DOE/SC-0170

Lignocellulosic Biomass for Advanced Biofuels and Bioproducts

U.S. DEPARTMENT OF ENERGY | Office of Science

Office of Biological and Environmental Research

Bioenergy Workshop

June 23–24, 2014, Washington, D.C.

Convened by

U.S. Department of Energy
Office of Science
Office of Biological and Environmental Research

Co-Chairs

Erich Grotewold, Ph.D
Ohio State University

Kristala L. Jones Prather, Ph.D
Massachusetts Institute of Technology

Organizer

Biological Systems Science Division

Kent Peters, Ph.D.
Kent.Peters@science.doe.gov
301.903.5549

http://genomicscience.energy.gov/biofuels/lignocellulose/

Mission

The Office of Biological and Environmental Research (BER) advances world-class fundamental research programs and scientific user facilities to support the Department of Energy's energy, environment, and basic research missions. Addressing diverse and critical global challenges, the BER program seeks to understand how genomic information is translated to functional capabilities, enabling more confident redesign of microbes and plants for sustainable biofuel production, improved carbon storage, or contaminant bioremediation. BER research advances understanding of the roles of Earth's biogeochemical systems (the atmosphere, land, oceans, sea ice, and subsurface) in determining climate so that it can be predicted decades or centuries into the future, information needed to plan for energy and resource needs. Solutions to these challenges are driven by a foundation of scientific knowledge and inquiry in atmospheric chemistry and physics, ecology, biology, and biogeochemistry.

Cover Credits

Brown Column: From top to bottom. Poplar tree tops (Oak Ridge National Laboratory). Switchgrass (U.S. Department of Agriculture's Agricultural Research Service). Atomic force micrograph showing nanometer-scale detail of interwoven rope-like, lignocellulosic microfibril bundles in a switchgrass cell wall (BioEnergy Science Center and National Renewable Energy Laboratory). Scanning electron micrograph of *Clostridium cellulolyticum* cells growing on switchgrass (BioEnergy Science Center and National Renewable Energy Laboratory). ***Green Background:*** Biorefinery at top (iStock) and switchgrass at bottom (Oak Ridge National Laboratory).

Suggested citation for this report: U.S. DOE. 2015. *Lignocellulosic Biomass for Advanced Biofuels and Bioproducts: Workshop Report*, DOE/SC-0170. U.S. Department of Energy Office of Science. http://genomicscience.energy.gov/biofuels/lignocellulose/.

DOE/SC-0170

Lignocellulosic Biomass for Advanced Biofuels and Bioproducts

Workshop Report

Published: February 2015

U.S. DEPARTMENT OF ENERGY | Office of Science

Office of Biological and Environmental Research

Table of Contents

Executive Summary v
Introduction 1
Biomass Development 9
- **State of the Science** 9
- **Remaining Biological Research Challenges** 11
- Understanding Cell Wall Synthesis, Structure, and Function 11
- Optimizing Carbon Delivery for Lignocellulose Production 12
- Increasing Biomass Yields 13
- Ensuring Sustainability 14
- **Key Research Opportunities** 15
- Short Term 15
- Long Term 15

Lignocellulose Deconstruction 17
- **State of the Science** 17
- **Remaining Biological Research Challenges** 22
- Pretreatments 22
- Enzymes 23
- Microorganisms 24
- Analytical Tools 25
- Biotechnological Tools 25
- Plants 26
- **Key Research Opportunities** 27
- Short Term 27
- Long Term 27

Specialty Fuels 29
- **State of the Science** 29
- **Remaining Biological Research Challenges** 32
- Predicting Process Feasibility—Picking Products, Pathway Designs, Yield, Selectivity, and Rate 32
- Increasing the Diversity of Microbial Platforms—Building and Testing Strains 34
- **Key Research Opportunities** 38
- Short Term 38
- Long Term 38

Bioproduct Development from Biomass 39
- **State of the Science** 39
- **Remaining Biological Research Challenges** 42
- Utilizing Bioproducts Resulting from Underutilized Byproduct Streams from Biofuels Production 42
- Developing Bioproducts from Sugar Streams Generated During Deconstruction 45

Key Research Opportunities 47
Short Term 47
Long Term 47
Summary and Conclusions 49
Appendices 53
Appendix 1. Workshop Agenda 54
Appendix 2. Workshop Scope 56
Appendix 3. Workshop Tasks, Charge Questions 57
Appendix 4. Workshop Participants 59
Appendix 5. Bibliography 61
Acronyms and Abbreviations ..... Inside back cover

Executive Summary

Multiple societal benefits underlie the U.S. Department of Energy's (DOE) support for a viable and sustainable domestic lignocellulosic advanced biofuels and bioproducts industry. These benefits include ensuring future energy security, lowering greenhouse gases to mitigate climate impacts, diversifying the range of available products, producing less toxic chemicals and byproducts, creating jobs in rural areas, and improving the trade balance. A DOE workshop sought ways to realize these benefits by accelerating the emergence of a robust, new cellulosic ethanol industry. The resulting report, *Breaking the Biological Barriers to Cellulosic Ethanol* (U.S. DOE 2006), outlined a path toward this future, emphasizing integrated research from feedstock development to conversion technologies.

Since then, DOE's Office of Biological and Environmental Research (BER), operating within the Office of Science, has supported transformational bioenergy research through the vertically integrated DOE Bioenergy Research Centers and development of biomass feedstocks and biofuels-relevant microbes. A number of important breakthroughs have resulted from this fundamental research and include the (1) demonstration that lignin composition and deposition can be genetically engineered to reduce plant cell wall recalcitrance without impacting plant viability; (2) development of effective pretreatments that can be adapted commercially to lower costs; (3) discoveries of novel microbes and enzymatic pathways for more efficient deconstruction of lignocellulosic biomass; (4) proof-of-concept research for consolidated bioprocessing (CBP; i.e., production of ethanol and other biofuels by naturally cellulolytic microbes); (5) metabolic engineering of microorganisms and plants for biological production of numerous advanced biofuels or their immediate precursors; and (6) identification of hundreds of new plant genes and developing an understanding of their role in cell wall biosynthesis.

To assess the current state of the science regarding lignocellulosic biofuels and identify remaining basic research challenges in establishing a viable domestic biofuels and bioproducts industry, BER convened the Bioenergy Workshop on June 23–24, 2014, in Washington, D.C. The workshop brought together 45 experts from industry, academia, and DOE national laboratories and included presentations and breakout discussions organized around the themes of (1) biomass development, (2) lignocellulose deconstruction, (3) specialty fuels, and (4) bioproduct development from biomass. Key workshop findings are summarized in the following paragraphs.

Biomass Development. Establishing a sustainable, lignocellulosic biomass-based bioeconomy will require a fundamental shift in how feedstocks are produced, processed, and transported to mills and biorefineries. Specific lignocellulosic biofuel crops are only now being deployed and tested in the field and have yet to be fully developed, unlike food crops that have been optimized over centuries of cultivation and breeding. A number of suitable biofuel

crop candidates—switchgrass, *Miscanthus*, energy cane, and poplar, to name the most prominent—are being improved for biofuel traits through a combination of natural variant selection, genotype-assisted breeding, and genetic engineering. Prioritized traits include reduced biomass recalcitrance, improved water and nutrient utilization, and delayed flowering. Recalcitrance of plant cell walls to conversion into biofuels and bioproducts remains a major challenge. This recalcitrance is dependent on cell wall structure, the synthesis of which likely involves more than a thousand genes and a majority of them remains poorly characterized or unidentified. Consequently, identifying these genes and determining their functions and the regulatory mechanisms responsible for their expression are a high priority for enabling rational engineering of biomass characteristics for advanced biofuels and bioproducts production. Understanding how the cell wall senses and responds to structural perturbations also is key to formulating strategies for improving biomass traits.

Lignocellulose Deconstruction. Current pretreatment techniques and materials include acid hydrolysis, alkaline wet oxidation, steam explosion, ammonia fiber expansion, organic solvents to solubilize lignin and hemicellulose, ionic liquids, sulfite, and ozone. None of these deconstruction methods is universally advantageous over the others, and they vary in their outcomes depending on the type of feedstock, downstream process configurations, and a variety of other factors. Several of these methods, however, show promise and can benefit from further development. Conversely, CBP, in which a single microbe or microbial consortium is used to deconstruct lignocellulose and convert it directly to product, may require no pretreatment. CBP has shown great potential with the utilization of new genetic engineering tools for natively cellulolytic thermophiles such as *Clostridium* sp. and *Caldicellulosiruptor* sp. Further research is needed to make deconstruction processes low cost, low energy, ecologically friendly, and capable of converting a range of lignocellulosic biomass types into hydrolysates that contain as much of the cellulosic or hemicellulosic sugars as possible for conversion into fuels and chemicals. Additionally, technologies are required to convert the relatively large fraction of carbon found in the lignin portion of lignocellulosic biomass into biofuels and chemicals. Stronger linkages between advances in biomass development and fuels production will strengthen these deconstruction efforts.

Specialty Fuels. Significant progress has been made in the development of tools for synthetic biology and metabolic engineering (U.S. DOE 2012). These advances have resulted in an expanded suite of accessible molecules beyond ethanol to potentially serve as biofuels. However, selecting appropriate target molecules based on meaningful evaluation of accessible markets remains a challenge. Furthermore, predictive modeling and integrated analysis capabilities are needed to reliably guide the development of new microbial strains for producing biofuel molecules or commodity chemicals, as well as associated process engineering. Strain development and optimization will be greatly accelerated with methods to test thousands of pathway variants in a high-throughput manner, significantly increasing the rate of discovery. Predicting a microbe's performance in industrial-scale fermentation based on benchtop-scale experiments also remains difficult, but improved analytical and modeling tools will facilitate such predictability.

Bioproduct Development from Biomass. The workshop's focus also included bioproduct development. This new focus recognizes the environmental benefits to be gleaned from producing chemicals from biomass, which currently are derived from petroleum, and the potential of the unbounded diversity of new molecules that could be produced from biomass. The synergies between the methods and approaches for fuel and bioproduct synthesis create an opportunity to leverage basic research in biofuels development with broader possibilities toward advancing a biobased economy (OSTP 2012). The most obvious cross-cutting technologies revolve around synthetic biology, metabolic engineering, strain optimization, and computational modeling (U.S. DOE 2012). For example, synthetic biology approaches enable the assembly of new pathways and reengineering of central metabolism to envision new suites of products that can be generated from biomass, yet the challenge of product selection noted for specialty fuels also applies to bioproducts. Although BER's focus is on biological processes, bioproduct development has an increased potential to incorporate novel thermochemical conversion methods for sugar and lignin transformation.

In 2014, a few lignocellulosic biorefineries came online in the United States. These first-generation biorefineries will serve as a testing ground for developing economic and agronomic models for an efficient and sustainable lignocellulosic advanced biofuels and bioproducts industry. Additionally, bioenergy research goals are shifting based on this progress and expanding from those established in 2006. These goals have matured from the economical production of lignocellulosic ethanol to the economical production of lignocellulosic advanced biofuels and bioproducts. Of particular interest is the potential for aromatic products derived from lignin because they offer an attractive alternative to petroleum-derived aromatic compounds; lignin-derived products use less toxic starting materials and potentially can be tailor-made by plants.

To date, much progress has been made in overcoming several barriers to the production of lignocellulosic biomass and its transformation to ethanol, and these successes can now be leveraged in the production of advanced biofuels and bioproducts. BER's integrative approach is uniquely well positioned to address the basic research challenges associated with the establishment of an economically competitive and sustainable domestic biofuels and bioproducts industry. Significant advances in plant breeding, molecular genetics, and genomic technologies provide new opportunities to build on existing knowledge of plant biology and more confidently predict and manipulate functional properties of biomass feedstock crops. Similarly, continuing advances in omics-enabled technologies and synthetic biology approaches for microorganisms provide opportunities to further develop nonmodel microorganisms for applications in industrial biotechnology and for conversion of biomass into biofuels and bioproducts. Most importantly, integrating plant and microbial systems biology research with cutting-edge research in chemical and process engineering, synthetic biology, and computational biology facilitates the kind of scientific breakthroughs needed to foster the development of a sustainable bioeconomy (OSTP 2012).

Future opportunities for basic research in support of a sustainable and commercially viable advanced biofuels and bioproducts industry include, but are not limited to:

- Gaining a fundamental understanding of plant biology to develop a broader set of biomass crops that are economically viable and environmentally sustainable over a range of geographically distinct field conditions.
- Determining the role of microbial interactions with plants in conferring resistance to abiotic and biotic stress and controlling nutrient availability.
- Defining robust, feedstock-agnostic pretreatment and separation systems to more efficiently deconstruct and separate plant biomass into its various components for more efficient downstream biofuels and bioproducts production.
- Developing broad metabolic engineering techniques to enhance production efficiency of advanced biofuels; leveraging these techniques to design new metabolic networks for concurrent production of bioproducts and specialty fuels from plant biomass.
- Developing new, broad-based genetic systems to access a greater diversity of microorganisms and plants for bioenergy purposes.
- Assembling computational biology tools and models to help glean understanding from complex plant and microbial datasets, formulate experimentally testable hypotheses, and aid biosystems designs for bioenergy purposes.

Introduction

In his January 2006 State of the Union address, President George W. Bush cited America's "addict[ion] to oil" and called for federal investment in renewable alternative energy, including cellulosic ethanol, to help alleviate this dependency. One month prior to this address, a group of scientists and engineers convened by the Department of Energy (DOE) participated in a workshop to outline the key technical challenges that needed to be overcome to enable the emergence of a robust new cellulosic biofuels industry. The resulting report, *Breaking the Biological Barriers to Cellulosic Ethanol* (U.S. DOE 2006), aimed to provide a basic research roadmap to accelerate the rise of cellulosic ethanol just as federally mandated targets for renewable fuels were being considered as part of the revised Renewable Fuel Standard (RFS) under the new Energy Independence and Security Act (EISA) legislation. This roadmap emphasized integrated research from feedstock development to conversion technologies and resulted in a funding opportunity announcement that culminated in the selection of three major science centers, the DOE Bioenergy Research Centers (BRCs; see sidebar, BER Bioenergy Assets, p. 2). The BRCs' ultimate goal is to better understand the biological mechanisms underlying biofuel production so that these mechanisms can be redesigned, improved, and used to develop novel, efficient bioenergy strategies that can be replicated on a mass scale.

In 2006, plant biomass recalcitrance was identified as the core barrier to cellulosic ethanol. There was an urgent need to understand the chemical and physical structures of plant cell walls, how they are synthesized, and, importantly, how they can be deconstructed (see Fig. 1. What is Lignocellulosic Biomass? p. 3). The basic research roadmap that emerged addressed the issue of plant biomass recalcitrance but also outlined several broad basic science goals for bioenergy research. These goals included (1) sustainable, effective, and economical methods for feedstock production, harvest, deconstruction, and conversion to ethanol; (2) creation of a new generation of energy crops with enhanced sustainability, yield, and composition; (3) research to advance the enzymatic breakdown of cellulosic biomass to its component 5- and 6-carbon sugars and lignin; (4) research on cofermentation of sugars to ethanol; (5) the consolidation and integration of processes to reduce costs, improve efficacy, and reduce generation of and sensitivity to inhibitors; and (6) research to improve overall yields and economic viability in biorefinery environments.

