Deconstruction of Lignocellulosic Biomass to Fuels and Chemicals

Shishir P. S. Chundawat,1,2,* Gregg T. Beckham,3,4,6,7,* Michael E. Himmel,5,8 and Bruce E. Dale1,2

1Great Lakes Bioenergy Research Center, East Lansing, Michigan 48824; email: chundawa@msu.edu
2Department of Chemical Engineering and Materials Science, Michigan State University, East Lansing, Michigan 48824
3National Bioenergy Center, 4National Advanced Biofuels Consortium, and 5Biosciences Center, National Renewable Energy Laboratory, Golden, Colorado 80401; email: gregg.beckham@nrel.gov
6Department of Chemical Engineering, Colorado School of Mines, Golden, Colorado 80401
7Renewable and Sustainable Energy Institute, Boulder, Colorado 80309
8Bioenergy Science Center, Oak Ridge National Laboratory, Oak Ridge, Tennessee 37831

Abstract

Plants represent a vast, renewable resource that is well suited to provide sustainably for humankind’s transportation fuel needs. To produce infrastructure-compatible fuels from biomass, two challenges remain: overcoming plant cell wall recalcitrance to extract sugar and phenolic intermediates, and reduction of oxygenated intermediates to fuel molecules. To compete with fossil-based fuels, two primary routes to deconstruct cell walls are under development, namely biochemical and thermochemical conversion. Here, we focus on overcoming recalcitrance with biochemical conversion, which uses low-severity thermochemical pretreatment followed by enzymatic hydrolysis to produce soluble sugars. Many challenges remain, including understanding how pretreatments affect the physicochemical nature of heterogeneous cell walls; determination of how enzymes deconstruct the cell wall effectively with the aim of designing superior catalysts; and resolution of issues associated with the co-optimization of pretreatment, enzymatic hydrolysis, and fermentation. Here, we highlight some of the scientific challenges and open questions with a particular focus on problems across multiple length scales.

Keywords

thermochemical pretreatment, enzymatic hydrolysis, biofuels, heterogeneous catalysis

Changes may still occur before final publication online and in print.
INTRODUCTION

Global consumption of crude oil and the impacts of climate change induced by greenhouse gas emissions have led to intensive research efforts to develop renewable and sustainable transportation fuels and industrial chemicals (1, 2). However, development of a renewable fuels industry will require significant R&D to minimize risks associated with its implementation (1–3). In the near-term, nonfood, plant biomass such as agricultural residues, switchgrass, and poplar are likely to be the primary feedstock for deconstruction to reactive intermediates (sugars and phenolics) that can be upgraded to fuels.

Relative to petroleum refining, lignocellulosic biomass conversion offers new logistic and scientific challenges that span many spatiotemporal scales (Figure 1). First, because biomass energy content per hectare is low, harvesting and consolidation of biomass is a major economic issue (4). The transportation distance of biomass thus becomes a major limiting factor in the sizes of lignocellulosic biorefineries (3). Once consolidated, biomass conversion is a considerable technical challenge because the cell wall is a heterogeneous solid composed of a carbohydrate fraction tightly interlinked with a complex alkyl-aromatic fraction. The difficulty associated with gaining access to these cell wall polymers for conversion to reactive intermediates is termed biomass recalcitrance (1). Moreover, the carbohydrate and aromatic polymers in plants have higher oxygen contents than crude oil; hence, reduction to higher energy density molecules is a key challenge in producing biofuels that are compatible with the current transportation infrastructure.

The multiple, near-term routes for overcoming biomass recalcitrance (Figure 2) are broadly separated into biochemical and thermochemical conversion methods (1, 5). Thermochemical conversion is typically delineated into two regimes based on the operating temperature of pyrolysis and gasification, which use heat and pressure to convert the biomass to bio-oils and synthesis gas, respectively. The advantages of thermochemical conversion are low residence time and the

![Figure 1](image-url)

**Figure 1**
Deconstruction of plants into fuels and chemicals through a biochemical or thermochemical route is a challenge that spans various spatiotemporal scales. Both macroscale (e.g., environmental impact, harvesting, and biomass consolidation) and microscale (e.g., lignin-carbohydrate complexation, cellulose crystallinity) factors influence production of lignocellulosic biofuels. Adapted from Reference 75. The first two images are courtesy of DOE/NREL.
ability to handle varied feedstocks in a continuous manner; however, the conversion process is not selective. Biochemical conversion, alternatively, offers high selectivity in deconstructing biomass to desired end products. Biochemical conversion first uses low-severity thermochemical treatment (pretreatment) at temperatures between 100 and 200°C to partially break down the cell wall and improve enzymatic accessibility. Many options exist for pretreatment of biomass (6, 7); the leading examples use liquid catalysts such as sulfuric acid, ammonia, or water, which penetrate the cell wall and alter its chemistry and ultrastructure. Elucidating the physicochemical effects of the many possible pretreatments upon subsequent hydrolysis and fermentation has proven to be a considerable challenge.

The pretreatment step is followed by application of enzymatic or microbial catalysts to convert the carbohydrates to soluble sugars, which are then converted to fuels. Because the plant cell wall is a solid, composite material, enzymes must work directly at the solid-liquid interface and engage in surface depolymerization of individual cellulose chains to hydrolyze carbohydrate polymers. This surface ablation process results in a reaction rate several orders of magnitude slower than freely diffusing enzymatic reactions because conversion is limited by substrate accessibility (8).

The combined steps of pretreatment and enzymatic hydrolysis are responsible for overcoming biomass recalcitrance during biochemical conversion. Reduction of the carbohydrate streams to fuels, which is the last step, is an area of active research spanning cofermentation of pentose and hexose sugars to ethanol (1), metabolic pathway engineering for production of higher alcohols and hydrocarbons (9, 10), and application of catalytic routes to fuels (11, 12). We stress that many of the reduction options available rely on the economic production of monomeric carbohydrates (or other reactive intermediates) derived from biomass. To summarize, overcoming biomass recalcitrance to produce sugars is the crucial first step that supports many downstream biorefinery options for fuel production from ethanol to hydrocarbons.

In this review, we focus on the technical challenges associated with overcoming biomass recalcitrance selectively across multiple length and time scales. First, we describe models of the plant cell wall and highlight unanswered questions associated with the ultrastructural organization of cell wall polymers that impacts its recalcitrance. Options for pretreatment are discussed in light of the current state of knowledge and the effect of pretreatment on the overall conversion process. Second, we review experimental and theoretical efforts to describe the steps that processive enzymes, the workhorses of industrial cocktails, undergo to deconstruct recalcitrant crystalline
cellulose to glucose. We also discuss the role of other proteins (e.g., hemicellulases and other helper proteins) and protein engineering strategies to enhance cellulase performance. Third, we briefly highlight reduction processes such as fermentation and possibilities for process consolidation. We end with our outlook for the future of biochemical conversion of lignocellulosic biomass.

