JOURNAL OF THE ROYAL SOCIETY Interface

Bioenergy research: a new paradigm in multidisciplinary research

Udaya C. Kalluri and Martin Keller

J. R. Soc. Interface published online 11 June 2010
doi: 10.1098/rsif.2009.0564

JOURNAL OF THE ROYAL SOCIETY Interface

FirstCite® e-publishing

J. R. Soc. Interface
doi:10.1098/rsif.2009.0564
Published online

Review

Bioenergy research: a new paradigm in multidisciplinary research

Udaya C. Kalluri[1,3] and Martin Keller[2,3,*]

[1]*Environmental Sciences Division,* [2]*Biological and Environmental Sciences Directorate, and* [3]*BESC BioEnergy Science Center, Oak Ridge National Laboratory, Oak Ridge, TN, USA*

The field of biology is becoming increasingly interdisciplinary and cross-cutting. This changing research atmosphere is creating the way for a new kind of enquiry that while building upon the traditional research establishment is providing a new multidisciplinary framework to more effectively address scientific grand challenges. Using the US Department of Energy sponsored BioEnergy Science Center as an example, we highlight how impactful breakthroughs in biofuel science can be achieved within a large cross-disciplinary team environment. Such transformational insights are key to furthering our understanding and in generating models, theories and processes that can be used to overcome recalcitrance of biomass for sustainable biofuel production. Multidisciplinary approaches have an increasingly greater role to play in meeting rising demands for food, fibre, energy, clean environment and good health. Discoveries achieved by diverse minds and cross-applications of tools and analytical approaches have tremendous potential to fill existing knowledge gaps, clear roadblocks and facilitate translation of basic sciences discoveries as solutions towards addressing some of the most pressing global issues.

Keywords: bioenergy; ethanol; plant; microbe; integrated science

1. INTRODUCTION

The field of biology is undergoing a revolution and is becoming increasingly interdisciplinary. This interdisciplinarity is not simply in the context of subdisciplines of biological sciences such as physiology, biochemistry and molecular biology but, more strikingly, among fields as disparate as biology, chemistry, physics, mathematics and engineering. The accessibility of recent technological and conceptual innovations such as genome sequencing, transcriptome and proteome profiling technologies, easy-to-use experimental kits, high-resolution imaging and robotics has spurred design and fruition of many ambitious molecular characterization experiments in the animal, plant and microbial worlds.

The US National Academy of Sciences recently called for realizing a new era of biology called 'New Biology' (National Research Council (US) 2009; Kaiser 2009). The report urges complementing the traditional, fundamental curiosity-based single-researcher-driven structure with New Biology investigations to be conducted by interdisciplinary teams at the interface of the biological, chemical, physical, mathematical, computational sciences and engineering fields. In others words, New Biology is applicable to reductionist as well as systems biology studies, although it requires that the investigative approach be cross-disciplinary. Thus, New Biology is capable of filling gaps in our scientific knowledge to address some of the pressing global challenges as well as producing new inventions, spin-off products and industries, impacting economies in unimaginable ways.

Decades ago, Max Delbruck, a leading physicist engaged in biological research, identified a fundamental issue that still holds true. 'Biology is a very interesting field (because of) the vastness of its structure and the extraordinary variety of strange facts...but to the physicist it is also a depressing subject, because...the analysis seems to have stalled around in a semi-descriptive manner without noticeably progressing towards a radical physical explanation...we are not yet at the point where we are presented with clear paradoxes and this will not happen until the analysis of behaviour of living cells has been carried into far greater detail' (Delbruck 1949). In addition to Delbruck, other scientists with close ties to physics contributed to the birth of molecular biology. Warren Weaver, a mathematical physicist, originated the term 'molecular biology' in 1932; he foresaw 'that the moment would arrive when the distinction between chemistry and physics and even mathematics on the one hand and biology on the other would be so illusory' (Keller 1990).

In the past few decades, biology has made tremendous progress towards analysis of living systems

*Author for correspondence (kellerm@ornl.gov).

Received 23 December 2009
Accepted 10 May 2010

through a collection of powerful analytical methods. These include fast and inexpensive sequencing methods, technologies to measure RNA expression levels, the measurement and quantification of proteins present within cells and imaging methods approaching single-molecule resolution (Keller *et al.* 2002; Keller & Zengler 2004; Banfield *et al.* 2005; Keller & Ramos 2008; Moore & Weiss 2008; Morozova & Marra 2008; Goedhart & Gadella 2009; Hamady & Knight 2009; Keller & Hettich 2009). Probably, the most drastic advances were generated around DNA technologies such as accessing DNA from environments with low-abundance organisms (Abulencia *et al.* 2006), to sequencing DNA isolated and amplified from a single micro-organism (Zhang *et al.* 2006; Podar *et al.* 2007), to obtaining metagenomics data from very complex microbial populations (Tringe *et al.* 2005; Goldberg *et al.* 2006), to rendering higher eukaryotes with large genome sizes amenable to sequencing (The *Arabidopsis* Genome Initiative 2000; Venter *et al.* 2001; Tuskan *et al.* 2006), to simplifying genome-wide transcript and proteome profiling of model plants (Fukushima *et al.* 2009; Jorrin-Novo *et al.* 2009; Kalluri *et al.* 2009; Sakata *et al.* 2009).

One of the bottlenecks in our current investigative methods is in efficient generation, deposition, integration, analysis and mining of large and diverse datasets. Cross-disciplinary collaborations between biologists, instrument scientists and computational modelling and statistical experts are needed to address this bottleneck in the creation of 'virtual organism' growth models. New Biology ways can be powerful but are presently not fully realized as they are in their infancy, inadequately supported and delivering only a fraction of their potential (National Research Council (US) 2009). The changing atmosphere in life sciences is making way for a new kind of enquiry that, while building upon the traditional research establishment, provides a new interdisciplinary framework to take on grand science challenges in a more effective manner.

One such grand challenge is the sustainable production of biofuels. Ever-increasing demand for energy across the globe, volatile and decreasing sources of petroleum and concerns about climate change have created a need for focused biofuels research across many countries (Perlack *et al.* 2005; Himmel *et al.* 2007). Currently, biofuels are predominantly produced from corn grain starch and simple sugars from beets and sugar cane through a biological fermentation process. As highlighted in the workshop report, *Breaking the Biological Barriers to Cellulosic Ethanol* (DOE/SC-0095 2006), a critical barrier to developing a sustainable biofuels industry is the resistance, also 'recalcitrance', of lignocellulosic biomass to efficient enzymatic breakdown into sugars fermentable into ethanol or next-generation biofuels (DOE/SC-0095 2006). The processes to convert lignocellulosic biomass into liquid transportation fuel of desired chemistry can be broadly divided into two major steps: the breakdown of biomass into available sugars followed by a fermentation step to convert these sugars into biofuels. It has become clear that a key factor impeding large-scale biological conversion of cellulosic biomass into liquid fuels is the expense in processing the feedstock (Lynd *et al.* 2008). Therefore, research-guided advancements for the conversion of biomass to sugar offer a cost advantage when compared with advancements for converting sugar into fuel. To address the grand challenge of efficient and sustainable biofuel production, the US Department of Energy (DOE) created three bioenergy research centers (BRCs) in 2007. A peer-reviewed and competitive process resulted in the selection of Tennessee's Oak Ridge National Laboratory to lead the DOE BioEnergy Science Center (BESC), DOE's Lawrence Berkeley National Laboratory to lead the DOE Joint BioEnergy Institute and the University of Wisconsin–Madison to lead the DOE Great Lakes Bioenergy Research Center. Each of these centres would receive $135 million over 5 years to research biological solutions for bioenergy needs. Further, each BRC is a partnership among experts from multiple disciplines and Institutes (DOE/SC-0166 2009; Miller & Keller 2009).

