergy and Bioproducts Industry: the Technical Feasibility of a Billion-ton Annual Supply, DOE/GO-102005-2135, Oak Ridge National Laboratory, Oak Ridge, TN (http://feedstockreview.ornl.gov/pdf/billion ton vision.pdf).
3. Perlack, R.D. and Stokes, B.J. (2011) U.S. Department of Energy. *U.S. Billion-Ton Update*: *Biomass Supply for a Bioenergy and Bioproducts Industry*. R.D. Perlack and B.J. Stokes (Leads), ORNL/TM-2011/224. Oak Ridge National Laboratory, Oak Ridge, TN. 227 p. (http://www1.eere.energy.gov/biomass/pdfs/billion_ton_update.pdf).
4. Dale, V.H., Kline, K.L., Wright, L. *et al.* (2011) Interactions between bioenergy feedstock choices and landscape dynamics and land use. *Ecological Applications*, **21** (4), 1039–1054.
5. US DEO (2006) Breaking the Barriers to Cellulosic Ethanol: a joint research agenda (June, 2006) DOE/SC-0095. US Department of Energy Office of Science and Office of Energy Efficiency and Renewable Energy. http://www.doegenomestolife.org/biofuels/.
6. Elander, R.T., Dale, B.E., Holtzapple, M. *et al.* (2009) Summary of findings from the Biomass Refining Consortium for Applied Fundamentals and Innovation (CAFI): corn stover pretreatment. *Cellulose*, **16**, 649.
7. Zheng, Y., Pan, Z.L., Zhang, R.H. *et al.* (2007) Evaluation of different biomass materials as feedstock for fermentable sugar production. *Applied Biochemistry and Biotechnology*, **137**, 423.
8. Albersheim, P., Darvill, A., Roberts, K., Sederoff, R., and Staehlin, A. (eds) (2010) *Plant Cell Walls*, Garland Science, Taylor & Francis Group, LLC.

9. Somerville, C. (2006) Cellulose synthesis in higher plants. *Annual Review of Cell and Developmental Biology*, **22**, 53–78.
10. Lehtio, J., Sugiyama, J. *et al.* (2003) The binding specificity and affinity determinants of family 1 and family 3 cellulose binding modules. *Proceedings of the National Academy of Sciences of the United States of America*, **100** (2), 484–489.
11. Wiselogel, A., Tyson, S., and Johnson, D. (1996) Biomass feedstock resources and composition, in *Handbook on Bioethanol – Production and Utilization* (ed. C.E. Wyman), Taylor and Francis, Washington, DC, pp. 105–118.
12. Viamajala, S., Donohoe, B.S. *et al.* (2010) Heat and mass transport in processing of lignocellulosic biomass for fuels and chemicals, in *Sustainable Biotechnology* (eds O.V. Singh and S.P. Harvey), Springer, Netherlands, pp. 1–18.
13. Esau, K. (2006) *Esau's Plant Anatomy: Meristems, Cells, and Tissues of the Plant Body: Their Structure, Function, and Development*, John Wiley & Sons, Inc., Hoboken, New Jersey.
14. Zeng, M.J., Mosier, N.S. *et al.* (2007) Microscopic examination of changes of plant cell structure in corn stover due to hot water pretreatment and enzymatic hydrolysis. *Biotechnology and Bioengineering*, **97** (2), 265–278.
15. Studer, Michael E., DeMartini, Jaclyn D., Davis, Mark F. *et al.* (March 2011) Lignin content in natural *Populus* variants affects sugar release. *PNAS*, **108** (15), 6300–6305.
16. Lyndon, R.F. and Francis, D. (1992) Plant and organ development. *Plant Molecular Biology*, **19** (1), 51–68.
17. Bonawitz, N.B. and Chapple, C. (2010) The genetics of lignin biosynthesis: connecting Genotype to Phenotype. *Annual Review of Genetics*, **44**, 337–363.
18. Donaldson, L.A. (2001) Lignification and lignin topochemistry – an ultrastructural view. *Phytochemistry*, **57** (6), 859–873.
19. Jarvis, M.C. (1992) Self-assembly of plant-cell walls. *Plant Cell and Environment*, **15** (1), 1–5.
20. McCann, M.C., Wells, B. *et al.* (1990) Direct visualization of cross-links in the primary plant-cell wall. *Journal of Cell Science*, **96**, 323–334.
21. Cannon, M.C., Terneus, K. *et al.* (2008) Self-assembly of the plant cell wall requires an extension scaffold. *Proceedings of the National Academy of Sciences of the United States of America*, **105** (6), 2226–2231.
22. Mohnen, D. (2008) Pectin structure and biosynthesis. *Current Opinion in Plant Biology*, **11** (3), 266–277.
23. Brown, R.M. (2004) Cellulose structure and biosynthesis: What is in store for the 21st century? *Journal of Polymer Science Part A-Polymer Chemistry*, **42** (3), 487–495.
24. Carpita, N., Sabularse, D. *et al.* (1979) Determination of the pore-size of cell-walls of living plant-cells. *Science*, **205** (4411), 1144–1147.
25. Vian, B., Roland, J.C. *et al.* (1993) Primary-cell wall texture and its relation to surface expansion. *International Journal of Plant Sciences*, **154** (1), 1–9.