**UNDERSTANDING HOW CHEMICAL PRETREATMENTS AFFECT PLANT CELL WALLS**

The first step in biochemical conversion of lignocellulosics is thermochemical pretreatment to enhance the rate of subsequent enzymatic and microbial catalysis. However, it is important to understand the underlying composition and architecture of cell walls to appreciate the physico-chemical impacts of pretreatment that result in a reduction of native cell wall recalcitrance.

**Compositional and Ultrastructural Organization of Cell Walls**

Plants are composed of at least 35 different cell types that are distinct in composition, structure, and ultrastructure (13). However, all cells have a thick (0.1 to 10 μm) cell wall that provides rigidity to the cell and prevents attack by pathogens. Cell walls typically are composed of three layers, the middle lamella, primary cell wall, and secondary cell wall. Secondary cell walls, which have further sublayers (S1, outer; S2, middle; and S3, inner), are present only in certain tissues (e.g., thickened cells that constitute the vascular bundles) and mature generally after cessation of growth, unlike the primary walls that are ubiquitous to all cells. Cellulose (20–50% on a dry weight basis), hemicellulose (15–35%), and lignin (10–30%) are the primary constituents of cell walls, whereas proteins (3–10%), lipids (1–5%), soluble sugars (1–10%), and minerals (5–10%) are minor components (14). This chemical composition of cell walls differs significantly between monocots (e.g., corn stover, switchgrass) and dicots (e.g., *Arabidopsis*, poplar), which ultimately influences their susceptibility to deconstruction (see Supplemental Information 1 for more information on differences between monocots and dicots; follow the Supplemental Material link from the Annual Reviews home page at http://www.annualreviews.org).

Cellulose is a complex macromolecule composed of linear β-1,4-glucan chains that tightly aggregate into microfibrils (3 to 5 nm in diameter) held together via strong intra- and intermolecular hydrogen bonds and van der Waals forces resulting from pyranose ring stacking. The degree of polymerization of cellulose varies, depending on its source, between 100 and 10,000 (15, 16). Native cellulose is degraded to a length of approximately 150 nm fairly rapidly, beyond which severe chemical or enzymatic treatment is necessary to hydrolyze it completely (17). The steric hindrance of glucan chains packed tightly in this solid, crystalline morphology is responsible for the low saccharification rate of cellulose (18). The most abundant crystalline polymorph found in higher plants is cellulose Iβ, which has a two-chain monoclinic unit cell (19). Thermochemical treatments can transform cellulose Iβ into other polymorphs (15), namely, cellulose II by NaOH (20), cellulose III by amines or ammonia (21, 22), and cellulose IV by glycerol (23). Differences in glucan chain packing in these polymorphs have been shown to influence their hydrolysis rates (20, 22, 24); however, a molecular-level explanation of the observed differences in digestion rates between cellulose polymorphs remains an open question.

Hemicelluloses are polysaccharides that are extractible by alkaline solutions. In contrast, pectins are a major component of the compound middle lamella (see Supplemental Information 1 for more information on the impact of pectins and ferulates on cell wall recalcitrance) that can be extracted with hot water and chelating agents (25). We now know that these complex heteropolysaccharides can be classified into four structurally distinct classes: (a) xylans (β-1,4-xyllosyl

6.4 Chundawat et al.
backbone with arabinose, uronic acid, and acetyl side chains), (b) mannans (β-1,4-mannosyl or glucosyl-mannosyl backbones with galactose side chains), (c) β-glucans with mixed linkages (β-1,3-1,4-glucosyl backbone), and (d) xylglucans (β-1,4-glucosyl backbone with xylose side chains) (26). Unlike cellulose, hemicellulose composition varies depending on cell tissue and plant species and differs in type of glycosidic linkages, side chain composition, and degree of polymerization (27, 28). The most abundant hemicelluloses found in monocots (e.g., corn stover, switchgrass) and dicots are glucuronoroarabinoxylans and galactoglucomannans, respectively.

Lignins are complex, phenyl-propanoid polymers derived from three basic monomeric units (monolignols): p-hydroxyphenyls (H), guaicyls (G), and syringyls (S), which vary between species and cell tissue type (29, 30). Lignin structures are hypothesized to arise from free-radical polymerization of phenoxy radicals (β-O-4-linked aryl ether linkages are most common) formed by oxidative enzymes in the cell wall (30). Hardwood lignins are predominantly G and S monolignols with trace amounts of H units. Softwood lignins are composed of mostly G units, whereas monocots incorporate equivalent amounts of G and S units along with significantly higher amounts of H monolignols.

The self-assembly and architectural organization of cell walls is an area of intense research (16, 31, 32). However, most research has traditionally focused on the primary cell walls, resulting in far less understanding of secondary cell walls (16, 33, 34), which constitute at least 70% to 80% of the stem internode mass (1, 24, 35). Secondary cell walls are also significantly more recalcitrant than primary cell walls to biological deconstruction (1, 36). The current models for cell walls envision cellulose microfibrils (composed of 30 to 36 hydrogen-bonded glucan chains) surrounded by a matrix of hemicellulose and lignin (Figure 3; 16, 34, 37, 38). The cellulose microfibrils in primary cell walls are organized in successive lamellae, forming a web-like matrix, that are separated by hemicellulose and pectins that control the overall wall porosity (<10-nm pore size) (38, 39). However, microfibrils in secondary cell walls are more closely associated with each other to form macrofibrillar lamellae that are oriented in a direction depending on their location within various secondary wall sublayers (34, 36). Unbranched hemicellulose (xylglucans, homoxyans, and mannans) forms hydrogen bonds with the surface of cellulose fibrils, whereas the side chains of branched hemicelluloses (e.g., uronic acids and arabinose) are covalently bonded to hemicellulose or lignin to create enzyme-impenetrable cross-links, also known as lignin carbohydrate complexes (LCCs). The majority of LCC linkages in monocots are ester linkages between hemicellulose side chains and phenolic acids (e.g., ferulate and diferulate) that constitute a portion of the noncore lignin (28, 40, 41). LCCs are thought to form inclusion complexes that exclude water and prevent chemical or enzyme-catalyzed deconstruction of cell walls, the mechanism for which is poorly understood (42).

Consortium for Applied Fundamentals and Innovation–Based Thermochemical Pretreatments

Before the discovery of Trichoderma reesei cellulases, concentrated acids were used to hydrolyze lignocellulose to fermentable sugars directly, which typically resulted in poor yields and extensive sugar degradation (43). However, today the availability of aggressive enzyme preparations permits the use of lower severity acidic pretreatments. The primary goal of any pretreatment used currently is to overcome the lignin-hemicellulose barrier to increase enzyme accessibility (44). Some pretreatments can also alter cellulose crystallinity to enhance its depolymerization rate. Pretreatments can be classified into four categories (i.e., physical, chemical, biological, and solvent-fractionation) and have been reviewed extensively recently (6, 7, 45). Despite efforts to develop novel pretreatments and optimize pretreatment conditions to maximize biomass digestibility, there has been a
Figure 3
A generic ultrastructural model for native and pretreated monocot grass-based secondary cell walls. Adapted from References 33, 75.