Using BESC as a primary example, we highlight how a conducive environment in the form of new funding models; promotion of large cross-institutional, yet integrated, collaborations at the interface of physical, biological, computational and chemical sciences; and adoption of cutting-edge technologies is resulting in significant progress towards making biofuels a viable alternative energy option.

The BESC team is made up of 10 institutional partners as well as individual researchers from ten more universities and laboratories. BESC's goal is to unravel the factors defining recalcitrance such that lignocellulosic biomass can be converted easily to sugars (Davison *et al.* 2009). Addressing the roadblock of biomass recalcitrance will require a greater understanding of how plant cell walls are formed and how they are deconstructed by micro-organisms and enzymes. Therefore, the BESC is organized into three focus areas, namely biomass formation and modification, biomass deconstruction and conversion and a crosscutting focus area in characterization and modelling, to overcome biomass recalcitrance by developing a fundamental understanding of plant cell walls from synthesis to deconstruction. This understanding would generate models, theories and finally processes that will be used to understand and overcome biomass recalcitrance.

2. BIOMASS (PLANT CELL WALL) FORMATION

Plant cell walls, made primarily of cellulose, hemicelluloses, lignin, pectin and glycosylated proteins, are rigid yet flexible structures that play central roles in plant growth and development, water and solute conduction, defence against pathogens and strength of plant form (Higuchi 1997; Pauly & Keegstra 2008). It is estimated that the net CO_2 fixation by land plants per year is approximately 56×10^9 tons. The worldwide biomass production by land plants is $170–200 \times 10^9$ tons, 70 per cent of which is estimated to be made up of plant cell walls (Lieth 1975; Duchesne & Larson 1989; Poorter & Villar 1997; Field *et al.* 1998; Pauly &

Keegstra 2008), two per cent of which, is estimated to be consumed by human use (Schubert 2006).

Knowledge of formation, function, remodelling and breakdown of the complex plant cell wall chemistry and structure is far from complete. Most information is currently available from molecular genetic studies (Somerville *et al.* 2004; Geisler *et al.* 2008; Taylor 2008; York & O'Neill 2008; Caffall & Mohnen 2009). In the model plant *Arabidopsis*, genome sequencing revealed over 700 genes encoding putative glycosyltransferases or glycosyl hydrolases, as well as several hundred more genes coding for proteins functioning in cell wall biosynthesis or remodelling (Henrissat *et al.* 2001). Whole genome sequencing of model biofuel plants such as *Populus* has provided an opportunity to comprehend the extent of relative expansion of gene families involved in plant cell wall biosynthesis (Geisler-Lee *et al.* 2006; Tuskan *et al.* 2006). Recent advances in imaging and spectroscopy methods, including immunochemical labelling, spectroscopy (mass spectroscopy (MS), Fourier transform infrared, Raman and nuclear magnetic resonance (NMR) spectroscopy), electron microscopy and X-ray crystallography, have made possible the characterization of complex wall carbohydrates. The catalytic function of the predicted genes is generally inferred based on sequence homology. However, the precise functions of most of these proteins are unknown, calling for a broad systems biology approach.

In order to be able to improve cell wall deconstruction approaches, it is imperative to have a sound basal knowledge of cell wall assembly and function. Therefore, BESC is seeking to improve the feedstock characteristics of two model bioenergy crops, *Populus* and switchgrass, based on the understanding gained from integrated, multidisciplinary research on biomass formation and modification (Miller & Keller 2009). Towards that end, a wide range of molecular, genetic, genomic, biochemical, imaging, engineering and bioinformatics tools are being employed to further our understanding of cell wall structure, biosynthesis and the biological consequences of plant cell wall modification.

2.1. Plant cell wall polysaccharides

There are three major plant cell wall polysaccharides, broadly classified into cellulose, hemicelluloses and pectin. Cellulose is composed of unbranched $\beta(1\rightarrow4)$-linked D-glucan chains that are synthesized at the plasma membrane and aggregate laterally through H-bonding and Van der Waals forces to form crystalline cellulose. Hemicelluloses, amorphous structures composed of branched mixed-sugar polysaccharides, can form H-bonds with the surface of cellulose fibrils. Pectins, found in primary walls and characterized by the major component uronic acid, are thought to play significant roles in determining the porosity of cell walls and adhesion by adjoining cells and in controlling the cell wall's ionic environment (Carpita & Gibeaut 1993; Jarvis *et al.* 2003; Caffall & Mohnen 2009).

A major goal within the plant cell wall biosynthesis focus area is to identify genes that may affect biomass production and to then use this new biological information to generate plant materials with reduced recalcitrance (i.e. plants that are more easily deconstructed for biofuel production). One such approach that the biomass formation group has undertaken is a broad, non-biased *Populus* cell wall profiling approach. The approach is non-biased in that stem cores from several hundred individual poplar plants collected from wild, natural populations in the US Pacific Northwest were characterized for lignin, C6 sugar and C5 sugar distributions using molecular beam MS. The study found striking variations in lignin quantities and S : G ratios existing in the wild, whose implication for recalcitrance was tested using a standardized high-throughput sugar release assay and are being correlated with genotypic differences among individuals (Decker *et al.* 2009; Tuskan *et al.* 2010, unpublished data). This discovery was made possible on the basis of collaborations among biologists, ecologists, chemists, engineers and computational scientists, all funded through a common source and driven by a single ambitious goal.

It is also unknown how the ultrastructure of cellulose varies in relation to genetic and environmental variables in model bioenergy crops. In an attempt to address this, Foston *et al.* sought to determine the per cent crystallinity and lateral fibril dimensions of cellulose in stem core samples of 18 individual poplar plants from aforementioned collection of natural population. To characterize the lignocellulosic supramolecular and ultrastructure features in poplar, a ^{13}C cross-polarization magic angle spinning NMR was used on isolated cellulose. The study, a cross-collaborative BESC effort, found that there was no significant correlation between the sampling site and any of the measured properties for cellulose and lignin, thus suggesting the significant aspects affecting lignocellulosic structure and morphology may be genetic in nature (Foston *et al.* 2009).

2.2. Reducing biomass recalcitrance through genetic engineering of plant cell walls

The suitability of genetic manipulation studies in yielding fundamental insights into biosynthesis of plant cell wall components has previously been demonstrated through many different studies. Lignin content, composition, and cross-linking within the plant cell wall is believed to be a vital piece of the puzzle of biomass recalcitrance (Pu *et al.* 2008), and several attempts have been made at genetically engineering low-recalcitrance plants with decreased or altered lignin content and/or lignin components (Davison *et al.* 2006; Chen & Dixon 2007; Nakashima *et al.* 2008). Davison *et al.* demonstrated that changing lignin content increased xylose release in *Populus* upon dilute sulphuric acid hydrolysis (Davison *et al.* 2006). Chen and Dixon studied alfalfa plants downregulated for six lignin biosynthesis enzymes, showing improved fermentable sugar yields without acid pretreatment as compared with a wild-type that had included an acid pretreatment (Chen & Dixon 2007). In addition, the plants demonstrated increased digestibility with different lignin concentrations, suggesting that factors

other than the lignin content in plant cell walls determine digestibility. These results further strengthen the potential to generate feedstock plants that might be easily digestible without harsh pretreatment methods, which in turn would dramatically change the economics of biofuel production. This finding also serves as an example to demonstrate the importance of a multidisciplinary collaborative approach to investigating the factors underlying biomass recalcitrance (involving the aforementioned examples of chemical engineers and plant biologists engaged with BESC).