26. Donohoe, B.S., Decker, S.R. *et al.* (2008) Visualizing lignin coalescence and migration through maize cell walls following thermochemical pretreatment. *Biotechnology and Bioengineering*, **101** (5), 913–925.
27. Kristensen, J., Thygesen, L. *et al.* (2008) Cell wall structural changes in wheat straw pretreated for bioethanol production. *Biotechnology for Biofuels*, **1** (1), 5.
28. Kumar, R., Mago, G. *et al.* (2009) Physical and chemical characterizations of corn stover and poplar solids resulting from leading pretreatment technologies. *Bioresource Technology*, **100** (17), 3948–3962.
29. Chundawat, S.P.S., Donohoe, B.S. *et al.* (2011) Multi-scale visualization and characterization of lignocellulosic plant cell wall deconstruction during thermochemical pretreatment. *Energy & Environmental Science*, **4** (3), 973–984.
30. Abuja, P., Schmuck, M. *et al.* (1988) Structural and functional domains of cellobiohydrolase-i from trichoderma-reesei – a small-angle x-ray-scattering study of the intact enzyme and its core. *European Biophysics Journal with Biophysics Letters*, **15** (6), 339–342.
31. Beckham, G.T., Bomble, Y.J., Bayer, E.A. *et al.* (2011) Applications of computational science for understanding enzymatic deconstruction of cellulose. *Current Opinion in Biotechnology*, **22**, 231–238.
32. Parthasarathi, R., Romero, R.A., Redondo, A., and Gnanakaran, S. (2011) Theoretical study of the remarkably diverse linkages in lignin. *The Journal of Physical Chemistry Letters*, **2**, 2660–2666.
33. Durbeej, B. (2003) Spin distribution in dehydrogenated coniferyl alcohol and associated dilignol radicals. *Holzforschung*, **57**, 59–61.

34. Durbeej, B. and Eriksson, L.A. (2003) A density functional theory study of coniferyl alcohol intermonomeric cross linkages in lignin: Three-dimensional structures, stabilities and the thermodynamic control hypothesis. *Holzforschung*, **57**, 150–164.
35. Durbeej, B. and Eriksson, L.A. (2003) Formation of β-O-4 lignin models: A theoretical study. *Holzforschung*, **57**, 466–478.
36. Jarvis, M.W., Daily, J.W., Carstensen, H.-H. *et al.* (2011) Direct detection of products from the pyrolysis of 2-phenethyl phenyl ether. *Journal of Physical Chemistry A*, **115**, 428–438.
37. Beste, A., Buchanan, A.C. III, and Harrison, R.J. (2008) Computational prediction of/β selectivities in the pyrolysis of oxygen-substituted phenethyl phenyl ethers. *Journal of Physical Chemistry A*, **112**, 4982–4988.
38. Kim, S., Chmely, S.C., Nimlos, M.R. *et al.* (2011) Computational study of bond dissociation enthalpies for a large range of native and modified lignins. *Journal of Physical Chemistry Letters*, **2**, 2846–2852.
39. Parthasarathi, R., Romero, R.A., Redondo, A., and Gnanakaran, S. (2011) Theoretical study of the remarkably diverse linkages in lignin. *Journal of Physical Chemistry Letters*, **2**, 2660–2666.
40. Sangha, A.K., Parks, J.M., Standaert, R.F. *et al.* (2012) Radical coupling reactions in lignin synthesis. A density functional theory study. *The Journal of Physical Chemistry. B*, **116** (16), 4760–4768. doi: 10.1021/jp2122449
41. Bergenstråhle, M., Wohlert, J., Himmel, M.E., and Brady, J.W. (2010) Simulation studies of the insolubility of cellulose. *Carbohydrate Research*, **345**, 2060–2066.
42. Matthews, J.F., Skopec, C.E., Mason, P.E. *et al.* (2006) Computer simulation studies of microcrystalline cellulose Iβ. *Carbohydrate Research*, **341**, 138–152.
43. Matthews, J.F., Bergenstråhle, M., Beckham, G.T. *et al.* (2011) High-temperature behavior of Cellulose I. *The Journal of Physical Chemistry. B*, **115**, 2155–2166.
44. Beckham, G.T., Matthews, J.F., Peters, B. *et al.* (2011) Molecular-level origins of biomass recalcitrance: decrystallization free energies for four common cellulose polymorphs. *The Journal of Physical Chemistry. B*, **115**, 4118–4127.
45. Chundawat, S.P.S., Bellesia, G., Uppugundla, N. *et al.* (2011) Restructuring the crystalline cellulose hydrogen bond network enhances its depolymerization rate. *Journal of the American Chemical Society*, **133**, 11163–11174.
46. Besombes, S. and Mazeau, K. (2005) The cellulose/lignin assembly assessed by molecular docking. Part 1: adsorption of a threo guaiacyl β-O-4 dimer onto an Iβ cellulose whisker. *Plant Physiology and Biochemistry*, **43**, 299–308.
47. Schulz, R., Lindner, B., Petridis, L., and Smith, J.C. (2009) Scaling of multimillion-atom biological molecular dynamics simulation on a petascale supercomputer. *Journal of Chemical Theory and Computation*, **5**, 2798–2808.
48. Charlier, L. and Mazeau, K. (2012) Molecular modeling of the structural and dynamical properties of secondary plant cell walls: Influence of lignin chemistry. *The Journal of Physical Chemistry. B*, **116**, 4163–4174.