Key

Lignin–carbohydrate complex (LCC):

- Arabinoxylan
- Non-core lignin
- Common ester linkages cleaved during thermochemical pretreatments

Key elements:

- Homoxylan
- Glucan
- Core lignin

Figure 3
A generic ultrastructural model for native and pretreated monocot grass-based secondary cell walls. Adapted from References 33, 75.
lack of mechanistic understanding that integrates the molecular (nanometer) scale (e.g., kinetics and energetics) into the cellular/tissue (micrometer) scale (e.g., lignin and hemicellulose extraction and redeposition on outer cell wall surfaces) effects of pretreatments. Until recently, holistic assessment of pretreatments and their influence on upstream and downstream biorefinery processes using standardized methods was lacking (7). In 2000, several laboratories established the Consortium for Applied Fundamentals and Innovation (CAFI) to standardize protocols and conduct holistic assessment of pretreatments (46–48). Some of the leading pretreatments that have been studied (using corn stover, poplar, and switchgrass as feedstocks) as part of CAFI include dilute acid treatment (47, 49), steam explosion (50), hot water treatment (51, 52), ammonia fiber expansion (AFEX) (53, 54), ammonia recycle percolation (55, 56), and lime treatment (57). A detailed overview of CAFI pretreatments is provided in Figure 4.

Novel Pretreatment Approaches

Ionic liquids (ILs) were first reported to disrupt cellulose crystallinity in 2002 (58). Since this report, there have been attempts to fractionate lignocellulose and decrystallize cellulose using ILs, with some success (59–61). Both the anions and cations of the IL are thought to participate in cell wall and cellulose solubilization, with the former playing a more dominant role (62, 63). These results suggest that it should be possible to design more effective ILs through a better mechanistic understanding of polysaccharide-IL interactions (63). Currently research on recovery of ILs and isolation of the dissolved lignin-hemicellulose after pretreatment is lacking (60, 64). Native cellulases are severely inhibited by trace amounts of residual ILs, which has led to development of IL-tolerant enzymes (65). Zhang et al. (66) reported a phosphoric acid-acetone-water–based pretreatment (cellulose solvent and organic solvent lignocellulose fractionation or COSLIF) that can be conducted at mild conditions (50°C, 1 atm, 30 to 60 min). COSLIF fractionates lignin, hemicellulose, and acetic acid while decrystallizing cellulose to result in enhanced enzymatic digestibility compared with dilute acid–pretreated substrates (66, 67). With further process improvements related to reduced loading of expensive solvents and unit operations needed for recovery, the COSLIF and IL-based pretreatment methods may become economically more viable.

Pretreatments Alter Physicochemical Properties of Cell Walls

The primary physicochemical effects of chemical treatments (Figures 3 and 4) that result in enhanced cell wall digestibility can be classified into three categories as follows: (a) LCC cleavage and hemicellulose removal, (b) lignin modification and redistribution, and (c) cellulose decrystallization.

Cleavage of LCCs facilitates the extraction and removal of cell wall polymers (e.g., hemicellulose and lignin), which in turn increases enzyme accessibility to the intact carbohydrates (24, 68). One of the common LCCs includes ester linkages between arabinose and ferulic acid (28, 29, 41, 69, 70). However, to date there have been no detailed studies to determine the cleavage rates of ferulate and diferulate linkages during pretreatment. Model compound studies using experimental and computational approaches would elucidate the mechanisms and kinetics for these reactions. Ongoing work has revealed that the rate of ammonolysis and hydrolysis of diferulate ester linkages during AFEX depends on reaction conditions and type of diferulate linkage (S.P.S. Chundawat, R. Vismeh, A.D. Jones, & J. Ralph, unpublished data). Cleavage of diferulates (which cross-link lignin to polysaccharides) during AFEX facilitated removal of lignin/hemicelluloses and hence increased enzymatic accessibility (24, 53). Determining the susceptibility of various
### Pretreatment category

<table>
<thead>
<tr>
<th>Pretreatment category</th>
<th>Temperature (°C)</th>
<th>Reaction time (min)</th>
<th>Pretreatment chemical</th>
<th>Catalyst loading (g/g BM)</th>
<th>Water loading (g/g BM)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dilute acid</td>
<td>160–220</td>
<td>1–30</td>
<td>H₂SO₄</td>
<td>0.01–0.02 g acid/g BM</td>
<td>3–5</td>
</tr>
<tr>
<td>Steam explosion</td>
<td>180–290</td>
<td>1–15</td>
<td>None or SO₂</td>
<td>0.03 g acid/g BM</td>
<td>4</td>
</tr>
<tr>
<td>Hot water</td>
<td>160–230</td>
<td>10–30</td>
<td>None</td>
<td>None</td>
<td>5–6</td>
</tr>
<tr>
<td>Lime</td>
<td>25–160</td>
<td>120 min–weeks</td>
<td>CaO (w/wo O₂)</td>
<td>0.07–0.2 g CaO/g BM</td>
<td>2–10</td>
</tr>
<tr>
<td>ARP</td>
<td>160–180</td>
<td>5–45</td>
<td>NH₄OH</td>
<td>0.5 g NH₃OH/ g BM</td>
<td>2–3</td>
</tr>
<tr>
<td>AFEX</td>
<td>40–180</td>
<td>5–45</td>
<td>NH₃ or NH₄OH</td>
<td>0.5–1 g NH₃/ g BM</td>
<td>0–1</td>
</tr>
</tbody>
</table>

### Changes to Glucan and Lignin Conversion

<table>
<thead>
<tr>
<th>Pretreatment category</th>
<th>% crystallinity</th>
<th>% residual cellulose</th>
<th>% residual hemicellulose</th>
<th>% residual lignin</th>
<th>% drop in cellulose DP</th>
<th>% acetyl deesterification</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dilute acid</td>
<td>+</td>
<td>85–95</td>
<td>5–25</td>
<td>80–90</td>
<td>60–85</td>
<td>50–95</td>
</tr>
<tr>
<td>Steam explosion</td>
<td>+</td>
<td>95–99</td>
<td>5–60</td>
<td>50–60</td>
<td>60–80</td>
<td>50–85</td>
</tr>
<tr>
<td>Hot water</td>
<td>+</td>
<td>90–99</td>
<td>45–60</td>
<td>NA</td>
<td>20–50</td>
<td>55–75</td>
</tr>
<tr>
<td>Lime</td>
<td>+</td>
<td>97–99</td>
<td>65–97</td>
<td>40–50</td>
<td>50–60</td>
<td>90–95</td>
</tr>
<tr>
<td>ARP</td>
<td>–</td>
<td>90–99</td>
<td>40–70</td>
<td>15–60</td>
<td>10–30</td>
<td>85–90</td>
</tr>
<tr>
<td>AFEX</td>
<td>–</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>5–20</td>
<td>80–95</td>
</tr>
</tbody>
</table>