3. BIOMASS (PLANT CELL WALL) CHARACTERIZATION

3.1. New characterization techniques to analyse biomass properties

Various natural features of lignocellulosic biomass such as the epidermal tissue, the comparative amount of sclerenchymatous tissue, the degree of lignification and degree of cross-linking and deposition of wall polymers are known to contribute to the recalcitrance of feedstock materials (Himmel *et al.* 2007). A new mode of atomic force microscopy (AFM) was recently reported to be suitable for nanoscale characterization of lignocellulosic biomass (Tetard *et al.* 2010). AFM as well as mode-synthesizing AFM were demonstrated to be suitable for conducting topography and subsurface imaging with ongoing efforts to integrate an additional FTIR capability. Frequent and fruitful interactions between a chemical engineer, a plant scientist and physicists have resulted in the development of this new nanoscale chemical imaging approach for characterizing biomass.

Further progress in scaling-up approaches to increase digestibility has been hindered by lack of a greater knowledge of the plant cell wall biosynthesis pathway and influence of genetically altered cell wall chemistry and structure on biomass recalcitrance (Pu *et al.* 2009). We do not fully understand how, for example, downregulation of enzymes involved in the lignin precursor molecules influences the lignin polymer structure. A multidisciplinary, multi-institutional group of BESC researchers (chemists, a chemical engineer and plant biologist affiliated with Noble Foundation, the Georgia Institute of Technology and Oak Ridge National Laboratory) used one- and two-dimensional NMR techniques to determine structural changes in the lignin polymer isolated from genetically modified plants (Pu *et al.* 2009). The downregulation of C3H and HCT genes in alfalfa generated notable changes in lignin. The amount of *p*-hydroxyphenyl units increased, whereas the amount of guaiacyl and syringyl units decreased. Additionally, β-O-4 linkages and methoxyl groups in the lignin were seen to decrease by 55–58 and approximately 73 per cent, respectively (Pu *et al.* 2009). It is unknown whether lignin distribution in cell walls is affected by the downregulation of lignin biosynthesis. In general, a deeper knowledge of the chemical and structural architectures of plant cell walls is required to address these questions, even in the context of wild-type plant cells. Therefore, there is a need for new, higher resolution chemical imaging tools.

Over the past few years, nonlinear optical microscopy, such as coherent anti-Stokes Raman scattering (CARS), has shown to be a powerful tool for chemical imaging of biological systems (Ding *et al.* 2008; Zeng *et al.* in press). CARS microscopy provides a contrast mechanism based on the molecular vibrations that are intrinsic to the samples as well as high spatio-temporal resolution and is free of background from one-photon-excited fluorescence (Cheng & Xie 2004). Characterization of lignin-downregulated alfalfa plants using this *in situ* imaging approach supported observation of overall reduction in lignin content previously reported using a destructive chemical analysis method. The greatest reduction of lignin content was seen in the corners of cells and in the compound middle lamella, which could account for the enhanced susceptibility to chemical and enzymatic hydrolysis (Zeng *et al.* in press; Zhao *et al.* in press). With such new insights, obtained via collaborations between plant geneticists (Dixon and group) and microscopists pushing the frontiers of physics (Dixon and group 2009, personal communication) and facilitated by BESC's interdisciplinary focus, informed genetic modification of plant cell walls and therefore targeted changes in plant cell wall polymers and biomass recalcitrance can be achieved.

3.2. Characterization during pretreatment and destruction of biomass

Plant cell wall material can be converted to biofuels in several different ways. Combustion and gasification of plant materials results in CO and hydrogen gas (known as syngas), which can then be converted to hydrocarbons by catalyzed chemical reactions (Fisher–Tropsch process; Schubert 2006). Alternatively, syngas can be converted to alcohols or short-chain fatty acids via syngas-fermenting micro-organisms (Henstra *et al.* 2007). A distinct, purely biologically driven approach uses enzymes as biological catalysts to degrade plant cell wall material to monosaccharides, followed by fermentation to liquid fuels (Schubert 2006). Though degradation of plant cell material is a natural process, it is yet to be easily transferred into a cost-effective industrial process. Various methods have the potential to make plant cell walls more digestible. One option is to increase solubility in water and, therefore, access of enzymes to polysaccharides. Alternatively, altering the ratio of less-soluble polysaccharides to more-soluble polysaccharides may enhance digestibility (Pauly & Keegstra 2008).

BESC researchers have recently applied one-dimensional 13C and 31P NMR as well as two-dimensional HSQC NMR to examine the structure of ball-milled lignin isolated from untreated and dilute acid pretreated switchgrass (Samuel *et al.* 2009). It was revealed that ball-milled lignin from switchgrass was HGS type and included cinnamate esters. Further analysis showed that the major inner unit was the β-O-4 ether linkage, with minor amounts of phenylcoumarin, resinol and spirodienone units also present. The data also showed that changes in ball-milled lignin structure owing to

dilute H_2SO_4 pretreatment included a decrease in syringyl units and β-O-4 ether bond cleavage (Samuel *et al.* 2009). Donohoe *et al.* undertook an effort to characterize the nature of lignin that remains with the solids after acid pretreatment of biomass. Scanning electron microscopy and transmission electron microscopy demonstrated a range of varying shaped droplets appearing on and within cell walls of pretreated biomass (Donohoe *et al.* 2008). FTIR, NMR, cytochemical staining and antibody labelling showed that these droplets contain lignin (Donohoe *et al.* 2008). Donohoe *et al.* speculate that the thermochemical pretreatments reach temperatures exceeding those for lignin phase transition, which causes lignin to merge into larger bodies with the potential to travel into and out of the cell wall, resulting in redeposition on the plant cell wall surface. They further speculate that this movement of the lignin could be as significant as the removal of lignin for improving the digestibility of biomass (Donohoe *et al.* 2008). This finding serves as an example of the impact that cross-disciplinary research, in this case a combination of chemical and structural characterization under the BESC context, can have on gaining fundamental advances.

3.3. Computational approach to understanding cell wall structure and dynamics

Computational modelling approaches are complementary to experimental approaches in gaining an understanding of the molecular-scale dynamics of lignocellulose. BESC biophysicists, Petridis & Smith (2008), reported that the primary step to an accurate computational model of lignin would be to derive a molecular mechanics force field that complements the CHARMM force field. CHARMM is both a set of molecular dynamics (MD) force fields and an MD simulation and analysis package developed by Martin Karplus and his team (Brooks *et al.* 1983; MacKerell *et al.* 1998). The parametrization of the model is based on reproducing the quantum mechanical data of model compounds. Partial atomic charges are derived from methoxybenzene : water interactions. Dihedral parameters are optimized by fitting to critical rotational potentials, and bonded parameters are found by optimizing vibrational frequencies and normal modes. Finally, the force field is validated by performing a model dynamics simulation of a lignin fragment molecule and by comparing simulation-derived structural features with experimental results. With the existing force field for polysaccharides, this new lignin force field will enable full simulations of lignocellulose (Petridis & Smith 2008). However, an accurate computer simulation of the lignin component in lignocellulose will present significant challenges. Compared with numerous other biological macromolecules, already studied with molecular simulation, both the chemical and three-dimensional structures of lignin are less well understood. Furthering collaborative research among biochemists, physicists, computational scientists and instrument specialists is important to closing this research knowledge gap.

Xylan, a hemicellulose polymer of the monosaccharide unit xylose, coats the cellulose cores of cell wall microfibrils and can make up approximately 30 per cent of plant matter (Nimlos *et al.* 2006). Acid pretreatment of biomass, used to increase digestibility of plant material, is known to hydrolyze hemicellulose (mostly xylan) to release short-chain xylo-oligomers and monomeric xylose. This process also leads to the decomposition of xylose, and, to a much lesser extent, glucose, leading to the generation of furfuraldehydes. This process, significant in its long industrial use, is not fully understood (Nimlos *et al.* 2006). Nimlos *et al.* studied this decomposition of xylose using quantum mechanical calculations supported by NMR data. Energies were determined using various quantum calculations. The barriers for dehydration of the aldose were high compared with intramolecular rearrangement. This suggested a direct intramolecular rearrangement of the protonated pyranose as the mechanism of furfural formation (Nimlos *et al.* 2006).