Today, as a result of the past 8 years of research, the scientific community has a much deeper understanding of plants, particularly cell wall composition and the effects of changing cell wall composition on plant physiology (Burton and Fincher 2014; Jung, Samac, and Sarath 2012). Additionally, researchers have gleaned much more insight into the chemical, enzymatic, and microbial deconstruction of plant cell walls, as well as an enabling understanding of how to engineer saccharolytic microbes. Other studies such as the *U.S. Billion Ton Update* (U.S. DOE 2011; USDA 2014) have provided important information looking at

BER Bioenergy Assets

The Department of Energy's (DOE) Office of Biological and Environmental Research (BER) supports basic research on microbes and plants to provide fundamental understanding needed for developing new bioenergy crops and improving biofuel production processes to be cost-effective and sustainable. Within the Office of Science, BER manages a bioenergy research portfolio that spans the DOE Bioenergy Research Centers (BRCs); national laboratory scientific focus areas (SFAs); and specific research programs on bioenergy feedstocks, microbial biofuel production, and sustainable bioenergy crop development. BER also supports enabling capabilities such as the DOE Systems Biology Knowledgebase (KBase) and national user facilities for advanced genome sequencing, interpretation, and structural analysis.

BRCs. DOE established three BRCs to focus the most advanced biotechnology-based resources on the biological challenges of biofuel production. Each center is pursuing basic research underpinning a range of high-risk, high-return biological solutions for bioenergy applications. The BRCs' ultimate goal is to better understand the biological mechanisms underlying biofuel production so that those mechanisms can be redesigned, improved, and used to develop novel, efficient bioenergy strategies that can be replicated on a mass scale.

DOE National Laboratory SFAs. BER supports biofuels research at DOE national laboratories through multidisciplinary, multiyear research projects. Research topics include dynamic visualization of lignocellulose degradation through the integration of neutron scattering imaging and computer simulation (Oak Ridge National Laboratory) with a systems biology approach to energy flow in hydrogen-producing microbial communities (Lawrence Livermore National Laboratory).

Plant Feedstock Genomics for Bioenergy. Integrating DOE's capabilities in genomic sequencing and biofuel production analysis with the U.S. Department of Agriculture's (USDA) long experience in crop improvement, DOE and USDA jointly fund projects to accelerate plant breeding programs and enhance bioenergy feedstocks.

Systems Biology–Enabled Research for Microbial Production of Advanced Biofuels. To harness the microbial world's biosynthetic processing power for advanced biofuels production, an expanded set of platform organisms is needed with appropriate metabolic capabilities and stress tolerance characteristics. This BER program's specific targets related to biofuel production are (1) promising new model organisms, (2) novel microbial functional capabilities and biosynthetic pathways along with strategies to overcome associated metabolic challenges resulting from pathway modification, and (3) novel analytical technologies or high-throughput screening approaches.

Systems Biology Research to Advance Sustainable Bioenergy Crop Development. To achieve reliable and sustainably high yields, bioenergy feedstocks must have the capacity to adapt and maintain productivity even in challenging environments (e.g., land that is less fertile, water stressed, and erosion prone). This BER program focuses on (1) systems-level research to better understand the molecular and physiological mechanisms that control bioenergy crop vigor, resource use efficiency, and resilience or adaptability to abiotic stress, as well as interactions with the surrounding environment, to increase biomass productivity under changing and sometimes suboptimal conditions; and (2) systems biology–enabled investigations into the role(s) of microbes and microbial communities in complex and multiscaled interactions of the plant-soil environment.

KBase. KBase is an open bioinformatics platform for predictive systems biology designed to accelerate understanding of microbes, microbial communities, and plants relevant to DOE's missions, including bioenergy. As a community resource, KBase's purpose is to integrate a wide spectrum of genomics and systems biology data, models, and bioinformatics tools to ultimately predict and design biological function. KBase allows researchers to collaboratively generate, test, and share new hypotheses about gene and protein functions; perform large-scale analyses on a scalable computing infrastructure; model interactions between relevant organisms; and propose new experiments to further refine the models.

DOE User Facilities. DOE's Office of Science creates, maintains, and operates state-of-the-art national user facilities that are key to continued U.S. leadership in physical and biological research. BER supports the *DOE Joint Genome Institute (JGI)*, which is one of the world's largest and most productive public genome-sequencing centers. JGI has sequenced numerous varieties of plants relevant to bioenergy production, as well as microbial communities with degradative abilities for processing plant biomass into advanced biofuels. BER also supports the *Environmental Molecular Sciences Laboratory (EMSL)*, which offers a powerful suite of instruments for characterizing biological organisms and molecules. Other user facilities supported by DOE's Office of Basic Energy Sciences include the *synchrotron light sources and neutron facilities*, which enable the tracking of biomolecular processes in real time and imaging of biological materials at atomic resolution. DOE's Office of Advanced Scientific Computing Research offers access to *supercomputing facilities and advanced scientific networks*.

broader biomass supply and process development issues needed to support a biofuels industry. These technical insights and scoping data are critical for developing a sustainable biofuels and bioproducts economy, yet several challenges remain.

Biomass recalcitrance is still the single most important factor impeding the development of low-cost biomass processing technologies. Recalcitrance directly impacts yield, and the basic scientific questions most relevant to the emergence of a cellulosic biofuels industry continue to revolve around increasing the yield of sugars from biomass, the concentration of these sugars in the fermentation medium, and the rate of enzymatic hydrolysis and fermentation processes.

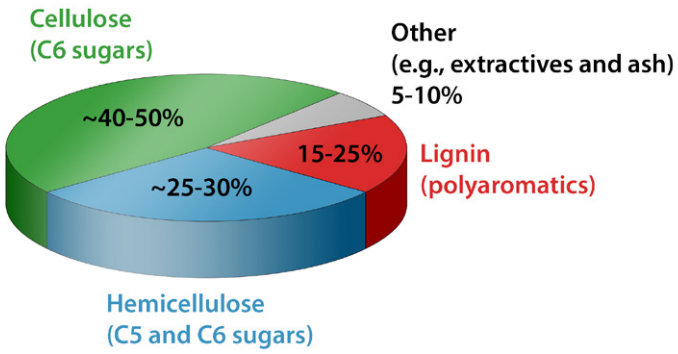


Fig. 1. What is Lignocellulosic Biomass? This pie chart represents the approximate distribution of the three primary components of plant cell walls—cellulose, hemicellulose, and lignin.

Multiple options for biomass feedstocks and bioprocessing exist today, and these options are relatively well understood, taking into account issues such as regional variation, relation between recalcitrance and cell wall structure, and feedstock-specific processing conditions that impact yield and cost. Because of the varied conditions under which biofuel crops will be cultivated, it is clear that there will be no single "ideal" feedstock, and thus basic research is continuing on multiple crop types. In the near term, coproducts of the feed and food enterprise, such as corn stover, serve as inputs for lignocellulosic ethanol fermentation, while dedicated bioenergy crops have yet to emerge at industrial scale.[1]

Pentose-fermenting yeasts are now in common use in laboratory research and pilot plant demonstrations. These yeasts will be a key technology in the emerging cellulosic fuels industry. As a result of progress over the last few years, researchers now envision biologically processing lignocellulosic biomass to fuels without either added enzymes or pretreatment, an outcome that would be truly transformative. With the first genetic systems for thermophilic, cellulolytic anaerobes having been developed and applied, the development of a new branch of biotechnology is anticipated based on techniques that enable rapid characterization and genetic manipulation of nonmodel microbes that have industrially valuable phenotypes. To ensure success, however, basic scientific research programs must integrate vertically from agronomy and plant breeding to molecular biology, from process and chemical engineering through metabolic engineering and chemistry, and from omics analytics to computational biology.

The roadmap anticipated that lignocellulosic ethanol produced at high rates and titers utilizing tolerant mesophilic and thermophilic organisms would be achieved, as well as organisms that, in fact, would enable efficient ethanol production in a single step—the so-called consolidated bioprocessing (CBP)

[1]"Liberty" switchgrass developed by the U.S. Department of Agriculture's Agricultural Research Service, which sponsored CenUSA Bioenergy, will be available to farmers in spring 2016.

concept. The development of strains with higher yields, an ability to predictively engineer ethanol tolerance, and fully predictive pathway models to enable model-driven design of cellular biocatalytic systems also were anticipated. CBP has nearly met these goals at the laboratory scale. Nevertheless, significant challenges remain before CBP reaches a state of commercial viability. Consequently, CBP approaches are still under basic development while also being actively pursued by industry.

Sustainability also is crucial to all aspects of a viable biofuels and bioproducts industry. Operated within the Office of Science, DOE's Office of Biological and Environmental Research (BER) recently held a workshop addressing how to build sustainability into the design of lignocellulosic biofuels, which requires a fundamental understanding of how biofuel crops interact with their environment—both biotic and abiotic—to affect sustainable outcomes. The resulting report, *Research for Sustainable Bioenergy: Linking Genomic and Ecosystem Sciences* (U.S. DOE 2014), describes multiple ways that recent advances in the genomic and other omic sciences can contribute to the knowledge needed to design sustainable systems. In particular, progress in plant genomics will enable the inclusion of sustainability traits in future feedstocks, and advances in genomics and metabolomics will allow insights into plant, microbe, and soil interactions that support plant productivity and vigor. Linking these advances to ecosystem science enables the use of systems biology in the fundamental design of sustainable biofuel systems.

As BER has supported advances in the basic science underlying lignocellulosic biofuels production, lignocellulosic ethanol plants using well-established technologies (several with large DOE grant support) have been anticipated since the early 2010s. The first bioconversion facility of more than 10 million gallons of annual capacity started up in 2013 in Italy (Beta Renewables). Others came online in 2013 in the United States (Fiberight and INEOS), with others following in 2014 in the United States (Abengoa, American Process, DuPont, and POET-DSM Advanced Biofuels) and in Brazil (GranBio). Additional facilities are expected in the near future (Peplow 2014). The initial biomass input for the U.S. plants is lignocellulosic biomass in the form of corn stover and hardwood, as well as mill, municipal, and yard wastes (see Table 1. Initial Production Capacity of U.S. Commercial-Scale Biorefineries, p. 5). Clearly, the capacities of these first commercial-scale cellulosic ethanol plants fall well short of the expectations of policymakers who set a target of 36 billion gallons of renewable fuels by 2022 under EISA.[2]

This production capacity represents a significant achievement for the industry, and the question now is how to accelerate the expansion of a lignocellulosic

[2]EISA included an RFS target of 36 billion U.S. gallons of biofuels by 2022, with a requirement that 21 billion gallons (58%) must be derived from non-cornstarch feedstocks. The U.S. Environmental Protection Agency's log of renewable fuel production in the RFS2 program shows 20,069 gallons of cellulosic ethanol produced in 2012, none in 2013 (or in the years prior to 2012), and 594,316 gallons produced as of December 8, 2014. [epa.gov/otaq/fuels/rfsdata/2014emts.htm]

biofuels and bioproducts industry. The startup of commercial demonstration projects summarized in Table 1 (this page) will drive further basic scientific research by identifying unanticipated challenges as new biological approaches are tested in these plants. Establishment of these commercial-scale plants will initiate a continuous cycle of technology improvements by informing biomass conversion science of additional challenges that must be overcome. This effort will help to prioritize areas of research—whether at the molecular, systems, or process levels.

Table 1. Initial Production Capacity of U.S. Commercial-Scale Biorefineries

Company	Location	Feedstock	Output	Annual Capacity (Millions of Gallons)*
Abengoa	Hugoton, KS	Agricultural residues, dedicated energy crops	Ethanol	25
American Process	Alpena, MI	Hardwood, mill waste	Ethanol	1
DuPont	Nevada, IA	Agricultural residues	Ethanol	30
Fiberight	Marion, IA	Municipal waste	Ethanol	6
INEOS	Vero Beach, FL	Yard waste, municipal waste	Ethanol	8
POET-DSM	Emmetsburg, IA	Agricultural residues	Ethanol	20

*Volumes as announced by the respective companies

In addition to DOE's investment in biofuels, the past decade has witnessed a tremendous level of private investment in the development of new biofuels processes that not only addressed the key barriers to cellulosic ethanol, but also "drop-in" biofuels that are compatible with existing engines. A few major examples include the $500 million Energy Biosciences Institute led by British Petroleum (BP) and located at the University of California-Berkeley and University of Illinois, and a broad range of established companies (e.g., DuPont, BP, and Monsanto) and private ventures totaling well over $1 billion and resulting in several public offerings (e.g., Amyris, Solazyme, Gevo, and Codexis).

Most of the gasoline now sold in the United States contains some ethanol. Gasoline with 10% ethanol content is referred to as E10. Currently, the U.S. market for E10 is saturated with corn and cane ethanol, and the E15 and E85 (gasoline with 15% and 85% ethanol content, respectively) markets have been slow to open up, limiting expansion of the bioethanol market. Thus, in the absence of policy or market incentives for more bioethanol, the focus shifts to the production of nonethanol biofuels. Such considerations also create opportunities for upgrading and converting biologically produced intermediates into finished products as well as hybrid biochemical-chemical processing options.

Determining how these opportunities might be realized—and, in particular, identifying the basic bioenergy science necessary to do so—was the primary driver for the Bioenergy Workshop held June 23–24, 2014, in Washington, D.C. The workshop included 45 participants from industry, academia, and DOE national laboratories, with goals to assess the state of the science regarding lignocellulosic-derived biofuels, identify remaining challenges that basic science can address, and explore the potential of bioproducts derived from biomass (see Appendices 1 to 4, beginning on p. 53).

The workshop included breakout groups in four topical areas: (1) biomass development, (2) lignocellulose deconstruction, (3) specialty fuels, and (4) bioproduct development from biomass. Each group was tasked with addressing a series of charge questions and also was encouraged to extend beyond this initial set of discussion topics to think deeply about both the specific focal area and cross-cutting issues that could be affected by other aspects of the commercial pipeline.

This report summarizes the workshop discussions, further articulates the bottlenecks and challenges that have inhibited commercial development of lignocellulosic biofuels and bioproducts, and proposes basic research strategies to address these challenges. Worth noting is that two of the breakout group areas, specialty fuels and bioproduct development, moved beyond the scope of the bioethanol roadmap (U.S. DOE 2006) and more toward development of capabilities that would support a broader-based bioeconomy (OSTP 2012). The goal of these two groups was to think beyond ethanol toward fuel molecules that either are more infrastructure compatible with gasoline or are suitable for alternative applications such as diesel and jet fuel and other bioproducts. Many of the same issues that confront cellulosic ethanol also will affect specialty fuels and bioproducts. One example is the availability of abundant, low-cost biomass-derived sugars and toxicity of inhibitory compounds present in hydrolysates to microbes and enzymes (Ximenes et al. 2011; Ximenes et al. 2010; Li et al. 2010; Kim et al. 2011; Kim et al. 2013). Thus, advances made in specialty fuels and bioproduct development can be applied to fuel production, adding significant revenues from the sale of value-added products in a "biorefinery" facility. For the purposes of the workshop and this report, a biorefinery is defined as a production facility that generates both fuels and nonfuel chemicals or materials for commercial use.

The report is structured in chapters that cover each of the four breakout groups. Each chapter describes the current state of the science in the specific area, followed by (1) a discussion of remaining scientific challenges that can be addressed by basic biological research and (2) key research opportunities. Throughout this report, an effort is made to distinguish between "short-term" (achievable within 5 years) and "long-term" (achievable beyond 5 years) objectives.

As the United States continues to develop a robust, thriving bioenergy research community, the research goals have shifted in some specifics since 2006. The overarching goal is to facilitate more economical production of lignocellulosic fuels, but this aim now extends beyond ethanol to fuel molecule targets and other bioproducts. This wider focus includes lessons learned and knowledge gained of fundamental molecular mechanisms of both biomass construction and deconstruction to capture value and generate revenue from coproducts that occur as a consequence of biofuels production. Bioproducts from lignocellulosic biomass have been added to the discussion because of their potential as carbon-neutral replacements for petroleum-derived chemicals and their potential to advance a broader-based bioeconomy (OSTP 2012). Past successes with the BRCs have shown that vertical integration of biological research and engineering is essential, and multidisciplinary teams are necessary to address the full scope of technical issues that confront lignocellulosic advanced biofuels and

bioproducts. Likewise, new research capabilities [e.g., new genetic engineering tools and faster high-throughput analytical tools (U.S. DOE 2012)] can contribute to accelerating the economical production of lignocellulosic advanced biofuels and bioproducts.