### Washing (W), Detoxification (D), and Nutrient Addition (N)

<table>
<thead>
<tr>
<th>Pretreatment category</th>
<th>%glucan conversion</th>
<th>%xylan conversion</th>
<th>Washing (W), detoxification (D), and nutrient addition (N)</th>
<th>MESP ($/gal ethanol)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dilute acid</td>
<td>92</td>
<td>93</td>
<td>W, D, N</td>
<td>1.35</td>
</tr>
<tr>
<td>Steam explosion</td>
<td>NA</td>
<td>NA</td>
<td>W, D, N</td>
<td>NA</td>
</tr>
<tr>
<td>Hot water</td>
<td>91</td>
<td>81</td>
<td>D, N</td>
<td>1.65</td>
</tr>
<tr>
<td>Lime</td>
<td>94</td>
<td>76</td>
<td>W</td>
<td>1.6</td>
</tr>
<tr>
<td>ARP</td>
<td>90</td>
<td>88</td>
<td>W</td>
<td>1.65</td>
</tr>
<tr>
<td>AFEX</td>
<td>96</td>
<td>91</td>
<td>None</td>
<td>1.4</td>
</tr>
</tbody>
</table>

### Degradation Products

<table>
<thead>
<tr>
<th>Degradation products (μg analyte/g substrate)</th>
<th>Untreated corn stover</th>
<th>AFEX treated</th>
<th>Dilute-acid treated</th>
</tr>
</thead>
<tbody>
<tr>
<td>Acetic acid</td>
<td>1,610</td>
<td>4,610</td>
<td>34,770</td>
</tr>
<tr>
<td>Levulinic acid</td>
<td>171</td>
<td>24</td>
<td>3,649</td>
</tr>
<tr>
<td>Furfural/HMF</td>
<td>72</td>
<td>645</td>
<td>23,640</td>
</tr>
<tr>
<td>Acetamide/phenolic amides</td>
<td>–</td>
<td>39,801</td>
<td>–</td>
</tr>
<tr>
<td>Pyrazine/imidazole derivatives</td>
<td>–</td>
<td>945</td>
<td>–</td>
</tr>
<tr>
<td>Syringaldehyde</td>
<td>3</td>
<td>11</td>
<td>149</td>
</tr>
<tr>
<td>Phenolic acids</td>
<td>196</td>
<td>1,183</td>
<td>3,151</td>
</tr>
</tbody>
</table>

Changes may still occur before final publication online and in print.
LCCs during pretreatment will prove critical to engineering bioenergy crops with reduced cell wall recalcitrance.

Unlike AFEX, acidic pretreatments achieve near-complete solubilization of hemicellulose to sugars (47, 71). The kinetics of hemicellulose hydrolysis during acidic pretreatments is biphasic, with the faster hydrolysis regime following first-order reaction kinetics (72). This biphasic behavior could be due to limited accessibility of hemicelluloses sheathed by hydrophobic lignin, but this theory has not been explained conclusively. Brunecky et al. (73) showed that xylan accumulates around the cell lumen and middle lamella after dilute acid treatment; however, the actual mechanism for in situ hemicellulose delocalization during acidic pretreatment and its effect on subsequent hydrolysis is unclear. Most chemical pretreatments result in significant xylan deacetylation that yields improved xylan hydrolysis (Figure 4b), but the impact of deacetylation on increased cellulose-xylan association is unknown. This phenomenon of increased association between deacetylated hemicellulose and cellulose is likely more important for lower severity pretreatments such as AFEX that do not solubilize the hemicellulose into a separate liquid stream. Use of ILs results in the rapid separation of the primary and secondary switchgrass cell walls from the middle lamella followed by complete solubilization (60). However, addition of water resulted in the precipitation of low-crystallinity cellulose II along with rejection of lignin and oligomeric hemicellulose into the supernatant (60, 74). Similar multifaceted characterization and visualization studies conducted for AFEX have revealed that the middle lamella and outer secondary grass cell walls are the most prone to disruption during pretreatment (24, 53, 75). Characterization of the relative recalcitrance of distinct cell wall regions (e.g., S1 versus S3 secondary wall) to pretreatment and hydrolysis would be of interest to efforts to engineer improved bioenergy crops.

The impact of pretreatment on lignin composition and redistribution within cell walls has been explored only recently (24, 44, 76, 77). Donohoe et al. (76, 78) used electron tomography to show that the droplets that appear following dilute acid treatment are enriched in lignin that extrudes out of cell walls at temperatures close to lignin’s glass transition temperature. These droplets have been shown to inhibit cellulase activity (79), but the mechanism of cellulase-lignin interaction is unclear. The exact composition of delocalized/coalesced lignin within different cell wall compartments and its impact on saccharification is also unknown. With the development of whole cell wall nuclear magnetic resonance (NMR) characterization techniques, it is now possible to analyze cell walls without modifying their composition during sample preparation (77, 80). Unlike other pretreatments, AFEX has been shown to alter subtly the distribution of lignin and hemicellulose via extraction/redeposition onto outer wall surfaces, without altering core lignin chemistry, to create an enzyme-porous cell wall (S.P.S. Chundawat, B.S. Donohoe, F. Lu, J. Ralph, unpublished data; 24, 75). Both dilute acid and AFEX pretreatment were found to significantly alter the ultrastructure of the compound middle lamella and the outer secondary cell walls of corn stover (24, 75, 76). These results suggest that mass transfer considerations for lignin and hemicellulose removal from cell walls are a major barrier to effective cell wall

Figure 4
Overview of Consortium for Applied Fundamentals and Innovation (CAFI) pretreatments using corn stover as feedstock: (a) range of pretreatment conditions employed; (b) major physicochemical impacts of pretreatment; (c) relative enzymatic digestibility, ease of fermentability and minimum ethanol selling price (MESP) for various pretreated substrates; and (d) major cell wall decomposition products formed during ammonia fiber expansion (AFEX) and dilute acid pretreatment. Adapted from References 47, 48, 50, 53, 71, 145, 146, 149. BM, biomass; ARP, ammonia recycle percolation; DP, degree of polymerization; HMF, hydroxymethylfurfural; NA, not available. Note: + or – signs indicate relative increase or decrease, respectively, in cellulose crystallinity with respect to untreated control.

www.annualreviews.org • Biochemical Conversion of Biomass 6.9
deconstruction; however, much remains to be learned about these processes. There have been no reports on real-time morphological (e.g., tissue disruption), ultrastructural (e.g., cellulose microfibril alteration), and chemical (e.g., LCC cleavage) changes occurring in cell walls during pretreatment that would enable us to obtain a multiscale understanding of the system. Haas et al. (81) have shown, using real-time microscopic imaging, that the structural-level complexity of unmilled plant cell walls impedes heat and mass transfer during thermochemical conversion to result in undesirable tar formation. Analogous real-time imaging and characterization studies for in situ cell wall pretreatment and enzymatic hydrolysis are critical to understanding and overcoming biomass recalcitrance.