In addition to the degradation of xylose to furfural, 5-(hydroxymethyl)-2-furaldehyde and other sugar degradation products, molecules of xylose and glucose can react with themselves or with each other in an acidic environment to form various disaccharides or oligomers, particularly at high initial sugar concentrations (Dong *et al.* 2009). These reversion reactions have been recognized as a cause of limited sugar yields from high-biomass solids, and moreover, having higher biomass solids is an important strategy for reducing biomass processing costs (Dong *et al.* 2009). So, in yet another example of application of a cross-disciplinary approach to studying biological processes at BESC, Car–Parrinello-based MD simulations combined with metadynamics simulations were used to determine the reaction energetics for the β-D-xylose condensation to form β-1,4-linked xylobiose in a dilute acid solution (Dong *et al.* 2009). The protonation of the hydroxyl group on the xylose molecule and the subsequent breakage of the C–O bond were found to be the rate-limiting step during the xylose condensation reaction. In addition, the water structure played an important role owing to the proton's high affinity for water molecules (Dong *et al.* 2009).

4. DECONSTRUCTION OF BIOMASS (PLANT CELL WALL)

Catalyst transport within plant biomass is thought to present a significant bottleneck during pretreatment and conversion of lignocellulose. Although reducing biomass size and homogenization can decrease limitations to catalyst transport, it adds cost and increases complexity to the process flow. Viamajala *et al.* followed the transport of an aqueous solution of a dye through intact corn internodes with high-resolution light microscopy (Viamajala *et al.* 2006). They found that both the air trapped inside cells and the waxy rind exterior were the significant blocks to dye transport. In contrast, the apoplastic continuum, vascular bundles and fissures formed during the drying process

provided the most frequently used paths for transport (Viamajala *et al.* 2006).

The biologically mediated hydrolysis of lignocellulosic plant material to accessible sugars is a critical step towards the development of a sustainable biofuels industry. The enzymatic degradation of cellulose is of central importance. The mechanism for enzymatic cellulose hydrolysis comprises a synergy of endoglucanases and cellobiohydrolases to hydrolyze solid cellulose on the surface to create soluble cellodextrins in the liquid phase, followed by β-glucosidase-mediated hydrolysis of cellobiose to glucose in the liquid phase (Zhang & Lynd 2006). The depolymerization reactions from long cellulose chains to short soluble cellodextrins are the rate-limiting steps for the overall hydrolysis process (Zhang & Lynd 2004). New biotechnological methods have been applied to further lower the cost of cellulases for biomass conversion. Most reductions in processing cost have come from increases in productivity of the enzymes, not in enzyme performance. Incremental cellulase performance improvements include engineering enzyme component mixtures, enzyme robustness and designing synergistic processing options (Ding *et al.* 2008).

Nature currently offers two kinds of cellulolytic systems: systems based on free, non-cell-bound, individual cellulases usually generated by aerobic fungi and bacteria, and systems based on complexes of cellulolytic enzymes, or cellulosomes, produced by some anaerobic bacteria (Ding *et al.* 2008; Wei *et al.* 2009). The free cellulase system of *Trichoderma* has received intensive attention owing to the high level of cellulase expression and secretion. The *Trichoderma reesei* cellulase mixture consists of many catalytically active proteins. At a minimum, two cellobiohydrolases (cellobiohydrolase I and II (CBH I and II)), five endoglucanases (EGI–V), β-glucosidases, and hemicellulases have been identified by two-dimensional electrophoresis (Vinzant *et al.* 2001). CBH I and CBH II and endoglucanase II (EGII) are the main components of the *T. reesei* cellulase system. CBH I (also called Cel7A) from *T. reesei* is one of the most effective exoglucanases known to date (Crowley *et al.* 2008).

4.1. Cel7A cellulolytic enzyme

Cel7A, a multi-domain enzyme, is made up of a large catalytic domain with an extensive active site tunnel and a small binding module (carbohydrate-binding module) joined to one another via a linker peptide. Although the precise mechanisms are unknown, it is believed that one chain of cellodextrins is removed from the crystal and is placed into the active site tunnel for hydrolysis to cellobiose. Following that step, Cel7A has been proposed to move along the target cellodextrin, cleaving a single cellobiose unit per catalytic event (Crowley *et al.* 2008). Experimental evidence for this proposed process is limited as is the detailed function of the carbohydrate binding module (CBM) and the linker peptide connecting the CBM to the catalytic domain. A group of BESC enzymologists and computational modellers have recently demonstrated that CBM recognizes and binds to the surface of cellulose (Nimlos *et al.* 2007). Zhao *et al.* surmised that the linker travels in a spring-like motion, enabling the enzyme to move on the cellulose surface while catalysis occurs. The hydrolysis of a cellulose strand compresses the length of the linker by the moving catalytic domain, and once the length is established to be short enough, potential energy allows the linker to move forward, releasing the binding module from the position on the cellulose surface (Zhao *et al.* 2008). To further improve the understanding of this enzyme, Crowley *et al.* created MD simulations of the binding and catalytic domains for cellulases with varying substrate configurations, solvation models and thermodynamic protocols based on deliberations with chemical engineers, physicists and biochemists. Most of the experiments required extensive alteration of the existing code and algorithms to use current terascale and petascale hardware to the degree of parallelism necessary to simulate a system of the desired size (Crowley *et al.* 2008). Zhao *et al.* demonstrated through a molecular simulation that the linker peptide has the ability to store energy much like a spring, supporting this processive motion of the Cel7A enzyme during cellulose depolymerization. Free energy calculations from these molecular simulations further support the idea. The linker has two stable states: at lengths 2.5 and 5.5 nm (during extension/compression), with a free energy difference of 10.5 kcal mol^{-1} between the two states separated by an energy barrier (Zhao *et al.* 2008). The aforementioned advances reported through BESC research underscore the great potential computational approach has in efforts to decipher biological phenomenon.

The CBMs of the cellobiohydrolases are recognized as being crucial for cellulase action on cellulose and are believed to have three primary functions, including proximity effects, substrate targeting and microcrystallite disruption (Gilkes *et al.* 1988; Kuutti *et al.* 1991; Reinikainen *et al.* 1992; Kruus *et al.* 1995; Linder & Teeri 1997; Boraston *et al.* 2004). Nimlos *et al.* used computational MD to research the interaction of the CBM from *T. reesei* CBH I with a model of the cellulose surface modified to display a broken chain. They found that the CBM was initially located in varying locations relative to the reducing end of the broken chain, and during the modelling, the CBM seemed to translate randomly across the cellulose surface, supporting its hypothesized role in processivity (Nimlos *et al.* 2007). Nimlos *et al.* also reported that the reducing end of a cellulose chain seems to cause a conformational change in the CBM. Their simulations demonstrated that the tyrosine residues on the hydrophobic surface of the CBM, Y5, Y31 and Y32, align with the cellulose chain at the reducing end, and that a fourth tyrosine residue, Y13, moves from its inner location to form Van der Waals interactions with the cellulose surface (Nimlos *et al.* 2007). In family I CBM, these four residues are highly conserved. As a result of this change in the residues near the surface, the CBM straddles the reducing end of the broken chain (Nimlos *et al.* 2007). This progress in understanding of CBM activity is due in part to the adoption of a computational approach to addressing a biological process.