Biomass Development

State of the Science

In 2006 a lignocellulosic biofuels industry was merely a vision, but ongoing national discussions on energy policy led to the creation of the Energy Independence and Security Act (EISA) of 2007, which set renewable fuel targets and fostered new research toward development of a domestic lignocellulosic-based biofuels industry. Specific and readily achievable targets for ethanol production, first established under the Renewable Fuel Standard (RFS) as part of the Energy Policy Act of 2005, were revised and updated in the 2007 EISA legislation. These ethanol targets sparked the development of a corn-based ethanol industry that helped drive record U.S. corn production but also presented a conflict between the use of grain for food or fuel. In the 2006 bioethanol roadmap report (U.S. DOE 2006), several alternative feedstocks were identified as potential substitutes for maize grain provided a conversion technology could be implemented to convert lignocellulosic biomass to sugars and, ultimately, to ethanol. These feedstocks included perennial grasses, woody shrubs and trees, and row crop residues. Feedstock transportation and densification (e.g., pelletization) processes also contribute significantly to biofuel production costs, but regional solutions in feedstock choice need to reflect climatic and geographic variation (e.g., arid versus wet; see Fig. 2. Approximate Geographic Distribution of Potential Dedicated Biomass Crops, p. 10) as well as transportation and processing infrastructure (e.g., rail, road, and refinery access). Although algae and other aquatic plants were not discussed in the 2006 report, they also have become potentially important sources of biomass (U.S. DOE 2010).

Transitioning to a lignocellulosic-based bioeconomy will require a fundamental shift in methods of feedstock production, processing, and transport to mills and biorefineries. The costs of growing, harvesting, and storing these crops are significant (between $50 and $125 per dry ton; NAE 2010; U.S. DOE 2006; U.S. DOE 2011). These costs and the bulky, unstable nature of lignocellulosic biomass can constrain the scale of biorefineries. Additionally, regulatory constraints could apply, particularly if feedstocks are genetically modified to enhance their bioenergy properties. Basic science and engineering research is needed to overcome these remaining bottlenecks in biomass development, ranging from improving lignocellulosic biomass properties to increasing conversion efficiencies and lowering production costs.

Advances over the past several years have resulted in a deeper understanding of the impact of lignin and polysaccharide structure on recalcitrance and cellulose digestibility (see Fig. 3. Three-Dimensional Illustration of Lignocellulose Meshwork, p. 11). These accomplishments include determination of the genomic sequences of a number of bioenergy crops and the identification of genes involved in cell wall lignin and polysaccharide synthesis. Concurrently, major gains have been made in the molecular science needed to redirect lignin and polysaccharide synthesis. The impact on plant growth by these engineered forms can be either minimally negative or positive (Baxter et al. 2014), while maximizing sugar yields

upon enzyme hydrolysis after the plant is harvested and providing lignin streams more amenable to product manufacturing (Ragauskas et al. 2014). Such engineered plants potentially could reduce the costs of biomass conversion (Baxter et al. 2014; Bonawitz et al. 2014; Wilkerson et al. 2014; Yang et al. 2013).

As this document was written (fall 2014), POET-DSM's Liberty Biorefinery was the first domestic facility to have a projected capacity greater than 20 million gallons per year. The biomass being converted to ethanol at this facility is corn stover. Several concerns have been raised regarding the use of such plant residues as feedstocks, including the accelerated deterioration of soil quality through increased erosion and removal of nutrient-rich organic materials. At the same time, however, changes in agronomic practices—such as successive, multiyear corn-on-corn plantings, increased planting densities, and low- or no-till plantings—are starting to make excessive stover a land management challenge on many farms. Thus, there are clearly more and less sustainable feedstock production and logistical practices. These first-generation biorefineries and their surrounding logistic infrastructure will be instrumental in serving as a testing ground for developing economic and agronomic models for efficient and sustainable biofuels production, as well as industrializing many processes that will become more efficient with scale-up. This

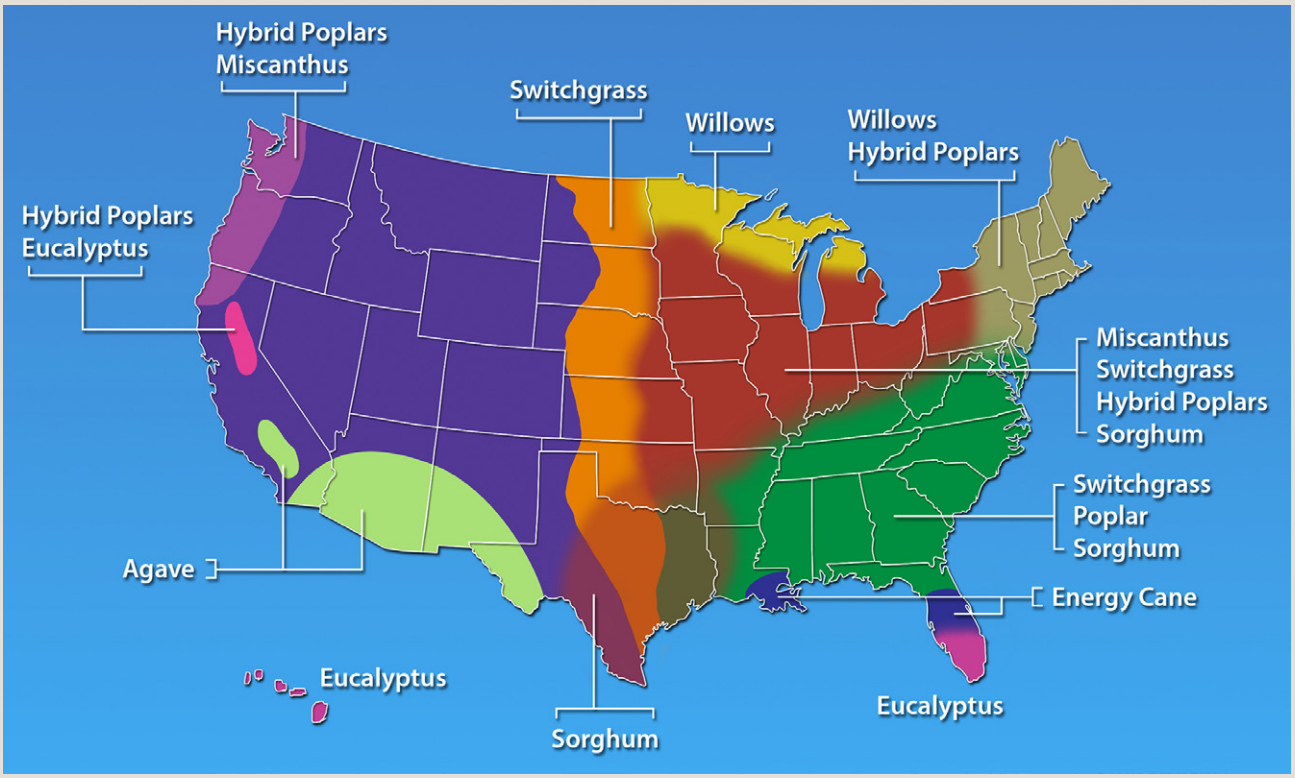


Fig. 2. Approximate Geographic Distribution of Potential Dedicated Biomass Crops. Multiple crop types designed for various agroecosystems will require continued development to realize biomass yields for large-scale production of biofuels and bioproducts. As research progresses, new crop types could be added and the boundaries of their likely ranges could change. Agricultural residues (e.g., wheat straw, rice hulls, and corn stover) are not included on this map.

development will include sustainable coproduction of lignocellulosic feedstocks on existing agricultural lands and the harvest, delivery, and storage of seasonally generated biomass materials for conversion facilities that must operate year round.

Remaining Biological Research Challenges

Following are some critical knowledge gaps impeding the utilization of lignocellulosic biomass for biofuels production, along with key research areas that can fill these gaps. This enhanced understanding will be expedited through the development of model systems that are genetically closely related to target feedstocks. Several model systems—such as *Setaria viridis* and *Brachypodium distachyon*, with their short generation times, small genomes, and robust transformation protocols—will enable detailed molecular characterizations of genes and pathways that are either not present in (e.g., C4 photosynthesis) or highly divergent from (e.g., cell wall structure) dicot models. These pathways must then be validated in biofuel crops.

Understanding Cell Wall Synthesis, Structure, and Function

Fermentation of sugars, whether derived from starch or sucrose, is a relatively robust and efficient industrial process. However, many challenges remain in developing a strong industrial process to deconstruct lignocellulosic biomass to liberate these sugars. Low-cost pretreatments and biological catalysts (as discussed in subsequent sections of this report) will solve part of this challenge; other solutions may be realized through the engineering of novel plant cell wall structures.

Over the past few years, dozens of genes involved in cell wall biosynthesis have been identified, but another thousand or more genes are likely involved, the majority of which are still unknown. Determining their functions is a high priority to enable rational engineering of ideal biomass characteristics for fuel production. Also not understood is how the cell wall senses and responds to structural perturbations—key knowledge for formulating strategies to modify biomass qualities. Recent exciting breakthroughs in designing modified cell walls include the Zip-lignins™, which incorporate a modified lignin monomer linkage that improves the lignin polymer's digestibility, greatly simplifying enzymatic digestion of cellulosic materials (Wilkerson et al. 2014). Extending this design principle to other primary and secondary cell wall polymers may provide additional avenues for improving plant cell wall

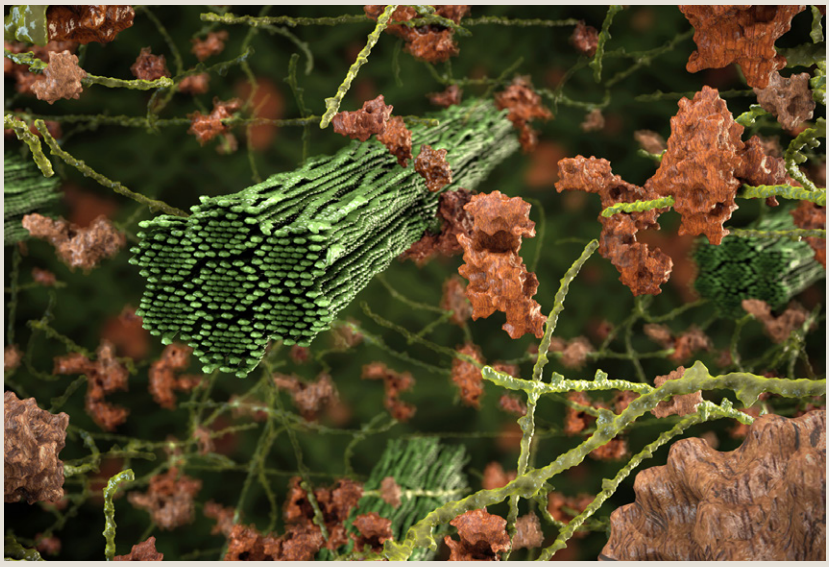


Fig. 3. Three-Dimensional Illustration of Lignocellulose Meshwork. Researchers are using computational modeling to gain a molecular-level understanding of the plant cell wall and its major components, including cellulose fibers (green), lignin molecules (brown wooden texture), and hemicellulose (light green). [Image courtesy Thomas Splettstoesser, www.scistyle.com, for Oak Ridge National Laboratory]

deconstruction. Additionally, changes in pectin polysaccharide structure resulted in positive effects on reducing recalcitrance and in promoting biomass productivity, something that likely would not have been discovered without multidisciplinary approaches. Results such as these highlight the importance of continued research on cell wall structure and biosynthesis to facilitate better understanding of biomass formation and quality for production of liquid fuels and bioproducts. Despite these advances, the recalcitrance of plant cell walls to conversion remains an important challenge.

Researchers also have made strides in generating functional genomic tools to analyze gene function *in vivo* using sequenced whole-genome mutant populations. However, the variety of plant species for which these population studies have been conducted needs to be expanded, especially in grasses. Furthermore, there still is little knowledge about regulation of the genes already discovered. Several transcription factors that directly bind to the promoters of cell wall biosynthesis genes are known (Gray, Caparrós-Ruiz, and Grotewold 2012), but the transcriptional regulatory hierarchy controlling the entire pathway remains enigmatic. For example, an exciting and unanticipated recent breakthrough was the discovery that the *Mediator* transcriptional regulatory complex specifically affects the functions of genes involved in lignin synthesis in *Arabidopsis*, showing that single transcription factors can regulate cell elongation and response to environmental stress (Bonawitz et al. 2014). These findings highlight the fact that important discoveries pertaining to biomass development and production will come from continued investments in basic plant research. Moreover, they indicate that fundamental knowledge gaps remain as to how biomass deposition is controlled, unknowns that must be understood to most efficaciously undertake rational engineering approaches to feedstock modification. Both biochemistry and nanometer- and tissue-scale composition and structure analyses are needed; other requirements are improved methods for rapid biochemical analyses at the appropriate scales to determine perturbations resulting from gene and regulatory changes, as well as validation of these changes.

Optimizing Carbon Delivery for Lignocellulose Production

From photosynthetic generation of sugars to their transport from the leaves and, ultimately, their partitioning to multiple sink tissues, the process of lignocellulose formation is dynamic and changes over time with plant development and in response to environmental stresses such as temperature, drought, or nutrient availability. How plants control assimilation, transport, and partitioning of carbohydrate resources from sources to sinks is not understood, but plants clearly actively adjust the rate of carbon delivery to sink tissues to maintain sufficient carbohydrate supplies for sustaining growth. Moreover, as sink tissues develop, they increase their rate of carbohydrate import to increase carbon storage (i.e., secondary cell walls in tree trunks and sucrose in sugarcane stems), but the mechanistic understanding of how sink strength is controlled and what limits cell wall synthesis is lacking. Additionally, sugars act as potent signals regulating plant gene expression, metabolism, and physiology, and they can feed back on both source and sink strengths to limit or promote further carbon accumulation

(Koch 1996; Roitsch 1999; Rolland, Baena-Gonzalez, and Sheen 2006; Bihmidine et al. 2013). Determining how plants modify and increase source and sink strengths and then engineering this process through integrated systems biology approaches will afford new opportunities for improving plant biomass quantity and quality. These strategies also will enhance efforts to sequester carbon (e.g., delivering more carbon to roots or tree trunks or depositing more carbon in the soil through root exudates) and will provide novel paths to engineer lignocellulosic feedstocks for enhanced biofuels production.

Increasing Biomass Yields

Heterosis (i.e., the phenomenon in which offspring of diverse parents have higher agronomic performance and yield than either parent) has been applied successfully in crop breeding efforts for food production, but its potential for breeding dedicated biofuel crops has not been sufficiently exploited. Additionally, many plants show increased growth and yield with higher genome ploidy levels, but the causal relationships between polyploidy and better biomass characteristics are not understood. Some plant composition alterations have shown significant negative effects on yield and field survival; other alterations targeting improved conversion, however, have shown beneficial impacts on growth and yield as well as minimal change in predation or disease. Insights into the molecular mechanisms underlying the impact of heterosis and polyploidy on yield will lead to improvements in biomass breeding strategies.

The yield potential of herbaceous perennials adapted to colder climates and a shorter growing season is limited by the rapid onset of flowering, seed production, and senescence, but biomass yield may be improved by breeding or engineering winter hardiness into high-yielding accessions normally grown in more temperate regions. Conversely, introducing delayed flowering into winter-hardy varieties may have a similar effect. The genetic basis of winter survival and cold tolerance in the overwintering crown and growing points and the integration of these traits at a molecular level with flowering and senescence are not well understood.

An ever important factor in plant biomass yield is the efficiency with which plants utilize the sun's radiant energy. This uptake is influenced by the biochemistry of light capture, as well as leaf and plant shape—as vegetation becomes denser, shading becomes a factor. Plant breeders must address such considerations as dedicated biomass crops are developed.