Most acidic and oxidative pretreatments result in a marginal increase in cellulose crystallinity and a reduction in its degree of polymerization (Figure 4b; 44). However, cellulose crystallinity measurements by X-ray diffraction of biomass are confounded by the presence of lignin and hemicellulose (82). Some pretreatments, such as concentrated acids (e.g., 85% phosphoric acid) (66), dimethyl sulfoxide (DMSO) (83), transition-metal complex solutions (e.g., Cadoxen) (84), and ILs (63, 85), can completely solubilize cellulose, which upon precipitation with antisolvents results in the formation of amorphous cellulose and cellulose II. Treatment of crystalline cellulose with anhydrous liquid ammonia also results in the formation of cellulose III$_1$ without producing significant amounts of amorphous cellulose, which has demonstrated a four- to fivefold higher rate of saccharification than cellulose I$_\beta$ (22, 24). Ongoing efforts aim to adapt conventional AFEX to produce cellulose III$_1$ and extract lignin simultaneously during pretreatment, as it would be significantly cheaper to recycle ammonia than to utilize the conventional chemicals used to produce amorphous cellulose (S.P.S. Chundawat, V. Balan, L. Sousa, A. Cheh, B.E. Dale, unpublished data). However, we lack a mechanistic understanding of the improved deconstruction kinetics of crystalline cellulose III$_1$ versus I$_\beta$.

Most thermochemical pretreatments result in the formation of decomposition products owing to degradation of carbohydrates and lignin; these products may inhibit downstream biological processing (Figure 4d; 53, 86). Recently, we conducted a detailed mass balance for more than 75 degradation products (e.g., amides, furans, imidazoles, phenolics) formed and/or released from corn stover cell walls during AFEX and dilute acid pretreatment (53). Identification of degradation products formed during pretreatment and elucidation of their inhibitory/stimulatory effect on enzymes and microbes is crucial to optimizing pretreatments and minimizing the impact of pretreatment-induced recalcitrance on biomass deconstruction.

UNDERSTANDING AND IMPROVING THE ENZYMATIC HYDROLYSIS OF BIOMASS

The second step in biochemical conversion of biomass is enzymatic hydrolysis to depolymerize the intact carbohydrate polymers to soluble sugars. These synergistic enzyme cocktails include exoglucanases (processive enzymes), endoglucanases (nonprocessive enzymes), and $\beta$-glucosidases (cellobiases) for depolymerizing cellulose as well as several classes of hemicellulases and accessory enzymes for depolymerizing hemicelluloses (87). Fungi and most bacteria utilize noncomplexed, secreted enzymes, whereas some bacteria tether their enzymes to large scaffolds in protein complexes termed cellulosomes (Figure 5; 88). Many authors have reviewed the differences between free and complexed enzyme systems (87, 88). Here, we focus on free fungal enzymes, as these have received significant attention recently (1, 89–92). Fungi are of interest commercially because they can secrete proteins to titers more than 100 g liter$^{-1}$ and their enzyme cocktails are naturally quite effective at biomass deconstruction (89).
Structural overview of the cellulosic substrate and biological catalysts responsible for its deconstruction. (a) Native (I) and synthetic (II, III) cellulose polymorph crystals. (b) Noncomplexed *Trichoderma reesei* exocellulase (Cel7A, Cel6A) and endocellulase (Cel7B) catalytic domains. (c) The *Clostridium thermocellum* exocellulase CelS catalytic domain as part of the large cellulosomal complex inclusive of CbhA and Cel5B (shown here as a coarse-grained model with the enzymes in different colors). Note: All four catalytic domains (Cel6A, Cel7A, Cel7B, and CelS) are shown with the cellodextrin residue (orange) bound within their respective enzyme active sites.

**The Development of the Cellulase System from *Trichoderma reesei***

During World War II, the U.S. military faced problems with canvas accoutrements rotting in the tropics and subsequently deployed scientists at the U.S. Army Natick Laboratory to study the biological agents responsible for this decay (93). A major outcome of this work was the classification of a particularly effective cellulose-degrading fungus, *Trichoderma viride* (now known as *T. reesei*). This fungus has become one of the most thoroughly studied cellulase-producing organisms to date and a cornerstone of modern industrial biotechnology (94). This discovery led to intense research into the mechanisms by which *T. reesei* degrades biomass via a synergistic enzyme cocktail...
(95–115). The *T. reesei* enzyme cocktail was shown to contain a reducing-end-specific, processive glycosyl hydrolase (GH) Family 7 (see http://www.cazy.org for classification of carbohydrate-active enzymes) cellobiohydrolase (Cel7A, formerly CBH I) and a counterpart nonreducing-end-specific Family 6 processive cellobiohydrolase (Cel6A, formerly CBH II). Together, these two enzymes comprise the majority (>50%, w/w) of the enzyme cocktail secreted by *T. reesei*. Recently, *T. reesei* was reclassified as an anamorph of *Hypocrea jecorina*.

**Molecular-Level Understanding of Cellulases**

Because it is the best-characterized cellulase and the major component of fungal enzyme cocktails, here we focus on studies of and insights gained about *T. reesei* Cel7A. Cel7A consists of a small Family 1 carbohydrate-binding module (CBM), an O-glycosylated linker, and a large catalytic domain (CD) containing a 50 Å tunnel for threading cellulose chains and three sites for N-glycosylation (104). A schematic of this enzyme is shown in **Figure 6**. The findings and insights

![Figure 6](image_url)

**Figure 6**

Steps involved in the mechanistic action of the *Trichoderma reesei* exocellulase (Cel7A) on crystalline cellulose. The yellow space-filling representation is O-glycosylation, the dark blue spacefill is N-glycosylation, the light blue schematic view is the Cel7A enzyme, and the green substrate is a cellulose microfibril. (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) formation of a catalytically active complex, (e) hydrolysis (the product is shown in pink spacefill), and (f) product expulsion and threading of another cellobiosyl unit. The catalytically-active complex structure is adapted from Reference 150.
gained for the action of Cel7A likely will extend to other processive cellulases (and chitinases) from fungi and bacteria because of structural and functional similarities between enzyme families. The probable steps involved in Cel7A action on cellulose include binding of the cellulase to biomass, recognition of a free cellulose chain end, initial threading of the chain into the active site tunnel and decrystallization from the substrate, the hydrolysis reaction, product expulsion, and reformation of the catalytically active complex (CAC). This overall process is shown in Figure 6; panels d–f illustrate the processive cycle. Each of these steps is reviewed here in turn.