Further light has recently been shed on the CBM and cellulose interactions using a nanotechnology-based approach where crystals of the family 2 CBMs from *Acidothermus celluloyticus* were labelled with quantum dots for tracking single-molecule motion (Liu *et al.* 2009). By tracking the movement of the single quantum dots, a linear motion along the axis of the cellulose fibre could be observed in contrast to quantum dots bound non-specifically to cellulose (Liu *et al.* 2009). This specific advancement was made possible by the coming together of computational scientists, biophysicists, enzymologists and instrument scientists to address a lingering biological question.

4.2. Composition and activity of cellulosome

Free cellulases are in a great contrast to complexes of cellulases found within cellulosomes. Results show that hydrolysis rates on model cellulosic substrates are higher for the complexed cellulase system of *Clostridium thermocellum* than for the non-complexed cellulase system of *T. reesei* (Ding *et al.* 2008). With the anaerobic thermophile, *C. thermocellum*, the cellulosome was first described in the early 1980s (Lamed & Bayer 1988) and was also characterized from various other anaerobic bacteria (Doi & Kosugi 2004). Composed of two types of subunits, the cellulosome has non-catalytic scaffoldins and catalytically active components. The cellulosome's assembly is aided by the highly specific binding between cohesin modules of the scaffoldin subunit and dockerin modules located on the enzyme. The scaffoldin may contain several cohesin modules, allowing multiple different enzymes to be assembled into the cellulosome complex. Some species have multiple scaffoldins, lending a higher level of complexity to the cellulosome assembly. Closer to 100 components can be assembled into an individual cellulosome of *A. cellulolytics* (Ding *et al.* 2008).

With the increasing number of genome sequences of cellulose-forming micro-organisms available, MS-based methods are being applied as a tool to identify and experimentally confirm the expression of cellulosomal proteins. Zverlov & Schwarz identified 32 cellulosomal protein components across four different cellulosomal samples from *C. thermocellum* cultures using MALDI-TOF/TOF MS (Zverlov & Schwarz 2008). Gold & Martin performed a metabolic isotope labelling in conjunction with liquid chromatography-tandem mass spectrometry to estimate quantitative changes in expression patterns of *C. thermocellum* cellulosomal subunits during growth on cellulose and cellobiose. Forty-one cellulosomal components were identified qualitatively between the two samples (Gold & Martin 2007). However, more than 20 per cent of these cellulosomal proteins have domains with no assigned function, and additional research is needed for experimental verification of more than 40 per cent of the currently hypothetical cellulosomal proteins (Gold & Martin 2007; Raman *et al.* 2009). In an attempt to shed further light, a group of BESC chemical engineers, microbiologists, spectroscopists and computational biologists performed a comprehensive comparison of cellulosomal compositional changes in *C. thermocellum* in response to different carbon sources (Raman *et al.* 2009). *Clostridium thermocellum* was grown on various carbon sources, including dilute-acid-pretreated switchgrass, amorphous cellulose, cellobiose, crystalline cellulose and combinations of crystalline cellulose with pectin or xylan. Fifty-nine dockerin- and eight cohesin-module containing components, including 16 previously undetected cellulosomal subunits, were identified by proteomic analysis. Further, many of these components showed differential protein abundance when grown in the presence of non-cellulose substrates (Raman *et al.* 2009). The precise functions of several of the cellulosomal subunits remain unknown. Therefore, heterologous expression of individual cellulosomal subunits and the *in vitro* assembly in artificial cellulosomes combined with computer simulations are being explored to provide the much-needed insights into the underpinnings of this high-activity, high-efficiency machinery.

5. CONCLUDING REMARKS

There is little doubt that life science research will have a prominent role to play in meeting the rising needs for food, energy, clean environment and good health. Leveraging science towards addressing these colossal goals, complicated further under the climate change scenario, is a challenge not limited to one scientist, group or country but to the global science community.

Multidisciplinary biofuel research teams are capable of achieving new scientific insights that were thus far not possible. This potential is facilitated by the presence of a conducive environment in the form of new funding models and promotion of large, integrated cross-institutional collaborations. Social engineering of such collaborations is no simple matter, and certain fundamentals are a necessity: inclusion of eager, enthusiastic or impressionable minds, a targeted research context, mutually interesting goals, collective solution finding and frequent and regular in-person or virtual opportunities to interact. With increasingly complex individual research portfolios, the challenge then becomes continued commitment and focus towards the larger research goals and timeframes in order to nurture and realize the potential of New Biology in achieving transformational scientific breakthroughs. It will, therefore, be important to embrace the excitement and unexplored potential in the interfacial space of the science and engineering fields and garner timely momentum in turning the tide towards New Biology.

Four predictions are made on how such science initiatives sponsored by interagency funding will be conducted in the future. There will be (i) wide-spread adoption of integrative computational and engineering tools in systems biology studies, (ii) escalation in the number of community resource and data repositories, (iii) emergence of newer biotechnological approaches and (iv) increase in large interdisciplinary collaborative efforts. The new era of biological research is expected to have increased sophistication in hypothesis generation, experimental design and toolset, and in the data analysis, modelling and simulation areas. The improvements

in advanced imaging will need to be brought about in a collaborative fashion with mathematicians and computational experts to carry out data extraction, analysis, integration and visualization along time and space gradients. Simulation screenings that are presently required in evaluating earthquake resistance of tall buildings and screening drug effects can in the future be commonplace to test genetically improved plants and microbes prior to field deployment. Such modelling of information flow from molecular to whole organism level will be important in New Biology research programmes, where simulation and modelling can be carried out to identify bottlenecks in feasibility, to identify feedback mechanisms and to predict non-target effects of perturbations.

Standardized guidelines in data collection, description and analysis will improve the power and ability in scaling the individual small-scale studies to large-scale understanding at systems level (Kitano 2002). The time is apt to formulate acceptable standards on experimental controls, number of replications, appropriate tool settings and data repository and access mechanisms as a way to facilitate integrated research efforts. The practical challenge in formulating standards is determining the scope of the effort; what should drive the effort, who should participate and who should adopt. If the ideal case scenario of global scientific standards is too challenging as a first step, then would it be acceptable to develop such standards on the basis on one research consortium, one science focus area, one agency or one country? How could the bases of such databases be flexible to emerging technologies and terminologies? It is timely to actively engage in these discussions within our national and international scientific societies.

In the present review, we have provided examples of research progress achieved through collaborations among scientists. It also needs to be brought to the forefront here that under the BESC context, as an example, there are regular communications among the academic, government and industry sectors. Such a layer of collaboration is highly significant in ensuring that the design and interpretation of laboratory results are extrapolated to real-world scenarios and in turn impact the return on investment of research funds positively (Tilman *et al.* 2009).

It is a calling to us scientists, across all scientific disciplines, to overcome the traditional scientific stovepiping and gain momentum towards arriving at creative multidisciplinary solutions. Discoveries springing out of such meeting of minds and cross-applications of tools and analytical approaches have tremendous potential to fill existing knowledge gaps, clear roadblocks and facilitate translation of basic sciences discoveries towards applied causes such as improved food and energy crops adapted to changing environment. Moreover, the concepts, frameworks and resources developed for a particular purpose can open new doors in other related fields as well. As an example, discoveries made to improve productivity and stress tolerance of biofuel crops in the Philippines could be readily explored for applicability in developing crops to combat poverty and stamp out hunger in sub-Saharan Africa or even in an unrelated topic such as personalized medicine and hence enabling the attainment of a worthy, yet unintended, goal.

The merits are foreseeable; however, the actualization of these will require scientists to find common language and vision with additional support in the form of changes within the current funding and organizational structures. Integration must be further promoted across the disciplines at institutional, funding agency and scientific community levels coupled with deliberations on the scientific, economic, environmental, societal and ethical implications so that the long-range goals of bio-based research efforts are best realized. Last but not the least, we must bear in mind the expectations on us from our community to meet rising global challenges as we prepare to teach yet another freshman student batch at our colleges.