Another important challenge to improving biomass yield is understanding how environmental variability and plant responses influence biomass characteristics associated with quality. For example, identical plant genotypes grown in different locales have different biomass properties (Zalesny et al. 2009), but how and why these differences occur are not understood. Additionally, biomass quality is impacted greatly by environmental stress (e.g., flooding, drought, salinity, heat, and frost; Vasilakoglou et al. 2011). Finally, alterations in cell wall characteristics are known to affect pathogen infection. Hence, designing plants to be more resilient to environmental and disease stresses, thereby producing high-quality biomass in the face of such stress, remains an important goal.

Ensuring Sustainability

Over the long term, dedicated bioenergy feedstocks will drive the biobased economy, using sustainable agronomic practices that include low inputs and high biomass–accumulating annual or perennial crops grown on marginal or low-productivity fields (e.g., topographically difficult to cultivate, low water availability, or saline soils). Water inputs, for example, impact productivity, thus indicating the need for drought-tolerant plant varieties. To accelerate development of such feedstocks, several critical challenges must be addressed. One of these challenges involves perennialism (the seasonal cycling of vegetative development), a growth habit exhibited by many high biomass–accumulating crops such as sugarcane, *Miscanthus*, and switchgrass. A useful attribute of some perennial crops is the remobilization of aboveground nitrogen and phosphorus to rhizomes (underground stems) at the end of the growing season. The remaining aboveground plant matter consists largely of carbon skeletons that can be harvested, while most of the valuable nutrients are retained in underground organs. However, the signals involved in triggering resource repartitioning from leaves and stems to rhizomes, as well as the mechanisms by which the nutrients are remobilized at the start of the next growing season, remain largely unknown.

Another important attribute of perennial crops is their establishment of extensive and deep root systems that scavenge soil nutrients and water and alter the composition of soil microbial communities. These microbial communities (the microbiome) perform diverse roles providing plants with mechanisms for attaining micronutrients, protecting against soil-borne pathogens, increasing the effective surface area for water uptake, and helping to fix nitrogen into forms that plants can utilize. Microbial communities also play dynamic roles in carbon and nitrogen cycling. These attributes are extremely important for assessment and possible improvement of feedstock production systems, especially those in marginal areas. A deeper exploration and understanding of the diversity, functions, and dynamics of these microbial communities will not only help in the establishment and maintenance of more productive fields of bioenergy crops, but also may afford new opportunities for sequestering carbon and mitigating the adverse impacts of agriculture on natural resources and the environment. Important unresolved questions include determining how plants with designed or engineered properties impact the soil-plant microbial community, soil carbon, and plant biomass. Capabilities for manipulating and measuring both the plants and the microbial community are important for this understanding and the subsequently deployed improvements.

Improving biomass sustainability and curtailing biofuel costs also require better understanding of the impact of biomass removal on soil carbon and other nutrients, as well as the impacts of environmental stresses. Similarly, strategies to recycle mineral nutrients and water from the biorefinery back to the field will be needed to minimize crop inputs. Additionally, efforts to minimize the water content of biomass harvested from the field are needed to reduce feedstock transportation costs. Ultimately, the development of a sustainable bioenergy sector likely will come from the deployment of multiple feedstocks (see Fig. 2, p. 10) tailored to a region's breeding and cropping systems.

Key Research Opportunities

To enable scientific breakthroughs that overcome the challenges discussed in this chapter, research into multiple technologies is needed. Following are some technological innovations and enhancements that will greatly improve biomass development.

Short Term

- Molecular breeding tools to improve biomass germplasm using naturally existing or induced genetic variation [e.g., genome-wide association studies (GWAS), marker-assisted selection, new mutant populations, genome sequencing, and genome selection models]. Some of these tools are ready for deployment in several bioenergy feedstocks, including maize, poplar, sorghum, switchgrass, and pine, but further development will increase their effectiveness.
- Improved transformation technologies for nonmodel crop species to enable engineering of entire pathways or to stack multiple individual genes with specific functions.
- Genome-editing tools to precisely engineer desired genetic changes [e.g., clustered regularly interspaced short palindromic repeats (CRISPR) and CRISPR-associated protein 9 (Cas9), transcription activator-like effector nucleases (TALENs), and zinc fingers (ZNFs)].
- Synthetic biology tools for building new gene functions and networks to introduce novel traits into biomass crops (i.e., cell- and tissue-specific, inducible and developmental stage-specific promoters, positive and negative feedback loops for controlling gene expression, and artificial minichromosomes as trait platforms).
- Rapid detailed biochemistry measurements of cell wall synthesis, including intermediates coupled with nanometer cell wall biochemistry measurements, to map suspected cell wall biosynthesis genes to biochemical function and to follow how pretreatment affects cell wall degradation.
- New techniques for outreach and communication about the use and containment of genetically modified species (Carter et. al. 2014).

Long Term

- Whole-genome predictive models that, with a high degree of accuracy, can identify genes working together in a particular biological process.
- Whole-genome functional validation tools (e.g., sequenced mutant populations) to facilitate rapid gene function analysis.
- Tools to define the full cell wall structure(s) and higher-level architecture that impart biomass recalcitrance and to identify the genes and proteins that synthesize these structures, including detailed chemical, biochemical, and physical analyses of native and reduced-recalcitrance biomass.
- New computational models to predict how genotypes will respond to different environments and the impacts of stacking multiple genes (on the order of dozens to hundreds, more than can be easily tested experimentally) on biomass characteristics, as well as physiological models, from photons to fuel, to identify bottlenecks that limit crop production.
- New tools to suppress self-incompatibility in grasses to significantly enhance capabilities for breeding perennial grasses.
- Field testing of potential bioenergy crops under environmentally relevant conditions across multiple geographic regions to assess viability and robustness.
- Enhanced understanding of microbial interactions with bioenergy crops to adapt to changing nutrient or environmental stresses.
- Shared efforts in controlled field trials, such as common gardens used in GWAS studies; rapid shared phenotyping techniques and data libraries.

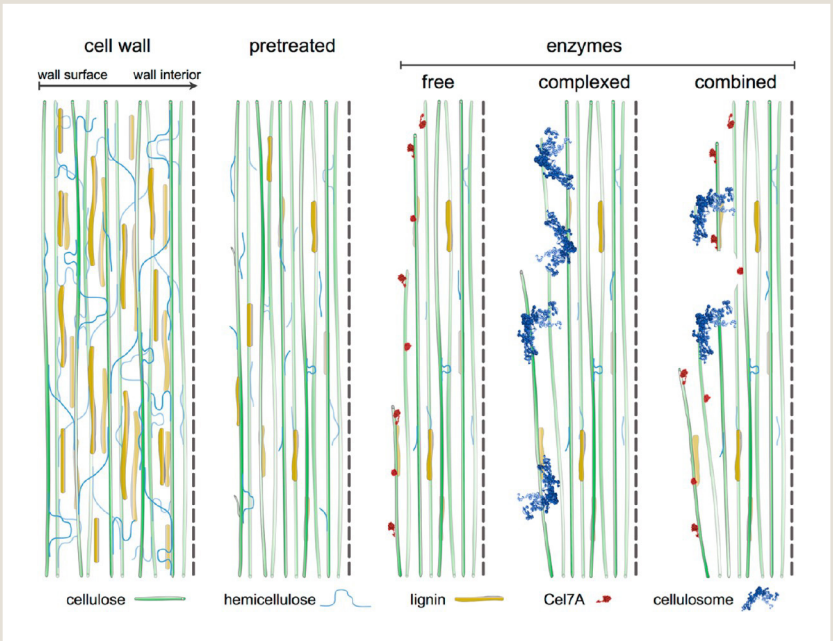
Lignocellulose Deconstruction

State of the Science

Deconstruction is the process by which the major components found in lignocellulosic feedstocks (i.e., cellulose, hemicellulose, and lignin) are converted into fermentable sugars and other desirable intermediate streams suitable for upgrading into finished products that can be sold in the marketplace (Chundawat and Beckham 2011). Through the biochemical routes currently envisioned, the deconstruction process represents the biggest cost in lignocellulosic biofuels production. As discussed in the Biomass Development section, p. 11, plant cell wall recalcitrance continues to be the major impediment to efficient and economic conversion of lignocellulosic feedstocks into sugars, fuels, and value-added coproducts and bioproducts. The key to overcoming recalcitrance still lies in the cell wall structure and using pretreatment to enhance its accessibility and susceptibility to deconstruction (see Fig. 4, Biomass Deconstruction, this page).

The conversion of lignocellulosic biomass into transportation fuels relies on the release of fermentable sugars by hydrolyzing the glycosidic bonds present in

Fig. 4. Biomass Deconstruction. Illustration of plant cell walls before and after pretreatment and models of hydrolysis by free (red) and complexed (blue) enzyme systems. Free enzymes with single carbohydrate-binding molecules and catalytic units hydrolyze cell wall polysaccharides by utilizing endoglucanases and cellobiohydrolases to react with accessible cellulose surfaces. Complexed enzymes with multiple catalytic and binding specificities likely have lower off-rates, and, once bound at multiple points of contact, disrupt the biomass surface, resulting in an increase in surface area. Combining these two enzyme paradigms on pretreated biomass synergistically deconstructs the cell walls by increasing the reactive surface area, allowing free enzymes to better diffuse and processively hydrolyze the substrate. Also, by removing the majority of the lignin and hemicellulose from the cellulose fraction, clean fractionation enhances the cellulose activity, enabling the benefits of combining these two deconstruction mechanisms. [Image reprinted with permission from Resch, M. G., et al. "Clean Fractionation Pretreatment Reduces Enzyme Loadings for Biomass Saccharification and Reveals the Mechanism of Free and Cellulosomal Enzyme Synergy," *ACS Sustainable Chemistry and Engineering* **2**(6), 1377–87. Copyright 2014 American Chemical Society]



cellulose and hemicellulose. Using enzyme mixtures for this process is an attractive approach, provided enzyme costs and reaction times can be reduced. These reductions require biomass pretreatment to enable enhanced substrate accessibility to enzymes, lignin removal, biomass fractionation, and perturbation of the microcrystalline domains found in cellulose (Blanch 2012).

Economic assessments of the advances in biomass deconstruction technologies can help quantify the impact of the different approaches, assist in prioritizing ongoing and future research and development activities (Tao et al. 2014; Klein-Marcuschamer et al. 2010), and identify strategies that promote effective plant biomass deconstruction leading to higher yields of fermentable sugars and value-added products. Some of the most powerful and informative work in this area to date has involved multiple institutions comparing different pretreatments on the same initial feedstock using the same protocols and analytical methods (Uppugundla et al. 2014). These efforts have generated new insights into the mechanisms and impacts of these pretreatments on the chemical and physical structures of lignocellulosic biomass and serve as a technical guide in the further development of advanced deconstruction technologies.

Several physical, chemical, microbial, and physicochemical methods have been developed to pretreat biomass prior to enzymatic hydrolysis, but each approach still has challenges that must be overcome. Physical methods include mechanical size reduction (comminution), such as via disk refining (Chen et al. 2013), and thermomechanical routes. Traditional chemical and physicochemical methods include acid or base addition at elevated temperatures. Dilute sulfuric acid (typically below 4%) is effective in breaking down and solubilizing hemicellulose, thereby providing improved accessibility for cellulose hydrolysis, but does not significantly perturb cellulose crystallinity (Lloyd and Wyman 2005). Another acid-based approach uses concentrated hydrochloric and sulfuric acids and anhydrous hydrogen chloride (HCl) gas to hydrolyze the holocellulose content of biomass. Crystalline cellulose is completely soluble in 72% sulfuric acid and in 42% hydrochloric acid. The lignin residue is removed by washing, and the HCl recovered by vacuum distillation. For any acid-based approach, there are concerns about spontaneous sugar dehydration into inhibitory compounds, acid neutralization and recovery, materials compatibility, process economics, range of biomass feedstocks that can be efficiently processed, and environmental impacts.

Alkaline pretreatments can be conducted at lower temperatures than those employed for acid pretreatments, with less degradation of sugars into inhibitory compounds, but pretreatment times are longer. The addition of alkaline swelling agents improves the accessibility of the biomass to enzymes. Mild swelling agents such as sodium hydroxide (NaOH), hydrazine, and anhydrous ammonia (NH_3) reduce cellulose crystallinity and increase enzyme accessibility (Kim and Day 2013). Alkaline agents are thought to saponify the uronic ester linkages of 4-*O*-methyl-D-glucuronic acids attached to the chain of xylan hemicelluloses, producing a charged carboxyl group and reducing cross-linking to lignin and other hemicelluloses. In the case of aqueous or gaseous NH_3 (commonly referred to as ammonia fiber expansion or AFEX™), ammonolysis of the same linkage produces an uncharged amide and consequently less swelling than with NaOH.

A recent variant on these techniques is extractive AFEX™, referred to as E-AFEX™, a process whereby lignin is extracted and cellulose I is converted to cellulose III. Similar to acid-based approaches, remaining challenges for the NH_3-based approaches are materials compatibility, environmental impacts, process economics, and range of biomass feedstocks that can be efficiently processed.

Other chemical pretreatment methods alter the lignin component of lignocellulose as well as the polysaccharides. Solvents such as ethanol or methanol, mixed with an aqueous inorganic acid catalyst, are very effective in delignifying biomass, and this pretreatment reduction in lignin content usually results in improved enzymatic conversion. Recently, certain ionic liquids (ILs) were shown to dissolve biomass, and the addition of an antisolvent (such as water) enables the cellulosic components to be precipitated prior to enzymatic hydrolysis (Brandt et al. 2013). A number of ILs are widely known for their ability to dissolve cellulose at high concentrations, but fewer can dissolve lignin and cellulose under similar processing conditions. Of the ILs able to dissolve lignocellulosic biomass, those based on imidazolium cations with halides, acetate, or (methyl) phosphates as the anions have been studied the most. Certain ILs are effective on a wide range of feedstocks, including those that have been pelletized to increase energy density. The resulting polysaccharide fraction is very susceptible to enzymatic hydrolysis. To be cost competitive with other pretreatment technologies, however, efficient IL recovery and recycling are required. Recent reports have shown lower-cost ILs, including those derived from hemicellulose and lignin that provide significant cost savings while maintaining performance. In addition to recovery and recycling, other remaining challenges include materials compatibility, process economics, and environmental impacts (Socha et al. 2014).

Historically, the type, number, and cost of glycoside hydrolase (GH) enzyme mixtures required for polysaccharide hydrolysis have been problematic for large-scale deployment at commercial biorefineries. Although significant and dramatic cost reductions have been obtained over a short period of time, GH enzyme mixtures remain one of the largest costs to biofuels production. This results from a combination of the enzyme loading necessary to achieve desired sugar-yield targets, persistence of recalcitrance linkages after pretreatment, and substrate accessibility. Moreover, the pretreatment method will define the types of activities needed during hydrolysis. For instance, in the case of dilute acid pretreatment, where the majority of hemicelluloses are converted into fermentable sugars, no hemicellulase activity is needed. However, after AFEX™ pretreatment, both cellulases and hemicellulases are needed to liberate fermentable sugars from the polysaccharide-pretreated solids. This process complexity and diversity inherently pose risks for demonstration and deployment activities.

Further improvements and discoveries also are needed to optimize and engineer enzymes capable of polysaccharide hydrolysis at desired operating conditions and environments. This is particularly true for breaking down lignin into smaller molecular weight intermediates suitable for upgrading. Currently, little is known outside of the major classes of fungal oxidative enzymes (Leonowicz et al. 2001), and there are no known lignolytic enzyme mixtures that can perform this function within the context of biorefining.