**Binding to cellulose via the carbohydrate-binding module.** Boraston et al. (116) provided an excellent review of and classification system for many CBM families. For fungal cellulases, most CBMs are the small Family 1 CBMs. An NMR structure was solved for the CBM of *T. reesei* Cel7A, and the sequence homology of CBMs in Family 1 is quite high; thus, most Family 1 CBM structures are inferred via homology modeling to the *T. reesei* Cel7A CBM (108, 114). Many open questions remain regarding CBM interactions with cellulose, which has important consequences for biomass conversion because CBMs are responsible for increasing the catalyst surface concentration of cellulases on cellulose. Experimental studies of the Cel7A CBM interaction with cellulose demonstrated that the CBM prefers the hydrophobic face of cellulose Iα (117). As the hydrophobic faces of cellulose Iβ and Iα are almost identical (19, 118), this observation likely holds for both polymorphs.

The thermodynamic nature of Cel7A CBM binding is unknown. The binding event in a Family 2 bacterial CBM from *Cellulomonas fimi* has been shown with isothermal titration calorimetry (ITC) to be entropically driven and likely occurs via surface dehydration (i.e., the hydrophobic effect) (119). However, because Family 2 CBMs are larger than Family 1 CBMs, the driving force for binding of Family 1 CBMs may be more enthalpic in nature. A recent simulation demonstrated that hydrogen bonds (enthalpic contributions), rather than the typically hypothesized hydrophobic interactions, are important to CBM behavior on the hydrophobic face of cellulose (114).

The relationship between CBM binding affinity and catalysis efficiency is poorly understood. For free cellulases and cellulosomes, a higher binding affinity has been demonstrated to yield a higher cellulose conversion rate (101, 120). This is not unexpected because cellulose conversion by enzymes is a surface reaction. However, the literature commonly compares overall cellulase activity across enzymes with different binding affinities. Given that cellulase action on cellulose is a heterogeneous catalysis process, an essential kinetic parameter is catalyst surface concentration. Using ITC to measure the binding affinities of CBMs and cellulases in general is crucial to compare their intrinsic cellulase activities. We also lack a clear understanding of and differentiation between productive and nonproductive binding of cellulase to its substrates. This partly explains the ambiguity in the literature and highlights the relationship between overall binding and hydrolysis yield (121).

Other significant questions about CBMs relate to their disruption of cellulose crystallinity and their (and other proteins') ability to enhance conversion rates. In general, the literature states that the CBM-cellulose interaction disrupts hydrogen bonds on the cellulose surface, but this has not been demonstrated definitively. There have been reports of cellulose disruption via disruptor proteins (92, 122–124), but only one study to our knowledge has demonstrated cellulase synergy at relevant cellulase loadings (92). Harris et al. (92) showed that adding a (potentially misclassified) Family 61 GH to an industrial *T. reesei* cocktail resulted in pretreated corn stover glucan conversions equivalent to those achieved using twofold higher enzyme loadings in the absence of GH61. Interestingly, the structure of this GH61 enzyme from *Thielavia terrestris* exhibits structural homology to a chitin-binding protein (CBP21) that has a similar synergistic effect on chitin (125). The mechanisms of *T. terrestris* GH61 and *Serratia marcescens* CBP21-catalyzed decomposition of...
cellulose and chitin, respectively, have not yet been determined. In both cases, a metal ion binding site is located near the protein surface, and removal of divalent metal ions during hydrolysis reduces the synergistic effect of these proteins (92, 125). Also, *T. terrestris* GH61 does not enhance conversion of *T. reesei* enzyme cocktails on isolated cellulose, e.g., Avicel; however, it does have this effect on biomass (92). CBP21 enhances conversion rates on chitin alone, which suggests that despite some structural homology, the mechanisms by which these enzymes work are different, or that GH61 acts on more accessible carbohydrates. If these proteins are indeed enzymatic and specifically hydrolytic, then it is odd that the reducing-sugar assays that are commonly used for measuring cellulase or chitinase activity have been unable to sufficiently demonstrate activity. Clearly, further characterization is needed to understand why GH61 enzymes (and other proteins including CBMs) are able to disrupt biomass and what molecular-level mechanisms they use to do so. Searching for other chemistries occurring (for GH61s) or examining at the molecular level the material properties of cellulose upon incubation with advanced surface characterization techniques will likely aid in determining the effect of biomass disruption. See Supplemental Material 1 for more information on the mechanism driving action of CBP21, the influence of the cellulose dipole on enzyme activity, cellulose polymorphs, and the role of lignin on CBM binding and cellulase activity.

**Surface diffusion of Cel7A on cellulose.** Once bound, cellulases diffuse on the surface of cellulose to locate a free chain for deconstruction. Studies of cellulase surface diffusion to date include several computational studies of the Cel7A CBM, an experimental study that established the binding faces on cellulose Iα, and a high-speed atomic force microscopy (HS-AFM) study (114, 115, 117, 126, 127). Two modeling studies predicted that the Cel7A CBM diffuses along a cellulose chain in discrete energy wells every ~1 nm, which corresponds to the cellobiose length. The residues responsible for this critical length scale are conserved across many Family 1 CBMs (114, 115). A recent, exciting study from Igarashi et al. (127) examined the diffusion of Cel7A on the surface of cellulose I with HS-AFM. Three enzymes were studied: the wild-type (WT) Cel7A; a catalytically inactive mutant; and the W40A mutant, which putatively does not thread cellulose (127). The authors showed that the WT Cel7A enzyme moves at 3.5 nm s⁻¹, although determining if the single enzymes under observation were hydrolyzing cellulose or diffusing along the surface without productive binding is not yet possible. That these enzymes were reported to travel only in a single direction along a cellulose fibril is a promising suggestion that they are hydrolyzing cellulose, because nonengaged cellulases should be able to diffuse in any direction on the surface. The catalytically inactive mutant binds to cellulose and does not move on the timescale of observation (on the order of minutes), and the nonthreading mutant (W40A) exhibited similar behavior in that it was observed to bind but not translate. This observation suggests that the catalytically inactive mutant threads a cellulose chain and then neither disengages nor reacts. We discuss the enigmatic W40A mutant results in the next section.

**Recognition of a reducing end of a cellulose chain via the catalytic domain and initial threading of a cellodextrin chain.** Koivula et al. (95) examined cellulose recognition by the CDs of cellulases. With Cel6A, they found that mutating the tryptophan at the entrance to the CD tunnel (W272) reduced its ability to deconstruct crystalline cellulose, but the conversion rate for amorphous cellulose was unaltered. A later study mentioned that the same is true for Cel7A, presumably mutation of W40 at the CD tunnel entrance (96). The recent HS-AFM study by Igarashi et al. (127) examined the W40A mutant, which digested phosphoric acid swollen cellulose (PASC) at a rate equivalent to the WT Cel7A and crystalline cellulose at a rate roughly equivalent to the Cel7A CD alone. However, in the HS-AFM experiments, W40A was observed
to bind but not translate on the surface of cellulose. Furthermore, its binding time was observed to be less than that of the catalytically inactive mutant. This suggests again that this enzyme cannot recognize and thread a significant portion of a crystalline cellulose chain because the aromatic group at the entrance of the tunnel has been removed. The W40A mutant likely does not translate on the surface but does conduct some hydrolysis because it engages as a nonprocessive cellulase. In other words, the binding free energy to the celloextrin chain may not be high enough to stabilize the chain being fed into the CD tunnel, and thus the enzyme can disengage easily from the cellulose surface. This hypothesis can be tested with free energy calculations in which the W40A mutation is made in silico. Measuring the relative flexibility in a celloextrin chain and the free energy change upon mutation will quantify this effect.