The authors would like to thank Brian Davison and Meghan Drake for their suggestions and assistance in preparing this manuscript. The BioEnergy Science Center is a US Department of Energy Bioenergy Research Center supported by the Office of Biological and Environmental Research in the DOE Office of Science. ORNL is managed by UT-Battelle, LLC, under contract DE-AC05-00OR22725 for the US Department of Energy.

REFERENCES

Abulencia, C. B. *et al.* 2006 Environmental whole-genome amplification to access microbial populations in contaminated sediments. *Appl. Environ. Microbiol.* **72**, 3291–3301. (doi:10.1128/AEM.72.5.3291-3301.2006)

Banfield, J. F., Verberkmoes, N. C., Hettich, R. L. & Thelen, M. P. 2005 Proteogenomic approaches for the molecular characterization of natural microbial communities. *Omics J. Integr. Biol.* **9**, 301–333. (doi:10.1089/omi.2005.9.301)

Boraston, A. B., Bolam, D. N., Gilbert, H. J. & Davies, G. J. 2004 Carbohydrate-binding modules: fine-tuning polysaccharide recognition. *Biochem. J.* **382**, 769–781. (doi:10.1042/BJ20040892)

Brooks, B. R., Bruccoleri, R. E., Olafson, B. D., States, D. J., Swaminathan, S. & Karplus, M. 1983 CHARMM—a program for macromolecular energy, minimization, and dynamics calculations. *J. Comput. Chem.* **4**, 187–217. (doi:10.1002/jcc.540040211)

Caffall, K. H. & Mohnen, D. 2009 The structure, function, and biosynthesis of plant cell wall pectic polysaccharides. *Carbohydr. Res.* **344**, 1879–1900. (doi:10.1016/j.carres.2009.05.021)

Carpita, N. C. & Gibeaut, D. M. 1993 Structural models of primary-cell walls in flowering plants—consistancy of molecular-structure with the physical-properties of the walls during growth. *Plant J.* **3**, 1–30. (doi:10.1111/j.1365-313X.1993.tb00007.x)

Chen, F. & Dixon, R. A. 2007 Lignin modification improves fermentable sugar yields for biofuel production. *Nat. Biotechnol.* **25**, 759–761. (doi:10.1038/nbt1316)

Cheng, J. X. & Xie, X. S. 2004 Coherent anti-Stokes Raman scattering microscopy: instrumentation, theory, and applications. *J. Phys. Chem. B* **108**, 827–840. (doi:10.1021/jp035693v)

Crowley, M. F., Uberbacher, E. C., Brooks, C. L., Walker, R. C., Nimlos, M. R. & Himmel, M. E. 2008 Developing improved MD codes for understanding processive cellulases. *J. Phys. Conf. Ser.* **125**, 012049. See http://iopscience.iop.org/1742-6596/125/1/012049

Davison, B. H., Drescher, S. R., Tuskan, G. A., Davis, M. F. & Nghiem, N. P. 2006 Variation of S/G ratio and lignin content in a *Populus* family influences the release of xylose by dilute acid hydrolysis. *Appl. Biochem. Biotechnol.* **130**, 427–435. (doi:10.1385/ABAB:130:1:427)

Davison, B. H., Keller, M. & Fowler, V. S. 2009 The goals and research of the bioenergy sciences center (BESC): developing cost-effective and sustainable means of producing biofuels by overcoming biomass recalcitrance. *BioEnergy Res.* **2**, 177–178. (doi:10.1007/s12155-009-9057-7)

Decker, S. R., Brunecky, R., Tucker, M. P., Himmel, M. E. & Selig, M. J. 2009 High-throughput screening techniques for biomass conversion. *BioEnergy Res.* **2**, 179–192. (doi:10.1007/s12155-009-9051-0)

Delbruck, M. A. 1949 A physicist looks at biology. *Trans. Conn. Acad. Arts Sci.* **38**, 173–190.

Ding, S. Y., Xu, Q., Crowley, M., Zeng, Y., Nimlos, M., Lamed, R., Bayer, E. A. & Himmel, M. E. 2008 A biophysical perspective on the cellulosome: new opportunities for biomass conversion. *Curr. Opin. Biotechnol.* **19**, 218–227. (doi:10.1016/j.copbio.2008.04.008)

DOE/SC-0166 2009 U.S. Department of Energy's Bioenergy Research Centers, An Overview of the Science. See http://genomicscience.energy.gov/centers/brcbrochure.pdf.

DOE/SC-0095 2006 US Department of Energy. Breaking the biological barriers to cellulosic ethanol: a joint research agenda. See http://genomicscience.energy.gov/biofuels/

Doi, R. H. & Kosugi, A. 2004 Cellulosomes: plant-cell-wall-degrading enzyme complexes. *Nat. Rev. Microbiol.* **2**, 541–551.

Dong, H. T., Nimlos, M. R., Himmel, M. E., Johnson, D. K. & Qian, X. H. 2009 The effects of water on beta-D-Xylose condensation reactions. *J. Phys. Chem. A* **113**, 8577–8585. (doi:10.1021/jp9025442)

Donohoe, B. S., Decker, S. R., Tucker, M. P., Himmel, M. E. & Vinzant, T. B. 2008 Visualizing lignin coalescence and migration through maize cell walls following thermochemical pretreatment. *Biotechnol. Bioeng.* **101**, 913–925. (doi:10.1002/bit.21959)

Duchesne, L. C. & Larson, D. W. 1989 Cellulose and the evolution of plant life. *Bioscience* **39**, 238–241. (doi:10.2307/1311160)

Field, C. B., Behrenfeld, M. J., Randerson, J. T. & Falkowski, P. 1998 Primary production of the biosphere: integrating terrestrial and oceanic components. *Science* **281**, 237–240. (doi:10.1126/science.281.5374.237)

Foston, M., Hubbell, C. A., Davis, M. & Ragauskas, A. J. 2009 Variations in cellulosic ultrastructure of poplar. *BioEnergy Res.* **2**, 193–197. (doi:10.1007/s12155-009-9050-1)

Fukushima, A., Kusano, M., Redestig, H., Arita, M. & Saito, K. 2009 Integrated omics approaches in plant systems biology. *Curr. Opin. Chem. Biol.* **13**, 532–538. (doi:10.1016/j.cbpa.2009.09.022)

Geisler, D. A., Sampathkumar, A., Mutwil, M. & Persson, S. 2008 Laying down the bricks: logistic aspects of cell wall biosynthesis. *Curr. Opin. Plant Biol.* **11**, 647–652. (doi:10.1016/j.pbi.2008.08.003)

Geisler-Lee, J. *et al.* 2006 Poplar carbohydrate-active enzymes. Gene identification and expression analyses. *Plant Physiol.* **140**, 946–962. (doi:10.1104/pp.105.072652)

Gilkes, N. R., Warren, R. A. J., Miller, R. C. & Kilburn, D. G. 1988 Precise excision of the cellulose binding domains from two *Cellulomonas fimi* cellulases by a homologous protease and the effect on catalysis. *J. Biol. Chem.* **263**, 10 401–10 407.