The chemical and physical state of the lignin after pretreatment and saccharification is also of interest and remains largely unknown. Saccharification is a time-consuming process that can take up to 7 days to obtain targeted yields, especially at low enzyme-loading levels, therefore requiring numerous large vessels. Increasing the enzyme loading to increase the saccharification rate may reduce the residence time and associated capital costs but increases operating costs. Several commercial enzyme mixtures are available, but the discovery and realization of more active enzyme mixtures (Smith et al. 2013), or reductions in their production costs through improvements in heterologous protein expression techniques, could generate significant cost savings. Developing enzymes and fermentative microorganisms that perform in the same conditions could aid in combining the saccharification and fermentation steps most effectively, without sacrificing residence times of the respective operations.

Research is being performed to optimize enzyme cocktails to curb product and substrate inhibition, engineer enzymes to be more robust under industrial operating conditions, optimize enzyme production, and discover new enzymes and enzyme complexes capable of more efficiently liberating fermentable sugars (see Fig. 5. Cellobiohydrolases Acting on Cellulose, this page). Currently, the majority of GHs are derived from and produced in fungi, notably *Trichoderma reesei*, which produces a suite of saccharolytic enzymes. However, developing other production hosts and a comprehensive toolbox that enables the expression of enzymes isolated from a wide range of environments and microbial origins remains problematic.

Cost-effective upscaling of pretreatment systems for commercial production also must be considered, given that a 20 million gallon-per-year plant will require 700 to 1,000 tons (dry basis) per day of biomass. Hence, a pretreatment approach that uses as little as 5% added reagents in a pretreatment system would translate to 35 to 50 tons per day of reagent that would need to be added, then properly recycled or disposed. Practical approaches will require that minimal added chemicals be used and that the chemicals applied be

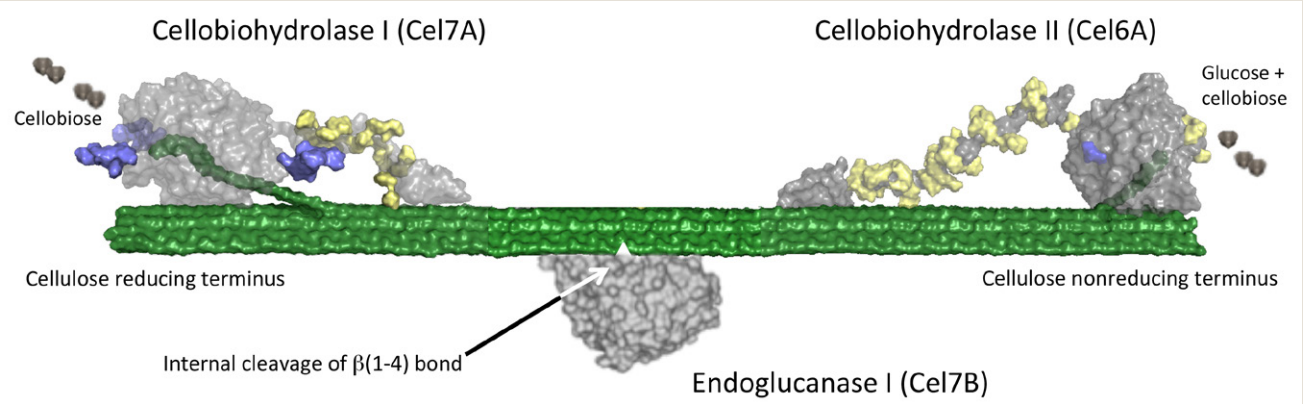


Fig. 5. Cellobiohydrolases Acting on Cellulose. Canonical depiction of cellobiohydrolases I and II (CEl7A and Cel6A) acting on cellulose. Enzyme binding, chain acquisition, chain translocation, cataysis, product expulsion, and recycling of these events are not well understood. [Image courtesy National Renewable Energy Laboratory]

approved for use through the Toxic Substances Control Act. Furthermore, the pretreatment system must be mechanically robust and capable of processing between 250,000 and 350,000 tons per year of biomass for a small-scale commercial demonstration facility. Since biomass can be shipped only relatively short distances (less than ~150 miles) to be cost-effective, capacity will determine both plant siting and geography by the type of biomass to be used (i.e., wood versus agricultural residues versus dedicated bioenergy crops). Feedstock processing (the first step shown in Fig. 6. General Biomass Processing Scheme, this page) also will contribute to the selection of plants on which to carry out fundamental research on their characteristics.

Once pretreatment is completed, the substrate is hydrolyzed by enzymes, and the resulting slurry may be introduced to fermentors that convert the formed sugars to biofuels, leaving behind a solid, lignin-rich residue. This residue is then separated and used as a boiler fuel to generate heat and electricity, both to provide energy for running the process and to sell as renewable electricity, because these residual solids generate more energy than is needed for operating the processing facility.

The making of ethanol from cellulose using thermochemical processes predates WWII, while the enzymatic saccharification of cellulosic biomass was first developed in the 1970s (Mandels, Hontz, and Nystrom 1974; Mandels, Andreotti,

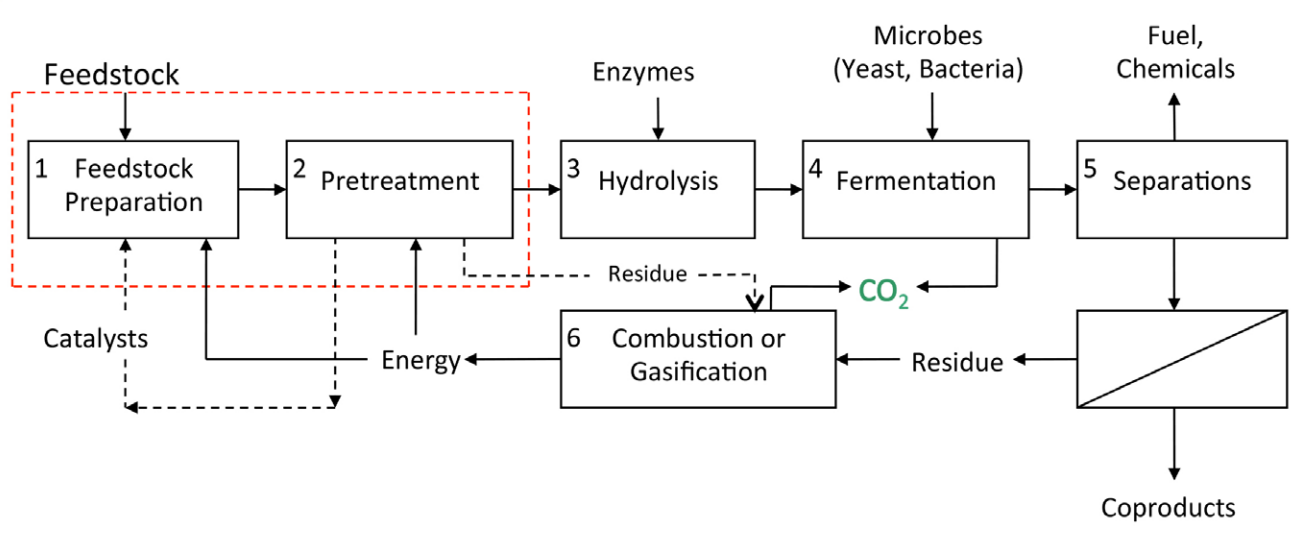


Fig. 6. General Biomass Processing Scheme. The basic unit operations in a biorefinery are (1) feedstock preparation, (2) pretreatment, (3) hydrolysis, (4) fermentation, and (5) separations (e.g., distillation or membrane separation and concentration of unfermented components by evaporation). Lignin is recovered either before or after fermentation and then gasified or combusted (6), making the process energy self-sufficient. Steps (3) and (4) are combined in processes where cellulose hydrolysis and fermentation occur simultaneously. Dilute acid, liquid hot water, and steam explosion pretreatments (2) result in process flows depicted by solid lines. Pulping processes and alkaline pretreatments extract a significant amount of lignin before the hydrolysis and fermentation steps [dotted line from step (2)], with pretreatment catalysts recycled to the beginning of the process (1). Lignin in excess of what is required for energy generation may be further processed into chemicals and drop-in biofuels. [Adapted from Ladisch et al. 2010. Reprinted with permission from the American Institute of Chemical Engineers Publication]

and Roche 1976). The bioprocessing route of converting cellulose to ethanol is conceptually simpler than for thermochemical processes and offers greater specificity. However, the specificity gained may be sacrificed in lower conversion rates. In the corn-to-ethanol industry, the introduction of TransFerm Yield+™ utilizes an advanced strain of yeast that expresses glucoamylase enzyme and reduces glycerol production. This consolidated bioprocessing (CBP) technology produced over 1 billion gallons of renewable fuels by December 2013, and industrial use has grown since then (Sapp 2013). Similarly, yeasts that produce enzymes for simultaneous hydrolysis of cellulose and hemicellulose and subsequent fermentation of the resulting glucose and xylose to ethanol have been engineered and the concept demonstrated at the pilot scale. Another approach to reduce operating costs (with enzymes being a major contributor) is to package hydrolytic enzymes in a proactive form in plant tissue or in the corn plant. These approaches all fit within the general biomass processing scheme (see Fig. 6, p. 21).

Remaining Biological Research Challenges

Adaptation of lignocellulosic processing technologies is likely to occur faster if processes can be developed that are low cost in terms of energy or reagents and produce value from as much of the biomass as possible. Thus, low-cost, low-energy technologies are needed to convert a suite of lignocellulosic biomass types (specific to different locations) into hydrolysates that contain as much of the cellulosic or hemicellulosic sugars as possible for conversion into fuels and chemicals. Technologies also are needed to convert the relatively large fraction of carbon found in the lignin portion of lignocellulosic biomass (25% or more) into biofuels and chemicals. Because of lignin's abundance and the difficulty in breaking it down into a form that can be processed into a product, lignin often is burned as fuel to power the processing facility.

Achieving these technologies will require further research advances in the areas of pretreatments, enzymes, microorganisms, analytical tools, biotechnological tools, and plants. Some of these knowledge gaps, as well as research opportunities to address them, are outlined in the remainder of this section.

Pretreatments

Effective substrate hydrolysis is dependent on pretreatments, whether by catalysts, heat, water, acids, bases, ILs, milling, or other means, all of which are designed to increase access and susceptibility of cellulose to enzyme hydrolysis (see Fig. 7. Effective Pretreatment for Optimal Enzyme Action, p. 23). Whereas these pretreatments improve hydrolysis, the inhibitors released from the inner parts of the cell wall structure, or through chemical modification of the lignin, will either deactivate or inhibit enzymes responsible for hydrolysis. At the same time, multiple enzymes are required for achieving full saccharification. Some plant materials, despite their generally equivalent composition of lignin, cellulose, and hemicellulose, have significantly different hydrolysis profiles. An example is corn stover, where the pith (soft inner part of the corn stalk) is much more readily hydrolyzed (even in the absence of pretreatment) than the rind, which has different physical

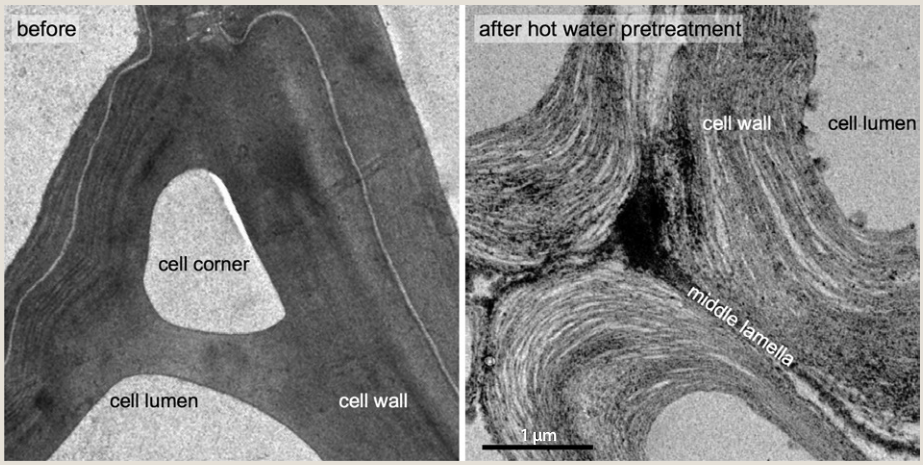
structures at the macroscopic, microscopic, and nanoscale levels. Consequently, a very important goal is to define the substrate and the enzyme systems that hydrolyze the substrate to gain a fundamental understanding of different pretreatment mechanisms, how pretreatments affect the substrate, and how the substrate affects the choice of pretreatment. While the substrate defines the range of conditions selected for investigation, the conditions should be among those that eventually can be used at the industrial scale.

Alternatively, development of feedstock-agnostic deconstruction technologies that can efficiently process a wide range of biomass feedstocks with minimal loss of performance would obviate the need for feedstock-specific processes. Furthermore, there is a growing recognition that the overall viability of any lignocellulosic biorefinery will be defined not only by its core biofuel products but by the coproducts it can generate. Needed is the development of a targeted and controllable biomass deconstruction and fractionation technology that enables downstream conversion of intermediates into finished products. This development will require an expansive knowledgebase of the impacts of different feedstock mixtures on performance and fractionation efficiency.

Enzymes

The sources and properties of enzymes used for polysaccharide hydrolysis may be fungal, bacterial, or formulations of multiple enzyme sources and types. To overcome recalcitrance and better understand cell wall deconstruction, a fundamental research goal is to understand how these enzymes carry out their functions of adsorption, decrystallization, and catalysis to hydrolyze both solid-phase and water-soluble substrates. This knowledge will help to define enzyme structures that achieve hydrolysis or prepare cell wall structures for deconstruction, as well as the role of accessory enzymes for effectively increasing cellulose hydrolysis. An example of an important new enzyme structure is the recently discovered

Fig. 7. Effective Pretreatment for Optimal Enzyme Action. Transmission electron micrographs of switchgrass (*Panicum virgatum*, Shawnee ecotype) cell walls are shown before and after a hot water pretreatment. The pretreated cell walls display extensive delamination throughout the secondary cell walls and evidence of lignin migration and coalescence into the cell corner. Such highly effective pretreatment hydrolyzes cross-linking hemicelluloses and provides a substantial increase in enzyme accessible surface area. [Unpublished data from Donohoe et al. 2011]



highly active cellulose from *Caldicellulosiruptor bescii*, a hot springs thermophile (Brunecky et al. 2013).

The interface between the soluble enzyme and the solid lignocellulose also is an area of important research. Of particular interest is understanding the manner in which the enzyme adsorbs to the surface and how the enzyme can be directed to attack areas of the cell wall while avoiding nonproductive binding at other areas. For GHs and ligninases, this understanding will require knowledge of protein structure and how different protein regions adsorb onto the surfaces that make up the cell wall.

Most of the attention on enzyme optimization and formulation has focused on GHs, with very little attention paid to lignolytic enzymes. However, if lignin is to be efficiently converted into other products beyond those created by burning lignin to generate waste heat and power, a lignolytic enzyme mixture may be needed to produce targeted intermediates suitable for upgrading. The current lack of understanding of lignin-degrading enzymes, especially for those enzymes not found in fungi, presents significant challenges and opportunities for the research community. This is especially true for the discovery and classification of lignin-degrading enzymes capable of breaking specific chemical bonds (e.g., etherases and esterases) needed to generate these targeted intermediates. Computational and bioinformatics tools relevant to lignin catabolism are nonexistent, and this gap must be addressed before the full potential of these systems can be realized in a biorefinery setting.

Despite some very impressive and isolated examples in the scientific literature (Kaul and Asano 2012; Bommarius, Blum, and Abrahamson 2011; Juturu and Wu 2012), another challenge is that current approaches to enzyme engineering and the prediction of amino acid changes to enhance biocatalysis and environmental stability remain elusive. More effort is needed to develop robust sequence- and structure-activity relationships that enable computational design of enzymes for targeted substrates and process environments. Furthermore, more mature computational models of enzyme-enzyme and enzyme-substrate interactions should be developed to generate multiscale information that would inform future enzyme-optimization activities.