The catalytic steps: hydrolysis, product expulsion, and processivity. Once the processive cellulase has recognized a free chain end, it threads the chain into the tunnel to form a CAC. Because cellulose decrystallization in water is free-energetically unfavorable, the tunnels or clefts of cellulase CDs contain hydrophobic and polar residues that form favorable contacts with a cellulose chain (102, 104, 107). A favorable ligand-binding free energy thus allows cellulases to form CACs despite the thermodynamic barrier to removal of a cellulose chain from the crystal. Several studies have mutated hydrophobic residues in the CD tunnels of cellulases and chitinases (chitinases are 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 (95, 127–130). Additionally, Horn et al. (129) and Vuong & Wilson (128) both 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 celloextrin chain, the hydrolysis reaction occurs usually via a retaining (Cel7A) or inverting (Cel6A) 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. The mechanism for these steps is unknown, although they are under intense investigation via simulation and experimental approaches. At the nanometer scale, it is likely that the cellulose polymorph (e.g., IIIβ versus Iα) and chain location in the crystal determines the work that a processive cellulase must do to decrystallize a cellobiose unit from the polymer crystal. As cellulose is insoluble, the ligand-binding free energy in a cellulase tunnel or cleft must be favorable to extract and process a cellulose chain from the crystal to the enzyme. For the reaction mechanism in Cel6A, Koivula et al. (97) used experimental and theoretical techniques to ascertain the catalytic residues and then confirmed the catalytic acid site originally hypothesized from the reported crystal structure (102). A recent study from Barnett et al. (98) probed the conformation of the pyranose ring in the active site of T. reesei Cel7A. Using density functional theory free energy calculations along a pucker coordinate, they elucidated the stabilized conformations in the CAC. As hydrolysis is likely the rate-limiting step overall, significant challenges still remain in elucidating the formation of the CAC and the elementary steps in the chemical reaction. Altering the crystalline structure of cellulose is thought to help overcome the rate-limiting step of enzymatic hydrolysis, but this theory has yet to be proven conclusively. A hybrid quantum-mechanics/molecular-mechanics (QM/MM) approach likely will be necessary to elucidate the reaction mechanism (131). For product expulsion, absolute ligand-binding free energy calculations (132) can measure the ligand-binding free energy of cellulases, which has significant relevance to product inhibition and the design of cellulase cocktails (133). Furthermore, several sets of simulations can help quantify the threading of a cellobiose unit including rare event simulations of threading, both in the presence of a cellulose crystal and on a cellulose chain in solution (134).
Role of the linker in catalysis. The role of the linker and CBM in catalysis has yet to be elucidated definitively. Until recently, Srisodsuk et al. (110) had conducted the only biochemical study of the \textit{T. reesei} Cel7A linker. The authors denoted the linker region nearest the CD as the flexible region and the region closest to the CBM as the stiff region because it contains significant O-glycosylation. They showed that removal of the flexible region does not change the activity, but the binding affinity was lowered at higher loadings. The authors were not able to explain their observations in terms of their original assessment of the stiff and flexible regions of the linker. We recently applied replica-exchange molecular dynamics simulations of the Cel7A linker domain with the experimentally determined glycosylation pattern (105) to illustrate that the linker is a disordered, flexible tether between the CBM and CD (135). Thus, in the limit of significant linker flexibility (135), these results (110) can be explained by reduction of the surface concentration of cellulases because of molecular crowding via a shortened, but still flexible, tether between the CBM and CD. The Igarashi et al. (127) study revealed the speed of the CD and intact Cel7A translating on the cellulose surface to be similar. Together, these results lend support to the linker functioning primarily as a flexible molecular tether between the CBM and CD.

Kinetic Models of Cellulase Action

Many kinetic models of cellulose digestion via enzyme cocktails have been developed (18, 99, 136–139). A crucial limitation of the models developed to date is that we still lack a comprehensive understanding of how a given cellulase cocktail, much less a single cellulase enzyme, works mechanistically. Thus we stress the importance of understanding the thermodynamics and kinetics of each elementary step of cellulase action in order for models to be predictive outside of the ranges in which parameters are fit. As cellulases undergo multiple steps to deconstruct cellulose, a systematic approach to probe each of these elementary steps with thermodynamic, kinetic, and biochemical measurements and molecular simulations will yield the insights necessary to improve models of cellulose deconstruction by cellulase cocktails.

Recently, Levine et al. (99) published an interesting cellulase kinetic model that highlights the incomplete molecular picture of cellulase action. The authors treated the adsorption to cellulose explicitly and separately from the formation of a CAC, a procedure that essentially accounts for cellulase surface diffusion as the induction time between the adsorption and CAC formation. This treatment is an improvement on previous models. The leveling off of the conversion rate as a function of time could be corrected by reducing the half-lives of Cel7A and the modeled endoglucanase II (EG-II) by an order of magnitude and by increasing the product inhibition constant significantly. The authors did not include a deactivation term in their model, which potentially could have made their model agree quantitatively with the available experimental data without the need to modify two parameters that experimental data indicate are different. Jalak & Väljamäe (140) treated the concept of deactivation in rate expressions for \textit{T. reesei} Cel7A acting on cellulose. They developed a technique to measure an observed rate constant by measuring the concentration of processive cellulases with a cellulose chain in the tunnel via a small molecule inhibitor. This technique was applied over a range of model and real substrates and over a large enzyme loading range. They found a rapid decrease in the observed rate as a function of time on all substrates. From these results, the authors developed a model in which the slow kinetics arises from deactivation of processive cellulases upon reaching steric obstacles along a cellulose chain. Thus the rate of reactivation of a given cellulase that has become immobilized is related, as discussed above, to the ligand-binding free energy, which is emerging as a key variable in cellulase engineering.
Outlook for Cellulase Improvements

There is significant financial impetus for improving cellulase activity for biomass conversion processes, and thus several groups have examined strategies for improving the activity of enzyme cocktails (89, 91, 141). Several strategies exist, categorized as follows:

- directed evolution;
- rational design and engineering; and
- addition of hemicellulases, accessory enzymes, and other helper proteins.