Goedhart, J. & Gadella, T. W. J. 2009 Fluorescence resonance energy transfer imaging of PKC signalling in living cells using genetically encoded fluorescent probes. *J. R. Soc. Interface* **6**, S27–S34. (doi:10.1098/rsif.2008.0340.focus)

Gold, N. D. & Martin, V. J. J. 2007 Global view of the *Clostridium thermocellum* cellulosome revealed by quantitative proteomic analysis. *J. Bacteriol.* **189**, 6787–6795. (doi:10.1128/JB.00882-07)

Goldberg, S. M. D. *et al.* 2006 A Sanger/pyrosequencing hybrid approach for the generation of high-quality draft assemblies of marine microbial genomes. *Proc. Natl Acad. Sci. USA* **103**, 11 240–11 245. (doi:10.1073/pnas.0604351103)

Hamady, M. & Knight, R. 2009 Microbial community profiling for human microbiome projects: tools, techniques, and challenges. *Genome Res.* **19**, 1141–1152. (doi:10.1101/gr.085464.108)

Henrissat, B., Coutinho, P. M. & Davies, G. J. 2001 A census of carbohydrate-active enzymes in the genome of *Arabidopsis thaliana*. *Plant Mol. Biol.* **47**, 55–72. (doi:10.1023/A:1010667012056)

Henstra, A. M., Sipma, J., Rinzema, A. & Stams, A. J. M. 2007 Microbiology of synthesis gas fermentation for biofuel production.*Curr. Opin. Biotechnol.* **18**, 200–206.

Higuchi, T. 1997 *Biochemistry and molecular biology of wood,* 1st edn. Berlin, Germany: Springer.

Himmel, M. E., Ding, S. Y., Johnson, D. K., Adney, W. S., Nimlos, M. R., Brady, J. W. & Foust, T. D. 2007 Biomass recalcitrance: engineering plants and enzymes for biofuels production. *Science* **315**, 804–807. (doi:10.1126/science.1137016)

Jarvis, M. C., Briggs, S. P. H. & Knox, J. P. 2003 Intercellular adhesion and cell separation in plants. *Plant Cell Environ.* **26**, 977–989. (doi:10.1046/j.1365-3040.2003.01034.x)

Jorrin-Novo, J. V. *et al.* 2009 Plant proteomics update (2007–2008): second-generation proteomic techniques, an appropriate experimental design, and data analysis to fulfill MIAPE standards, increase plant proteome coverage and expand biological knowledge. *J. Proteomics* **72**, 285–314. (doi:10.1016/j.jprot.2009.01.026)

Kaiser, J. 2009 Life sciences a new biology to mend society's woes. *Science* **325**, 1609–1609. (doi:10.1126/science.325_1609a)

Kalluri, U. C., Hurst, G. B., Lankford, P. K., Ranjan, P. & Pelletier, D. A. 2009 Shotgun proteome profile of *Populus* developing xylem. *Proteomics* **9**, 4871–4880. (doi:10.1002/pmic.200800854)

Keller, E. F. 1990 Physics and the emergence of molecular—a history of cognitive and political synergy. *J. History Biol.* **23**, 389–409. (doi:10.1007/BF00136376)

Keller, M. & Hettich, R. 2009 Environmental proteomics: a paradigm shift in characterizing microbial activities at the molecular level. *Microbiol. Mol. Biol. Rev.* **73**, 62. (doi:10.1128/MMBR.00028-08)

Keller, M. & Ramos, J. L. 2008 Microbial goods from single cells and metagenomes. *Curr. Opin. Microbiol.* **11**, 195–197. (doi:10.1016/j.mib.2008.05.003)

Keller, M. & Zengler, K. 2004 Tapping into microbial diversity. *Nat. Rev. Microbiol.* **2**, 141–150. (doi:10.1038/nrmicro819)

Keller, M., Zengler, K., Toledo, G., Elkins, J., Mathur, E. & Short, J. 2002 Microbiology: what is next?.*Geochim. Cosmochim. Acta* **66**, A390–A390.

Kitano, H. 2002 Systems biology: a brief overview. *Science* **295**, 1662–1664. (doi:10.1126/science.1069492)

Kruus, K., Lua, A. C., Demain, A. L. & Wu, J. H. D. 1995 The anchorage function of CipA (CelL), a scaffolding protein of the *Clostridium thermocellum* cellulosome.

Proc. Natl Acad. Sci. USA **92**, 9254–9258. (doi:10.1073/pnas.92.20.9254)

Kuutti, L., Laaksonen, L. & Teeri, T. T. 1991 Interaction studies of the tail domain of cellobiohydrolace I and crystalline cellulose using molecular modeling. *J. Chim. Phys. Chim. Biol.* **88**, 2663–2667.

Lamed, R. & Bayer, E. A. 1988 The cellulosome of *Clostridium thermocellum*. *Adv. Appl. Microbiol.* **33**, 1–46. (doi:10.1016/S0065-2164(08)70203-X)

Lieth, H. 1975 Primary production of the major vegetation units of the world. In *Primary production of the biosphere* (eds H. Lieth & R. H. Whittaker), pp. 203–215. Berlin, Germany: Springer.

Linder, M. & Teeri, T. T. 1997 The roles and function of cellulose-binding domains. *J. Biotechnol.* **57**, 15–28. (doi:10.1016/S0168-1656(97)00087-4)

Liu, Y. S., Zeng, Y. N., Luo, Y. H., Xu, Q., Himmel, M. E., Smith, S. J. & Ding, S. Y. 2009 Does the cellulose-binding module move on the cellulose surface? *Cellulose* **16**, 587–597. (doi:10.1007/s10570-009-9306-0)

Lynd, L. R. *et al.* 2008 How biotech can transform biofuels. *Nat. Biotechnol.* **26**, 169–172. (doi:10.1038/nbt0208-169)

MacKerell, A. D. *et al.* 1998 All-atom empirical potential for molecular modeling and dynamics studies of proteins. *J. Phys. Chem. B* **102**, 3586–3616. (doi:10.1021/jp973084f)

Miller, R. & Keller, M. 2009 The DOE bioEnergy science center-a US department of energy bioenergy research center. *Vitro Cell. Dev. Biol. Plant* **45**, 193–198. (doi:10.1007/s11627-009-9213-y)

Moore, A. M. & Weiss, P. S. 2008 Functional and spectroscopic measurements with scanning tunneling microscopy. *Ann. Rev. Anal. Chem.* **1**, 857–882. (doi:10.1146/annurev.anchem.1.031207.112932)

Morozova, O. & Marra, M. A. 2008 Applications of next-generation sequencing technologies in functional genomics. *Genomics* **92**, 255–264. (doi:10.1016/j.ygeno.2008.07.001)

Nakashima, J., Chen, F., Jackson, L., Shadle, G. & Dixon, R. A. 2008 Multi-site genetic modification of monolignol biosynthesis in alfalfa (*Medicago sativa*): effects on lignin composition in specific cell types. *New Phytol.* **179**, 738–750. (doi:10.1111/j.1469-8137.2008.02502.x)

National Research Council (US). 2009 Committee on a New Biology for the 21st Century: ensuring the United States leads the coming biology revolution. *A new biology for the 21st century*. Washington DC: National Academic Press.

Nimlos, M. R., Qian, X. H., Davis, M., Himmel, M. E. & Johnson, D. K. 2006 Energetics of xylose decomposition as determined using quantum mechanics modeling. *J. Phys. Chem. A* **110**, 11 824–11 838. (doi:10.1021/jp0626770)

Nimlos, M. R. *et al.* 2007 Molecular modeling suggests induced fit of Family I carbohydrate-binding modules with a broken-chain cellulose surface. *Protein Eng. Des. Sel.* **20**, 179–187. (doi:10.1093/protein/gzm010)

Pauly, M. & Keegstra, K. 2008 Cell-wall carbohydrates and their modification as a resource for biofuels. *Plant J.* **54**, 559–568. (doi:10.1111/j.1365-313X.2008.03463.x)

Perlack, R., Wright, L., Turhollow, A., Graham, R., Stokes, B. & Erbach, D. 2005 Biomass as feedstock for a bioenergy and bioproducts industry: the technical feasibility of a billion-ton annual supply, A joint study sponsored by the U.S. Departments of Energy and Agriculture. DOE/GO-102005-2135, ORNL/TM-2005/66, April.