Microorganisms

Microorganisms able to perform at industrial scale will be needed to carry out enzyme production and fermentation of biofuels and bioproducts. In some cases, enzyme production and bioproduct production may be separated into different processes as conditions for optimal productivity are different for these two processes. However, CBP could help cut cellulosic biorefinery costs by eliminating a dedicated process step for enzyme production (Sommer, Georgieva, and Ahring 2004; Lin et al. 2014; Yee et al. 2014). In nature, communities of mixed cultures are able to break down cell walls (e.g., in leaf-cutter ant colonies or rumens of cows and sheep) with conversions exceeding 50% of the structural carbohydrate. The use of microbial consortia for industrial production of bioproducts has not yet been studied to any great extent. However, the desire to maximize conversion

of the complex suite of organics present in lignocellulosic hydrolysates into fuels and coproducts provides opportunities for engineering microbial consortia. This could be accomplished either in a step-wise manner or simultaneously to mediate deconstruction of complex plant biomass and synthesis of advanced biofuel compounds and bioproducts. In addition, if photosynthetic or autotrophic microbes were part of these processes, they could be used to increase the carbon capture of these processes by capturing atmospheric carbon dioxide. Finally, technology is needed to assess the benefits of such systems operating in traditional batch, fed-batch, or continuous culture mode.

Basic microbiological research also is still needed to address inhibitor tolerance, which is complicated by the reality that different feedstocks, combined with different pretreatment methods, lead to variability in the inhibitor composition. This interdependence emphasizes the need for cross-cutting research projects organized across feedstocks, pretreatment, deconstruction, and conversion to understand the system as a whole.

In addition, the realization of robust hosts to produce recombinant lignocellulolytic enzymes at high titers (20 g/L) remains challenging because scientists have very few methodologies available to them in the field. The development of a robust and efficient toolbox that enables "protein expression on demand" in multiple microbes with desired expression levels, glycosylation, and environmental robustness is highly desired yet remains out of reach. The establishment of expansive, robust, accurate, and community-based libraries of lignocellulolytic enzymes from fungi and bacteria is needed before predictive approaches to enzyme design can reach their full potential. Methods to understand and then manipulate robustness also are needed to broaden the use of nonmodel microbes with unique useful traits.

Analytical Tools

One barrier to discovering and improving pretreatment strategies is the limitation in analytical tools that can characterize biomass structural properties and correlate them with pretreatments and enzyme hydrolysis. This limitation is true for both the structure of cell wall polysaccharides and lignin. Structures that are not effectively deconstructed by specific enzymes or microbes must be studied further to examine these residues and to understand how other combinations (e.g., biocatalytic and catalytic approaches) can be used. Strategies that promote effective plant biomass deconstruction leading to high-yield, value-added products should be identified. Moreover, computational tools that enable multiscale modeling of plant cell walls in deconstruction environments are highly desired because they would reduce the degree of experimentation needed.

Biotechnological Tools

Currently, genetic engineering can be applied only to a small fraction of the diverse microorganisms present in nature because most of them cannot be cultured and genetic transformation methodologies are nonexistent for all but a very few. Using approaches available today, development of transformation protocols for a

previously untransformed microbe relies on empirical approaches, often fails, and commonly requires an effort on the order of 5 person years or more. Reducing the effort required for such development by one to two orders of magnitude would be a revolutionary development. In particular, such a reduction would enable biotechnological applications to host organisms possessing phenotypes that are not practical to fully recapitulate in a chassis organism. Lignocellulose solubilization appears likely to be one such phenotype and represents an important proving ground for this emergent approach. What is needed is a widely applicable, likely bioinformatics-enabled, systematic approach to rapidly developing transformation systems. Development of such a general approach would be a major improvement over the current approach of transforming one microbe at a time.

Plants

Plant cell wall composition and architecture vary in different cell types as a function of species and developmental state. Optimal combinations of enzymes, pretreatments, and microbes for deconstructing biomass should be predictable from knowledge of the composition and architectural complexity of the specific biomass to be deconstructed. Although the major cell wall polymers are known, the nanoscale, molecular, and microscopic structures of the cell wall are not fully understood with respect to their role in attenuating or amplifying enzyme action. Complexity is further increased when pretreatments, required to make cell wall components accessible and susceptible to enzymes, change the physical structure, covalent cross-links, and noncovalent interactions in ways not yet predictable across different plant species (e.g., forbs, grasses, hardwoods, and softwoods), much less genetically modified variants of these species. Needed are fundamental cell wall composition studies that increase understanding of cell wall recalcitrance, dedicated investigations of the dynamic changes in cell wall structure that occur during deconstruction and pretreatment, and detailed studies of the pretreatment-enzyme-microbe mix used for deconstruction relative to biomass characteristics. These studies may include plant cell wall deconstruction at both the enzyme and microbe scales.

Stronger linkages between advances in biomass development and biofuels production also will strengthen deconstruction efforts. For example, this research should identify plant phenotypes that produce higher yields of fermentable sugars after deconstruction and higher yields of biofuels and coproducts after fermentation and upgrading (Wilkerson et al. 2014). These linkages can be realized only if research groups working in the various scientific fields collaborate to create a genotype-to-phenotype-to-intermediate-to-product knowledgebase that is accessible to the entire biofuels community.

Key Research Opportunities

Following are some important outstanding research opportunities that could significantly advance understanding for improving lignocellulose deconstruction.

Short Term

- Conduct experiments that measure the development of stage- and age-specific plant cell wall composition and architecture to inform the selection of pretreatment processes, including enzymes and microbes for feedstock-specific deconstruction of plant biomass.
- Define a mechanistic knowledgebase of deconstruction biochemistry to inform plant tissue structural studies. These studies would provide novel information about cell wall structure and aid in the identification, selection, and modification of enzymes and microorganisms to enhance their activities in the cell wall deconstruction process.
- Develop dynamic information flowpaths to enable better deconstruction and fundamental understanding of the process and basis for recalcitrance.
- Target bioinformatics studies of the expanding sequence and structure databases for new GH structures and families. New studies should include active-site and binding-site modeling capabilities.
- Develop pretreatments capable of efficiently fractionating biomass into targeted output streams with minimal inhibitor formation.
- Establish a basic toolbox for lignin catabolism to enable conversion of lignin into valuable products.
- Improve more general tools for manipulating nonmodel microbial isolates with unique traits to engineer additional traits as needed.

Long Term

- Improve cell wall deconstruction and enhance both yields and rates of processes involved—defined through pretreatments, development of robust lignocellulolytic enzyme mixtures, and discovery and genetic optimization of lignocellulolytic microorganisms; engineer plant cell wall synthesis to enable deconstruction, which can occur at the plant molecular biology level (Ragauskas et al. 2014). Subject to regulatory approval and constraints, these cell walls would be designed for facile deconstruction when exposed to processing conditions.
- Develop a feedstock agnostic deconstruction process that can efficiently convert a wide range of biomass feedstocks into targeted intermediates with equivocal performance to specific feedstock-converting technologies.
- Use multiscale modeling of plant cell walls in deconstruction environments, including bridging the gap between molecular dynamics and coarse-grained and finite-element mathematical models.
- Develop robust sequence- and structure-activity relationships for lignocellulolytic enzymes that enable predictive engineering for targeted substrates and environments.
- Develop preprocessing and pretreatment of biomass for recovering or generating value-added products. In this case, value-added products would be the coproducts derived at the same time or in the same process in which structural carbohydrates are converted to fuels.
- Articulate, gain supporting evidence for, and, ultimately, demonstrate one or more widely applicable, likely bioinformatics-enabled, systematic approaches to achieve rapid genetic transformation of nonmodel microbes (e.g., environmental isolates).
- Upgrade lignin to a better fuel or deconstruct lignin into its individual aromatic components, although lignin does not have the same defined structure as cellulose or hemicellulose. Significant effort will be needed to define the science that would enable the upgrade of lignin to a better fuel (i.e., diesel fuel) so that more of the carbon input from renewable biomass is captured as a marketable product.

Specialty Fuels

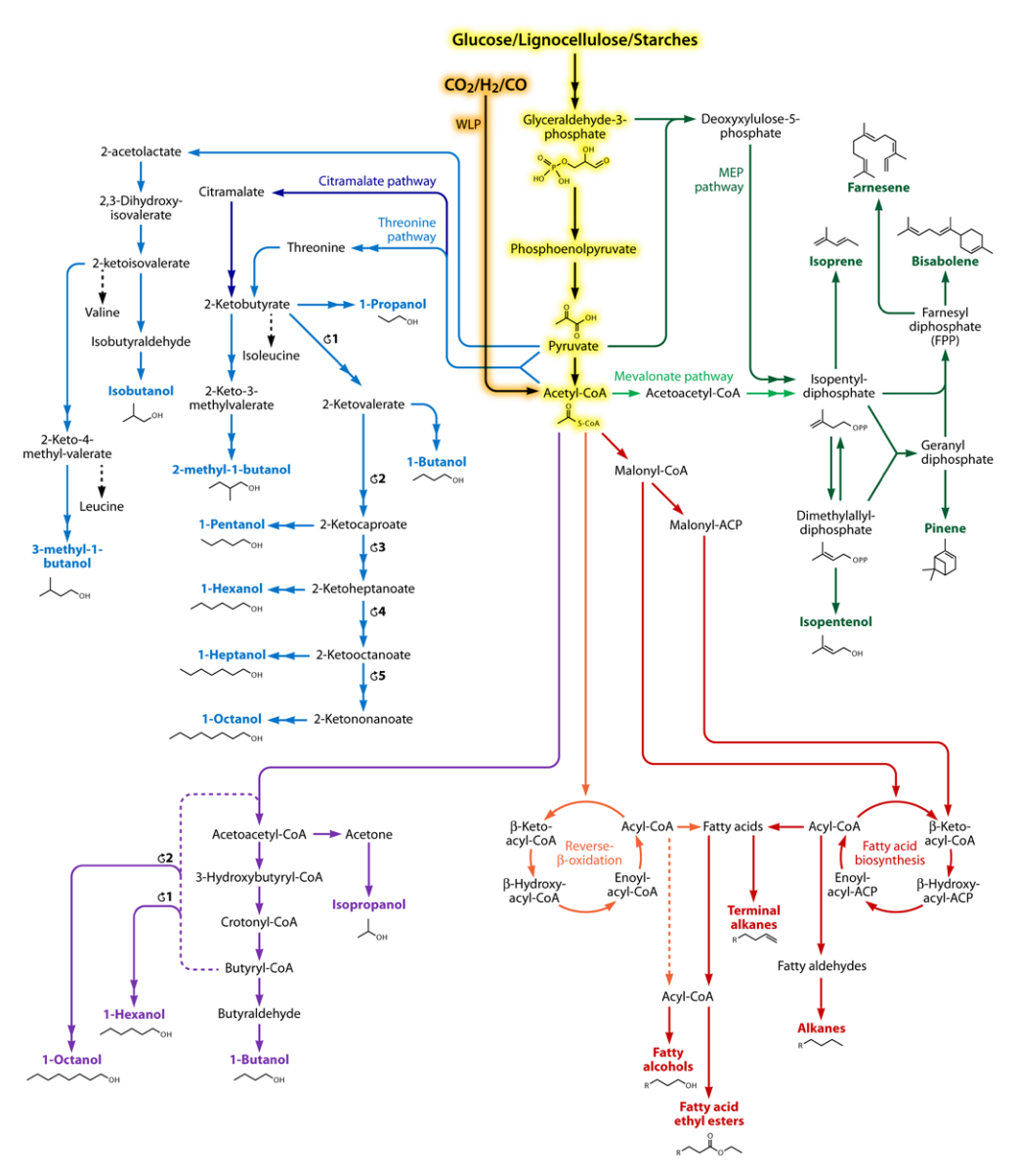
State of the Science

Specialty fuels are defined here as energy-dense fuel molecules other than ethanol, which is now a well-established biofuel. In the *Breaking the Biological Barriers to Cellulosic Ethanol* report (U.S. DOE 2006), ethanol was a dominant biofuel molecule, especially in the context of its production from lignocellulosic substrates. Butanol was not mentioned. In fact, the scientific literature (Ragauskas et al. 2006) did not discuss any specific biofuels other than ethanol and plant-derived biodiesel or the biological synthesis pathways of other fuel molecules. It did, however, discuss catalytic conversion of biomass to alkane molecules, then a nascent field of research.

Normal butanol has been produced biologically based on acetone-butanol-ethanol (ABE) *Clostridium* fermentation since the early 1900s. The ABE fermentation process is widely celebrated for its impact on the production of explosives used in WWI and WWII, as well as on the development of the automotive industry (Green 2011). Although not widely employed in the Western world since the late 1950s, the molecular biology and metabolic engineering of solventogenic *Clostridium* organisms (notably *C. acetobutylicum*) was an active field of research a decade ago, with an emphasis on butanol production as a commodity chemical or fuel (Paredes, Alsaker, and Papoutsakis 2005). Since 2006, advances in clostridial genetics and process developments have rejuvenated industrial interest in ABE fermentation. Isobutanol also has been produced at the pilot scale using engineered model microbes (Atsumi, Li, and Liao 2009).

Simultaneously, a significant amount of research has been conducted on alternative natural and semisynthetic pathways to produce candidate biofuel molecules. These higher energy dense molecules generally are intended to serve as "drop-in" biofuels that are compatible with existing engines, so they can compete with petroleum-derived gasoline, jet, and diesel formulations or be blended with existing fuels. The core pathways (see Fig. 8. Biosynthesis Pathways of Fuels and Related Chemicals, p. 30) utilized in these efforts include (1) the so-called reverse beta-oxidation route that mimics n-butanol synthesis and originates from acetyl coenzyme A (acetyl-CoA; Dellomonaco et al. 2011), (2) fusel alcohols derived from amino acid metabolism (Hazelwood et al. 2008), (3) isoprenoids derived from the mevalonate or nonmevalonate pathways (Peralta-Yahya and Keasling 2010; Kung, Runguphan, and Keasling 2012; Gronenberg, Marcheschi, and Liao 2013), and (4) fatty acids (Peralta-Yahya and Keasling 2010). The latter also have been used as precursors to deformylation reactions that generate alkanes (Schirmer et al. 2010). A 2004 report funded by the Department of Energy (DOE; U.S. DOE 2004) identified the 12 top value-added chemicals that could be produced from biomass. In light of the latest advances in conversion of biomass to biofuels, the time may be right to revisit this list, considering the value in dual-use chemicals such as isobutanol that can be used as fuels or precursors for other bioproducts.