For activity improvements, directed evolution is difficult to apply to cellulase activity because the sequence space is vast, the expression systems are complex because fungal cellulases often do not express well in other hosts, and the screens are confounded by the need to measure specific activity on crystalline cellulose as part of an enzyme cocktail (89). However, improving the thermal stability of fungal cellulases via high-throughput screening is the most popular strategy to improve cellulase performance. Heinzelman et al. (91, 100) used a computational screening approach to recombine segments of a Family 6 cellulase from several WT sequences into a new Family 6 cellulase with an improvement in \( T_m \) of 7 to 15°C. The modified Cel6A enzymes had superior activity on amorphous cellulose; however, the engineered enzymes were not tested on crystalline cellulose.

This approach to producing Cel6A mutants is interesting and will likely find application to the entire cocktail. Lantz et al. (90) recently used a high-throughput engineering approach to improve the thermal tolerance of \( T. reesei \) Cel7A by 14°C and Cel6A by 7°C. They screened enzymes on pretreated corn stover in an enzyme cocktail at process-relevant conditions. Future computational approaches may utilize, for example, Rosetta for in silico screening of more thermostable mutants before producing them experimentally (142).

To date, rational engineering of cellulases has yet to provide a significant specific activity improvement in fungal enzymes. Minor successes have come from improving the binding affinity of the CBM (101), but otherwise few processive and nonprocessive cellulases have been improved via a rational approach (89).

Another issue in cellulase improvement is glycosylation. We have shown that changes to glycosylation owing to expression in nonnative hosts usually leads to activity reductions (103, 111). Because nonnative hosts are often warranted for ease of purification for enzyme studies (91), particular attention should be paid to the role of O- and N-glycosylation of cellulases. Protease resistance and secretion have been attributed to cellulase glycosylation, and the glycans may be trimmed back after secretion, but a systematic study to probe this has not yet been conducted. Additionally, large glycosylated portions of a given CD or linker on a cellulase may interfere with or promote interactions with cellulose during hydrolysis, but again, this has not yet been studied.

With more advanced techniques to quantify the extents and chemistries of glycosylation under development (143), we anticipate that characterization of cellulase glycosylation will enable the use of industrial expression hosts designed to secrete enzymes with the optimal amount of glycosylation for protein stability and specific activity. See Supplemental Information 1 for more information on the role of hemicellulases and other accessory enzymes in improving cellulolytic activity.

CONVERSION OF SUGARS TO FUELS

Pretreatment and enzymatic hydrolysis offer near-term routes to overcoming recalcitrance. The product streams from these two steps are C5 and C6 sugars and lignin. The second challenge in biomass conversion to fuels is then conducting reduction chemistry to remove oxygen from the intermediates to increase the fuel value. Overall, the capability of pretreatment and enzymatic hydrolysis to yield these streams offers the potential to use many reduction chemistry processes in a plug-and-play type fashion. Perhaps the best-developed reduction chemistry option is the
cofermentation of C5 and C6 sugars in yeast/bacteria or the consolidated bioprocessing option in which cellulosomal bacteria, such as *Clostridium thermocellum*, are used to digest cellulose and produce ethanol simultaneously (87, 144). Two of the major challenges for efficiently fermenting the sugar hydrolysates are the presence of small molecule inhibitors (e.g., furans, phenolics) formed during pretreatment and the lack of suitable nutrients to support microbial growth (53, 145). Pretreatments, except for AFEX (146), strip essential nutrients (e.g., proteins, minerals) from the cell wall, which makes it necessary to supplement additional exogenously added nutrients, which negatively impacts process economics. Future research must examine the integration of on-site enzyme production (i.e., using pretreated biomass), pretreatment, hydrolysis, and fermentation to co-optimize the entire process rather than individual unit operations. In addition, coupling ecosystem-inclusive life-cycle analysis to technoeconomic analysis of biorefineries will help assess the true impact of cellulosic biofuels (see Supplemental Information 1 for more information on farm-to-wheel life cycle analysis and the impending food versus fuel dilemma).

Beyond fermentation and CBP, several new processing options exist; these are generally separated into biological and chemical routes. Biological reduction strategies include applications of synthetic biology to engineering nonnative metabolic pathways into bacteria for production of small molecules from sugars or other intermediates from biomass (147). A popular catalytic route from recent years is aqueous phase reforming, in which sugars are dehydrated to hydroxymethylfurfural (HMF) in the aqueous phase and separated via an immiscible organic phase (148). HMF is an attractive intermediate from biomass for the production of alkanes (11). This selective conversion and separation was first demonstrated for fructose, and many research groups are now attempting to demonstrate similar chemistry for glucose. It is anticipated that new routes for removing oxygen from monomeric carbohydrates or small-molecule intermediates will be developed in coming years, as there is great impetus to use continuous catalytic processes to upgrade biomass intermediates to fuels.

**CONCLUSION**

Here, we have outlined the fundamental scientific and engineering challenges associated with overcoming the recalcitrance of plant cell walls to biochemical conversion. Several key challenges limit lignocellulose utilization in the current petrochemical-dominated industry, namely: limited feedstock availability, rudimentary supply-chain logistics, high oxygen-to-carbon content, slow enzyme kinetics for catalysis of insoluble biomass to sugars, and lack of robust microbial catalysts. Many fundamental processes familiar to chemical engineers and physical chemists, such as heat and mass transfer, interfacial physics and chemistry, and catalyst design, are crucial to understanding and improving biomass deconstruction. Development of novel biochemical deconstruction processes along with multiscale, holistic modeling can address the recalcitrance and reduction issues facing lignocellulosic biomass utilization.

**DISCLOSURE STATEMENT**

The authors are not aware of any affiliations, memberships, funding, or financial holdings that might be perceived as affecting the objectivity of this review.

**ACKNOWLEDGMENTS**

S.P.S.C. and B.E.D. thank the DOE Great Lakes Bioenergy Research Center (DOE BER Office of Science DE-FC02-07ER64494) for funding. G.T.B. and M.E.H. thank the DOE Office of...
of the Biomass Program, the DOE Office of Science, Office of the Biological and Environmental Research through the BioEnergy Science Center (BESC), and the DOE Office of Science ASCR SciDAC program for funding. G.T.B. also acknowledges partial funding for this work from the National Advanced Biofuels Consortium (NABC), which is funded by DOE’s Office of the Biomass Program through recovery act funds. We thank Yannick Bomble for the cellulosome figure and other colleagues for fruitful discussions that contributed to this work. The first two images in Figure 1 are courtesy of DOE (http://genomics.energy.gov) and NREL (http://www.nrel.gov).

LITERATURE CITED


132. movements may still occur before final publication online and in print

RELATED RESOURCES

Great Lakes Bioenergy Research Center: http://www.glbrc.org
BioEnergy Science Center: http://www.bioenergycenter.org
Joint Bioenergy Institute: http://www.jbei.org
MSU Biomass Conversion Research Laboratory: http://www.everythingbiomass.org
Carbohydrate-Active Enzymes Database: http://www.cazy.org