Petridis, L. & Smith, J. C. 2008 Cellulosic ethanol: progress towards a simulation model of lignocellulosic biomass. *J. Phys. Conf. Ser.* **125**, 012055. See http://iopscience.iop.org/1742-6596/125/1/012055

Podar, M. *et al.* 2007 Targeted access to the genomes of low-abundance organisms in complex microbial communities. *Appl. Environ. Microbiol.* **73**, 3205–3214. (doi:10.1128/AEM.02985-06)

Poorter, H. & Villar, R. 1997 The fate of acquired carbon in plants: chemical composition and construction costs. In *Plant resource allocation* (eds F. A. Bazzaz & J. Grace), pp. 39–72. San Diego, CA: Academic Press.

Pu, Y. Q., Zhang, D. C., Singh, P. M. & Ragauskas, A. J. 2008 The new forestry biofuels sector. *Biofuels Bioproducts Biorefining* **2**, 58–73. (doi:10.1002/bbb.48)

Pu, Y., Chen, F., Ziebell, A., Davison, B. H. & Ragauskas, A. J. 2009 NMR characterization of C3H and HCT down-regulated alfalfa lignin. *BioEnergy Res.* **2**, 198–208. (doi:10.1007/s12155-009-9056-8)

Raman, B., Pan, C., Hurst, G. B., Rodriguez Jr, M., McKeown, C. K., Lankford, P. K., Samatova, N. F. & Mielenz, J. R. 2009 Impact of pretreated switchgrass and biomass carbohydrates on *Clostridium thermocellum* ATCC 27405 cellulosome composition: a quantitative proteomic analysis. *PLoS ONE* **4**, e5271. (doi:10.1371/journal.pone.0005271)

Reinikainen, T., Ruohonen, L., Nevanen, T., Laaksonen, L., Kraulis, P., Jones, T. A., Knowles, K. C. & Teeri, T. T. 1992 Investigation of the function of mutated cellulose-binding domains of *Trichoderma reesei* cellobiohydrolase I. *Proteins* **14**, 475–482. (doi:10.1002/prot.340140408)

Sakata, K., Ohyanagi, H., Nobori, H., Nakamura, T., Hashiguchi, A., Nanjo, Y., Mikami, Y., Yunokawa, H. & Komatsu, S. 2009 Soybean proteome database: a data resource for plant differential omics. *J. Proteome Res.* **8**, 3539–3548. (doi:10.1021/pr900229k)

Samuel, R., Pu, Y., Raman, B. & Ragauskas, A. 2009 Structural characterization and comparison of switchgrass ball-milled lignin before and after dilute acid pretreatment. *Appl. Biochem. Biotechnol.* **162**, 62–74. (doi:10.1007/s12010-009-8749-y)

Schubert, C. 2006 Can biofuels finally take center stage? *Nat. Biotechnol.* **24**, 777–784. (doi:10.1038/nbt0706-777)

Somerville, C. *et al.* 2004 Toward a systems approach to understanding plant-cell walls. *Science* **306**, 2206–2211. (doi:10.1126/science.1102765)

Taylor, N. G. 2008 Cellulose biosynthesis and deposition in higher plants. *New Phytol.* **178**, 239–252. (doi:10.1111/j.1469-8137.2008.02385.x)

Tetard, L., Passian, A., Farahi, R. H., Kalluri, U. C., Davison, B. H. & Thundat, T. 2009 Spectroscopy and atomic force microscopy of biomass. *Ultramicroscopy* **110**, 701–707.

The *Arabidopsis* Genome Initiative 2000 Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. *Nature* **408**, 796–815. (doi:10.1038/35048692)

Tilman, D. *et al.* 2009 Beneficial biofuels—the food, energy, and environment trilemma. *Science* **325**, 270–271. (doi:10.1126/science.1177970)

Tringe, S. G. *et al.* 2005 Comparative metagenomics of microbial communities. *Science* **308**, 554–557. (doi:10.1126/science.1107851)

Tuskan, G. A. *et al.* 2006 The genome of black cottonwood, *Populus trichocarpa* (Torr. & Gray). *Science* **313**, 1596–1604. (doi:10.1126/science.1128691)

Venter, J. C. *et al.* 2001 The sequence of the human genome. *Science* **291**, 1304–1351. (doi:10.1126/science.1058040)

Viamajala, S., Selig, M. J., Vinzant, T. B., Tucker, M. P., Himmel, M. E., McMillan, J. D. & Decker, S. R. 2006

Catalyst transport in corn stover internodes—elucidating transport mechanisms using Direct Blue-I. *Appl. Biochem. Biotechnol.* **130**, 509–527.

Vinzant, T. B., Adney, W. S., Decker, S. R., Baker, J. O., Kinter, M. T., Sherman, N. E., Fox, J. W. & Himmel, M. E. 2001 Fingerprinting *Trichoderma reesei* hydrolases in a commerical cellulase preparation. *Appl. Biochem. Technol.* **93**, 99–107.

Wei, H., Xu, Q., Taylor 2nd, L. E., Baker, J. O., Tucker, M. P. & Ding, S. Y. 2009 Natural paradigms of plant cell wall degradation. *Curr. Opin. Biotechnol.* **20**, 330–338. (doi:10.1016/j.copbio.2009.05.008)

York, W. S. & O'Neill, M. A. 2008 Biochemical control of xylan biosynthesis—which end is up? *Curr. Opin. Plant Biol.* **11**, 258–265. (doi:10.1016/j.pbi.2008.02.007)

Zeng, Y., Saar, B., Friedrich, M. G., Chen, F., Liu, Y.-S., Dixon, R. A., Himmel, M. E., Xie, X. S. & Ding, S.-Y. In press. Imaging lignin-down-regulated alfalfa using coherent anti-Stokes Raman scattering microscopy. *BioEnergy Res.* (doi:10.1007/s12155-010-9079-1)

Zhang, Y. H. P. & Lynd, L. R. 2004 Toward an aggregated understanding of enzymatic hydrolysis of cellulose: noncomplexed cellulase systems. *Biotechnol. Bioeng.* **88**, 797–824. (doi:10.1002/bit.20282)

Zhang, Y. H. P. & Lynd, L. R. 2006 A functionally based model for hydrolysis of cellulose by fungal cellulase. *Biotechnol. Bioeng.* **94**, 888–898. (doi:10.1002/bit.20906)

Zhang, K., Martiny, A. C., Reppas, N. B., Barry, K. W., Malek, J., Chisholm, S. W. & Church, G. M. 2006 Sequencing genomes from single cells by polymerase cloning. *Nat. Biotechnol.* **24**, 680–686. (doi:10.1038/nbt1214)

Zhao, Q., Gallego-Giraldo, L., Wang, H., Zeng, Y., Ding, S.-Y., Chen, F. & Dixon, R. A. In press. A NAC transcription factor orchestrates multiple features of cell wall development in *Medicago truncatula*. *Plant J.* (doi:10.1111/j.1365-313X.2010.04223)

Zhao, X., Rignall, T. R., McCabe, C., Adney, W. S. & Himmel, M. E. 2008 Molecular simulation evidence for processive motion of *Trichoderma reesei* Cel7A during cellulose depolymerization. *Chem. Phys. Lett.* **460**, 284–288. (doi:10.1016/j.cplett.2008.05.060)

Zverlov, V. V. & Schwarz, W. H. 2008 Bacterial cellulose hydrolysis in anaerobic environmental subsystems—*Clostridium thermocellum* and *Clostridium stercorarium*, thermophilic plant-fiber degraders. *Ann. NY Acad. Sci.* **1125**, 298–307.