Fig. 8. Biosynthesis Pathways of Fuels and Related Chemicals. Some key pathways leading to the formation of fuel molecules and related chemicals are depicted, starting with sugars (e.g., glucose), lignocellulose, starches, and waste gases [e.g., carbon dioxide (CO_2), hydrogen (H_2), and carbon monoxide (CO)]. The **yellow**-highlighted pathway is based on glycolysis. Use of $CO_2/H_2/CO$ is based on the **orange**-highlighted anaerobic Wood-Ljungdahl pathway (WLP) of acetogens, which can operate in tandem with glycolysis or similar pathways for utilization of sugars. **(Left)** The **blue** and **purple** pathways represent chemistries for the production of short- and medium-chain length alcohols. The alpha-keto acid pathways **(blue)** are based on typical native microbial amino acid biosynthesis routes, whereby decarboxylation and reduction (via the Ehrlich pathway) of the keto-acid intermediates of the valine, leucine, isoleucine, or threonine pathways (or the heterologous citramalate pathway) are employed. **(Lower left, purple)** Essentially, this is the clostridial pathway for the production of butyrate, butanol, acetone, and isopropanol, with the extension reactions possibly giving rise to 1-hexanol and 1-octanol. **(Right)** The **green**, **red**, and **orange** pathways represent chemistries for the production of alkanes and other hydrocarbons via fatty acid biosynthesis, the mevalonate (MEV) pathway, and/or 1-deoxy-D-xylulose-5-phosphate (DXP) or 2-C-methyl-D-erythritol-4-phosphate (MEP) pathways. Isoprenoids **(green)** form by successive condensation of the 5-carbon isomers isopentenyl-pyrophosphate (IPP) and dimethyl-allyl pyrophosphate (DMAP). IPP and DMAP are synthesized by either the plant MEV pathway or bacterial MEP pathway (or DXP pathway), employed by microbes such as *Escherichia coli* and cyanobacteria. Fatty acids **(red)** are synthesized from malonyl-coenzyme A (CoA) by fatty acid synthases. Reverse beta-oxidation was recently proposed as an alternative pathway that uses CoA as a carrier molecule and acetyl-CoA instead of malonyl-CoA to elongate the growing chain. Double-headed arrows indicate that multiple enzymatic steps are involved. [Adapted from figures 2 and 3 from Gronenberg, Marcheschi, and Liao 2013. Reprinted with permission from Elsevier]



The recent commissioning of cellulosic biofuel plants indicates that technologies are extant for commercial-scale production of ethanol or other compounds through fermentation of catabolic pathways. However, technologies to reduce the toxic effects of inhibitors in the hydrolysates (furans, aromatics, and chemicals used to solubilize biomass) or the fuels themselves (e.g., ethanol and isobutanol) on industrial microbes potentially could improve these fermentations and, thereby, further decrease the time, energy, or cost of conversion. In contrast, published papers on microbial production of biofuels (longer alcohols, fatty acids, and others) through anabolic pathways often report levels that are well below theoretical yields. Consequently, there is a need to increase the efficiency of microbial anabolic synthesis of fuels and chemicals, possibly under the low-dioxygen (O_2) or anaerobic conditions that often are desired in large-scale industrial fermentations, and to minimize the loss of reducing power derived from metabolism to a preferred electron acceptor like O_2.

Technologies available to engineer microbial systems also have progressed substantially in the past decade. In 2006, when synthetic and systems biology were in their infancy, there was a lack of information and molecular parts needed to execute more than simple attempts at engineering microbial pathways or networks. DNA sequencing technologies have advanced radically since then, with sequencing productivity doubling about every 6 months, and now can quickly and accurately sequence and assemble whole genomes. Additionally, automated annotation software has reached a level of speed and precision that enables quick and inexpensive basic genome annotation. Technologies also enable the identification of genome-level methylation data (Gonzalez et al. 2014; Clark et al. 2012) that can be used to identify and overcome restriction systems for developing effective transformation protocols and genetic tools. Today, sequencing, annotation, and even DNA methylation can be viewed as nonlimiting, readily accessible, and inexpensive enabling capabilities. DNA synthesis also has improved, but at a less exceptional pace.

As a result, it is now routine to sequence entire genomes, assess laboratory evolution experiments at depths of hundreds of millions of reads, and perform microbiome analyses of thousands of samples in multiplex. The on chip–based synthesis methods enable the production of millions of sequence-specific oligomers (~100 to 200 nucleotides in length), which can be pieced together to assemble increasingly large pieces of DNA as demonstrated by the synthesis and activation of the first complete microbial genome in 2010 (Gibson et al. 2008; Gibson et al. 2009; Gibson et al. 2010). Moreover, these short oligomers also can be used to dramatically rewrite microbial genomes when combined with editing technologies such as phage-based recombineering or clustered regularly interspersed short palindromic repeats (CRISPR), among others (Thomason et al. 2007; Mali, Esvelt, and Church 2013). For example, these approaches were used to construct an *Escherichia coli* strain with all 321 instances of the UAG codon removed (Lajoie et al. 2013)—a massive genetic engineering accomplishment that was not possible a decade ago on any reasonable timescale. At the most basic level, as capabilities for reading and writing DNA become even more robust, they no longer will be the rate-limiting step in microbial engineering

efforts for model organisms. Generally applicable genetic tools and methods for nonmodel microbes with unique complex traits (e.g., hydrolysis, robustness, or tolerance) are still needed, but the core challenge lies in genome design and, more generally, in the lack of understanding of genome and metabolic network structures and processes for design automation (U.S. DOE 2012).

Remaining Biological Research Challenges

Predicting Process Feasibility—Picking Products, Pathway Designs, Yield, Selectivity, and Rate

Several issues confound progress in bringing new products and processes to the industrial setting. While many advances have been made at the academic level in identifying potential products (chemicals as fuel molecules or commodity or specialty chemicals) and pathways to synthesize these products or their precursors, the industrial success of these processes still faces major challenges.

First is the choice of the product to be targeted for production using a biological process or a combination of biological and nonbiological processes. *Top Value Added Chemicals from Biomass* (U.S. DOE 2004) and several other scientific publications (USDA 2014) have listed a large number of target molecules as fuels or commodity chemicals, but these listings may be missing important molecules, especially molecules that can be produced by combining biological with nonbiological processes. One example is isobutanol, which until recently was a small-market chemical. The market for isobutanol expanded only after processes were developed to make other products from isobutanol. Such a market expansion is difficult to predict, making the rational and systematic identification of other target molecules challenging. Consequently, identification of additional target molecules will continue to rest with individual companies and investigators in an *ad hoc* manner.

Second, assuming that target products have been identified, actual industrial production requires cost-effective processes that make such production a profitable proposition. These challenges derive from the inability to achieve the important process characteristics or metrics necessary for industrial-scale success. Although many parameters will affect the industrial viability of various processes, the most important ones include:

- Ability to use inexpensive substrates at a high rate.
- Higher product formation rates, titers, yields, and selectivity [i.e., the ability to produce only the desirable product without byproducts (Dale and Ong 2012)].
- Ease of continuous or semicontinuous operation.
- Process integration with separation technologies for post-fermentation product purification.
- Industrial experience with the chosen production organism.

Significantly, it is the integration of all these parameters, several of which are interdependent, that makes predicting the likely success of a process very

difficult. What combination of these parameters must be advanced, and to what extent would it lead to industrial feasibility? To solve this multiparametric problem, robust algorithms or models are needed at various levels and scales to generate robust technoeconomical models that will provide the desirable answers. Examples of these model types include:

1. *Models to predict optimal pathways for production in common hosts (e.g.,* E. coli *and yeast cells) and thus to subsequently use thermodynamic and stoichiometric principles for calculating theoretical yields.* Much progress has been made over the last decade, particularly in the development of robust algorithms for predicting virtually all biologically possible pathways for synthesizing a potential molecule from common or rare biological metabolites and intermediates. For example, the Biochemical Network Integrated Computational Explorer (BNICE; Hatzimanikatis et al. 2005) framework is a systematic formulation of enzyme reaction rules based on the 1999 Enzyme Commission (EC) system. More integration of these model types is needed.
2. *Downstream processing models for estimating separation costs.* Needed is a framework for a comprehensive technoeconomical evaluation of a chosen candidate product and process. As an example, the Aspen models (www.aspentech.com; produced by DOE's National Renewable Energy Laboratory) and other simulation programs can handle traditional separation processes (e.g., distillation and traditional extractions), but these capabilities need to be expanded to include more recently proposed separation processes (e.g., *in situ* processes that combine fermentation with separations, or even more general processes based on the target molecule's thermodynamic principles and chemical properties).
3. *Fermentation process models, including approximate kinetic models that incorporate product inhibition and other off-target effects that result in productivity loss.* Needed are the development and validation of core fermentation process models that can reasonably simulate and predict real fermentation processes and thus be used for overall process simulation and technoeconomical studies. Such models should include all practical variations of fermentor operations (Zingaro, Nicolau, and Papoutsakis 2013): batch, fed-batch and repeated fed-batch, repeated batch with cell recycle (e.g., the Melle-Boinot process for yeast-based ethanol production), and continuous bioreactors (e.g., a series of chemostats with or without cell recycle). Robust kinetic models also should be implemented to assess the impact of product formation rate on process economics.
4. *Other models to predict process feasibility.* Appropriate modeling platforms are needed to enable fast and efficient preliminary technoeconomical simulations for assessing whether a process for a chosen product has the potential to be industrially feasible and, at the same time, to define critical process parameters for developing an industrially feasible process. Such technoeconomical simulations are widely used in the development of chemical processes and employ widely used simulators such as Aspen or SuperPro Designer (www.intelligen.com/superpro_overview.html).

Such simulation capabilities should be made easily accessible to both academic and small-company scientists so that the decision-making process can be accelerated on a more even footage, leading to more rational and comprehensive funding choices. Such models and the overall process simulation for a sufficiently robust technoeconomical analysis then should be able to assess the impact on industrial feasibility of core strain and fermentation parameters (e.g., rate of formation, titer, yield, and selectivity) and on strain characteristics (e.g., spectrum of substrates that can be used, simultaneous use of multiple substrates, and substrate flexibility). These models will aid in making decisions on platform organisms that potentially could meet these requirements. The Aspen models have been widely used and provide excellent examples of publically produced tools.

Increasing the Diversity of Microbial Platforms—Building and Testing Strains

In most cases, the choice of an organism for developing a fermentation process to produce a biofuel molecule or commodity chemical is driven by familiarity of the microbial system or ease of doing genetic modifications. Consequently, most new processes and products have been based on *E. coli* or a yeast system (largely *Saccharomyces cerevisiae*). The set of microorganisms serving as microbial hosts for producing biofuels and commodity chemicals needs to be expanded (U.S. DOE 2011; U.S. DOE 2012). Rational approaches for choosing a host have been discussed (Fischer, Klein-Marcuschamer, and Stephanopoulos 2008). The ultimate goal for an industrial microbial process is a combination of high yield, rate, and titer. However, these goals are difficult to judge *a priori* for a potential host microorganism, and a globally accepted set of parameters for choosing a host may not be available. Nevertheless, key parameters for selection and development include:

1. *Broad and flexible substrate utilization capability.* Although the most important, this parameter seems to be frequently ignored as substrate costs for these types of processes exceed 60% of total production costs. It should, however, override most of the parameters that follow.
2. *General genetics system.* Required is a basic genome-engineering demonstration system: a genetic system that includes gene knockin and knockout technologies, as well as recombineering, and the ability to integrate large DNA pieces into the genome. This genome-engineering toolkit should be expanded to work in a much broader range of organisms than is accessible today. Technologies such as multiplex recombineering and CRISPR have been shown to work in a range of microbes (Boyle et al. 2013; van der Oost et al. 2014). Further development and optimization of these and similar tools are likely to result in a greatly expanded capability for engineering biofuels-relevant microbes (U.S. DOE 2011; U.S. DOE 2012). Moreover, as *in vitro* methods for DNA assembly continue to advance, efforts to take advantage of the availability of large collections (1000+) of longer DNA assemblies (10 to 20 kilobase pairs) should be pursued in such microbes. Collectively, these efforts should provide a broad capability for performing

rational engineering of proteins, pathways, and genomes at throughputs that are orders of magnitude greater than currently is possible.

3. *Industrial competence.* Host microbes must be robust in an industrial setting and adapt easily to fermentation technology. They must demonstrate high-tolerance characteristics to substrates, product(s), and general bioprocessing stresses, including those derived from *in situ* product removal, but also tolerance to oxidative stress, pH extremes, and temperature and pressure variations. A few groups of organisms possess such traits, but, even in these organisms, the tolerance phenotype may not be sufficiently robust for industrial bioprocessing. Thus, beyond selecting for tolerant hosts, robust strategies for developing tolerant platform organisms will be equally important. Much progress has been made over the last few years in tolerance engineering, but larger efforts are necessary to solve this critical challenge. These efforts will require systematic and sustained investments with ambitious targets and milestones. A range of new technologies now exists for quantitatively mapping and engineering complex phenotypes, which collectively are well suited to address robustness at a scale well beyond what was previously possible. Ideally, such understanding will result in the development of robust "pathways or complexes" that could be accessed and engineered in a manner similar to how biochemical pathways currently are accessed and engineered.

Initially, efforts should focus on identifying a priority list of organisms based on the core traits discussed previously to target organisms that possess at least one of these desirable traits. This list should be reviewed and revised regularly. Clearly, metabolic engineering and synthetic biology of host microbes that already have key desired traits will be essential to overcoming (1) the primary challenges of producing a desired product and (2) increasing yield, rate, and titer while overcoming inhibition. A further goal is to build a minimal set of genetic tools to make use of the priority list of organisms. Although some of these tools will be organism specific, others may take a global approach.

Genome engineering entails the writing and editing of complete microbial genomes to encode new functions with predictable performance. As discussed previously, advances in DNA synthesis, genome modification, and DNA sequencing technologies have enabled whole new approaches to protein and pathway engineering, which, in turn, now have advanced to the genome scale. Yet key challenges to predictable genome design and construction remain. Several of these challenges involve efficient protein-, pathway-, and genome-scale construction efforts that could be addressed in the near term, given sufficient research and development efforts. The ability to obtain low production levels (<g/L) of a broad range of compounds is nearing maturity. Similar capability gains should be pursued for engineering organisms and processes that result in high-level performance across not only titer (g/L), but also productivity (g/L/hour) and yield (g/g). Moreover, such performance metrics should increasingly map to production costs, and a readily accessible and flexible industry-standard model that allows for the consideration of sustainability metrics (e.g., carbon costs) would be useful. Concurrently, algorithms that map performance from cultures operating at the microliter scale up to the

benchtop fermentation scale of 1 to 10 liters would greatly improve the relevance of optimization efforts occurring at the microliter scale.

Individual protein design and engineering tools still occupy a rate-limiting step in genome-scale engineering. Good methods are available for searching and analyzing genome databases, but annotation remains an issue, and methods for predicting function from sequence information alone need further development. Moreover, while methods for designing proteins with new or improved functions have advanced considerably in the past decade, further improvements in speed, accuracy, automation, complexity, and breadth are required (Nivón et al. 2014). These advances are becoming crucial as design goals expand from individual proteins to metabolic pathways, where balancing of multiple proteins, metabolites, and cofactors further complicates the engineering challenge. Many examples of successful basic pathway engineering have been reported in the past decade [i.e., cloning a short (< 5 to 6 genes) heterologous pathway for proof-of-concept production levels (e.g., mg/L to g/L) of a target compound is now routine], but optimization of flux through such pathways remains a major challenge. The key underlying reasons include a lack of understanding of microbial metabolism complexity, which has prevented the development of complete and broadly applicable metabolic models, and a lack of orthogonal and flexible DNA "parts" that enable the precise rewiring and control of metabolism. As designs further expand from individual pathways to larger collections of pathways and biomolecular complexes that collectively operate at the genome scale, performance continues to be challenged by an inability to routinely engineer robustness, resulting from a limited understanding of the genetic basis of tolerance and toxicity phenotypes. New approaches for mapping genes to traits have been reported, but many have not yet been demonstrated in organisms beyond *E. coli* or *S. cerevisiae*, even though there is considerable interest and a rationale for pursuing biofuels production in a number of additional microbes.

As capabilities for constructing complex designs advance, so too must capabilities for mapping the performance of such designs. In terms of genome-scale metabolite measurements, conventional approaches remain limited in throughput, sensitivity, and flexibility relative to design and construction technologies. Consequently, they must rely on the use of strong selections or high-throughput screens to reduce measurement loads or simply characterize only a small number of specific designs in so-called "rational" engineering efforts. Such reliance too often becomes the rate-limiting step in strain- and metabolic-engineering efforts. Profiling metabolites from thousands of samples per day is now possible [e.g., matrix-assisted laser desorption/ionization (MALDI)], but further development and integration of such approaches into the strain-engineering pipeline are needed. While similar challenges exist at the level of genome-scale protein measurements, multiplex RNAseq approaches have advanced in conjunction with sequencing capabilities and now match the sensitivities and throughputs of modern design and construction approaches.

As each of these technologies is developed, the overall throughput of the strain-engineering cycle will increase dramatically. Research will be needed to incorporate sophisticated engineering strategies that cycle through the stages of

design, build, test, and learn (see Fig. 9. Design-Build-Test-Learn Cycle, this page). Such strategies have the potential to rapidly increase knowledge of biofuel-relevant traits as they enable direct and rapid testing of hundreds to thousands of specific hypotheses in parallel. This capability shifts the emphasis from pursuing only the most well considered hypotheses to considering how to obtain more benefit from pursuing many different hypotheses. In this light, near-term efforts should focus on machine learning and other computational and statistical approaches to data analysis.

In the longer term, continued growth in DNA sequencing and synthesis capabilities likely will radically change strategies for engineering microbes for the production of biofuels and chemicals (U.S. DOE 2012). Genome-scale synthesis and assembly methods will become accessible to a broad spectrum of institutions ranging from private companies to large academic and national laboratory settings. Furthermore, genome-scale design software that connects individual designers to DNA synthesis and assembly institutions likely also will become readily available. Additionally, technology platforms for full-scale genome engineering should be developed. Such platforms would be designed to incorporate advances cited previously, as well as new technologies for quantitatively and iteratively mapping design performance (e.g., MALDI-based mass spectrometry). Such platforms also should enable the design and engineering of increasingly complex designs based on the predictable linking of multiple biomolecular and metabolic modules. Such modules could provide capabilities for consuming different sugars, producing different molecules, or tolerating nonideal conditions, among others.

Further efforts should enable full-genome synthesis and transplantation into a broad variety of microbes. Genomes already can be designed on a computer and then assembled outside the target organism [i.e., using *E. coli* or yeast (Gibson et al. 2008; Gibson et al. 2009; Gibson et al. 2010)]. This capability enables a complete shift in strategies for engineering nonmodel organisms. Given a collection of readily available methods for complete genome transplantation, flexible and robust strategies for genome assembly and transplantation should be pursued by extending the genome-engineering technology platform (described previously) to a very broad collection of microbes. Moreover, as such technologies are developed, strategies also should be pursued to efficiently isolate and interrogate a broad range of microbes from interesting environments. Such microbes could serve as important new chassis for genome-engineering efforts or provide insights into novel biochemistries and robustness phenotypes.

Alternatively, engineered microbial consortia offer a different strategy for bringing together a collection of desired traits. However, microbial responses in "non-natural" process schemes are poorly understood and will require significant investigation in model and applied settings to understand how to engineer and maintain stable and productive consortia.

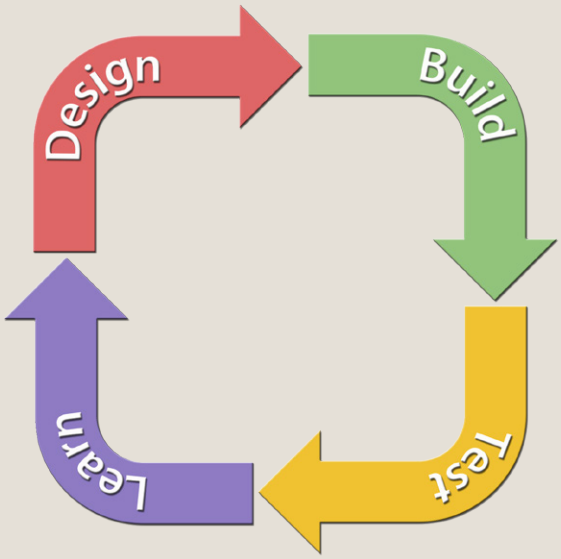


Fig. 9. Design-Build-Test-Learn Cycle. An iterative design strategy is based on the cyclic process of developing an initial design or prototype, testing that prototype, analyzing its performance against specific metrics, learning what worked and what did not work, designing a new protoype based on what was learned, and completing the cycle again. The goal is to improve design or prototype quality and functionality.

Key Research Opportunities

Goals for specialty fuels require a broad range of methods, models, and tools. Needs include the development of optimal genome-engineered models that can produce robust and reliable outcome predictions. Approaches, tools, and interfaces should be genome-scale, broadly applicable, flexible, accessible, and user friendly. Simulations should be able to assess impacts of capabilities, and applications should be scalable or conducive to future industrial use.

Short Term

- Develop generally applicable methods and genetic tools for use in nonmodel microbes so that process development is less dependent on well-established microbes and can more easily exploit nonmodel microbes that display unique or desired traits.
- Develop a priority list of 10 to 15 microorganisms with core required traits to be targeted for development into experimentally tractable systems. Necessary tools for organism development include genetic tools that enable single-gene knockins and knockouts, recombineering, and the ability to insert large DNA sequences into the genome, among others; metabolic models that fully articulate the metabolic complexity from genomic and metabolomic data; and fermentation behavior databases. Models that can predict behavior in scaled-up applications and yields will be important for industrial-scale adaptation.
- Develop high-throughput methods to screen or select high-performance strains or constructs to improve product formation rates, titers, yields, and selectivity (i.e., the ability to produce only the desired product without byproducts).
- Enhance microbial tolerance of toxins and thereby improve fermentation yields by developing a better understanding of the cellular and molecular bases of tolerance for each of the major chemical classes found in these processes. These chemical classes include alcohols, organic acids, aldehydes (furfural and hydroxymethylfurfural), and ketones (Mazumder et al. 2000; Nikolaou, Gaida, and Papoutsakis 2010; Lennen and Pfleger 2013).
- Establish standard assays and set metrics for assessing tolerance, including assays that assess production rates and titers in the presence of model toxic chemicals.
- Simplify the algorithmic complexity of models predicting optimal production pathways and develop user-friendly interfaces to make these capabilities more broadly accessible.
- Develop broadly applicable and flexible models based on state-of-the-art kinetic, stoichiometric, and thermodynamic principles, with integration of genome-scale models where possible, to enable robust and reliable estimations of fermentation outcomes and downstream separation outcomes; verify model reliability using multiple experimental case studies. These models should include options for utilizing various possible substrates alone, combined, or sequentially. Ultimately, such model simulations should be able to assess the impact of such capability options, enabling rational decisions on whether such capabilities should be pursued.
- Develop new techniques for outreach and communication about the use and containment of genetically modified species (Carter et. al. 2014).

Long Term

- Expand the priority list to 20 or 30 microorganisms to obtain a larger range of naturally occurring properties and develop tools for this expanded list of organisms, including a global toolkit for short-term goals 1, 3, and 4 [e.g., universal lambda-red system currently available for a few organisms (Poteete 2001)] that can be engineered into any and all organisms; build tolerant strains based on the knowledge gleaned from accomplishing the short-term goals.
- Develop advanced genome-engineering approaches for building tolerance to meet the previously stated metrics.
- Set targets to achieve production rates and titers of model toxic chemicals from each class that exceed the best possible values in 2014 by at least 25%.
- Expand the algorithmic capabilities of models to predict optimal production pathways and robust and reliable downstream separation outcomes. Ideally, the interface would be user friendly, and users should be able to specify mode of operation and organismal properties including product yields and selectivities and kinetic parameters.

Bioproduct Development from Biomass

State of the Science

The initial goal for biofuels production from biomass targeted the breakdown of lignocellulosic biomass into sugars that could be fermented to ethanol. As that technology has advanced, the concept of a biorefinery in which higher-value bioproducts are coproduced with biofuels has become increasingly important. These bioproducts can come from byproduct streams generated during biomass deconstruction or by creating added-value products from the generated sugar or lignin streams. A significant challenge, however, is determining the right target products, because they will need to have viable conversion and separation pathways from the feedstock stream selected, as well as a market appropriately sized to the feedstock stream being transformed.

Given the large number of potential bioproducts, applying comprehensive state of the science to all of them is difficult. Therefore, the range of bioproducts discussed in this workshop report is defined as products that can either directly or functionally replace petrochemicals or petroleum-derived materials.

Biofuels and bioproducts have a significant number of common technological challenges associated with their production, such as the research tools and methodologies needed for biocatalyst and chemical catalyst and separation technology development. However, they also have several important differences. Biofuels have outlets in three large markets: gasoline, jet fuel, and diesel fuel. The number of markets for potential bioproducts also is quite large, but the size of each market is significantly smaller than for fuels. Moreover, bioproducts generally require a specific chemical species with very high chemical purity. Fuels are less demanding in these two attributes. These differences impose technological challenges that are unique to bioproducts, particularly the fundamental research requirements for advancing bioproducts that can broadly enable a range of potential products.

The choice of target fuel or chemical molecules is critical since the economics and strategies of the chemical industry heavily depend on the availability of low-cost feedstocks and a market for the products that can be derived from these materials. In the past 8 years, major shifts have occurred in the feedstocks and products from the specialty chemicals sector. The newly expanding domestic fossil fuel supplies contain fewer aromatics and other chemicals than imported fossil fuels contain. At the same time, lignin, which currently is often a waste product, is now recognized as a potential source of these aromatics and other chemical compounds. This shift illustrates the need for genome-enabled technologies to assemble a catalog of robust microbial strains that can produce a suite of potential products from a relatively small and common set of strategic metabolic intermediates. The development of this technology will give industry the ability to adjust product types based on shifts in needs or product costs.

The general process concept for bioproducts from biomass byproduct streams was predicated on a separation step followed by or in conjunction with a chemical conversion step. For example, corn cobs were introduced to acidified aqueous reaction conditions that simultaneously released the C5 sugars and dehydrated them to furfural. This technology originally was developed in the United States, but the majority of furfural production now has moved to China. In the wood-processing industry, a key step in producing cellulose is removal of the lignin portion, which usually is burned for energy. Several companies are able to isolate this lignin stream in a form used for bioproducts. However, due to limited markets for these products and despite significant research being directed toward valorizing wood-derived lignin byproduct streams (Zakzeski et al. 2010), only a small portion of the lignin currently generated is used as a bioproduct.

Two commodity bioproducts that have been introduced, 1,3-propanediol (PDO) and lactic acid, were manufactured via fermentation using engineered microbes. Particularly in the case of PDO, a significant number of genetic interventions specific to the desired molecule were required to engineer the commercial microbe. In the case of lactate, the market for biologically produced lactic acid and its lactate-based polymers [e.g., the NatureWorks, LLC, technology (Vink et al. 2004)] took several years of development to reach the point of sustainable profitability. Lactic acid had no direct petrochemical competitor, but biologically derived PDO competed against a petrochemical-derived PDO and became the preferred process. These two bioproducts demonstrated that commodity chemicals produced from carbohydrates can compete economically with fossil carbon–derived chemicals.

Top Value Added Chemicals from Biomass (U.S. DOE 2004) proposed the use of "building blocks" in which carbohydrates are converted to intermediate platform chemicals, each of which subsequently could be converted to a range of biobased chemical products (see Fig. 10. Flowchart Comparing Potential Biomass- and Petroleum-Derived Products, p. 41). This building block concept is analogous to the petrochemical industry, which converts its intermediate molecules to petrochemical products (see Fig. 11. Products Made from a Barrel of Crude Oil, p. 42). The proposed building blocks could be obtained from sugars through either biological or chemical conversions. A follow-on Department of Energy report (U.S. DOE 2007) focused on lignin and its derivatives as the central building blocks.

Since 2004, a number of deconstruction pathways from biomass to biofuels have been explored, resulting in a range of potential byproduct streams. However, with the growth of the commercial biofuels industry, the primary new byproduct stream that became readily available for conversion to bioproducts was glycerol from biodiesel production. Extensive research was performed on this stream, including both biological (Clomburg and Gonzalez 2013) and chemical (Zhou et al. 2008) conversion approaches.

Also during the past decade, the number of companies undertaking research aimed at developing specific bioproducts has grown significantly. In many of these efforts, the feedstock streams being considered were sugars rather than byproduct streams from biofuels production. In some cases (e.g., succinic acid, adipic acid,

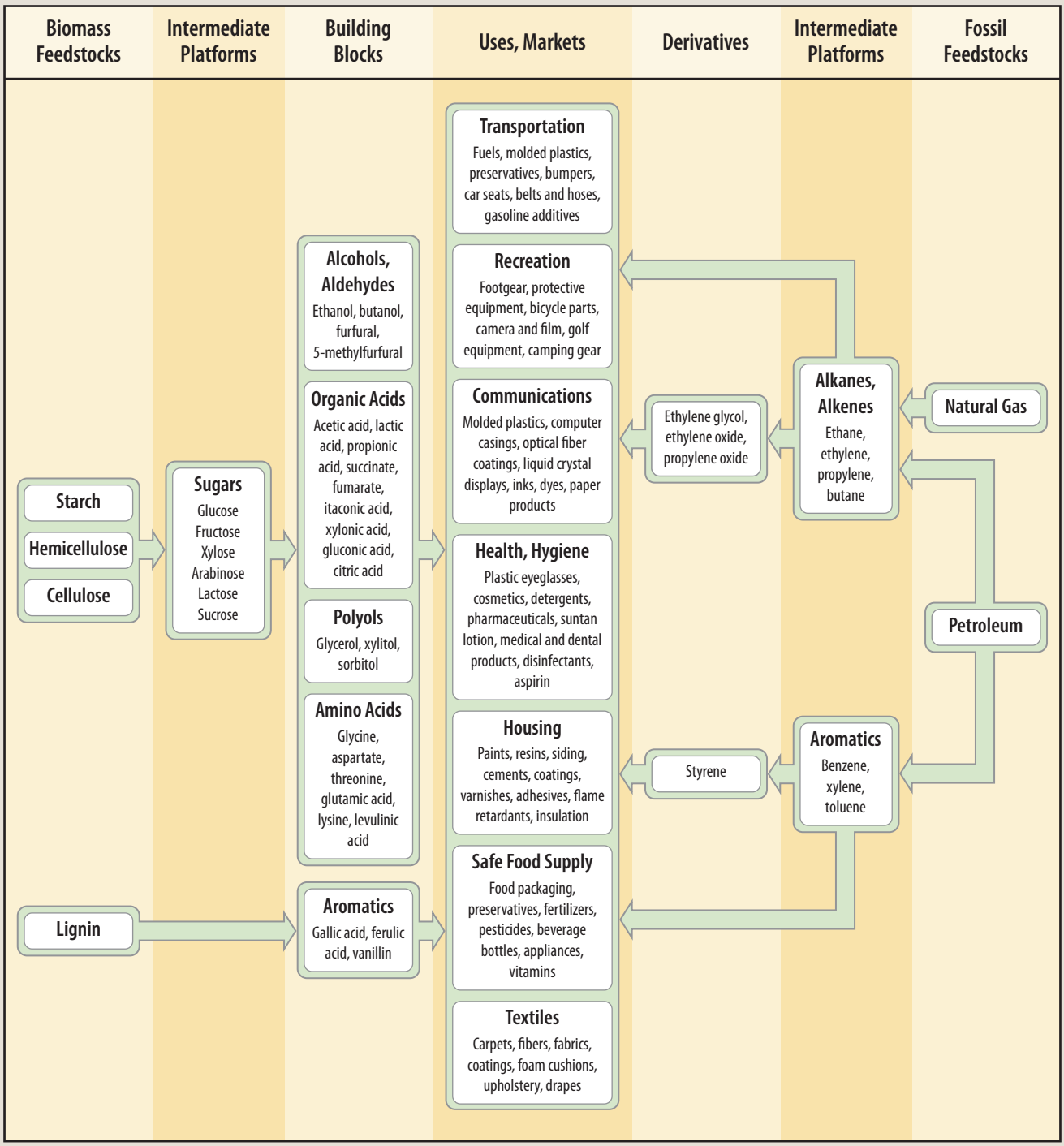


Fig. 10. Flowchart Comparing Potential Biomass- and Petroleum-Derived Products. Today, petroleum-derived products are found in virtually all facets of human life, including transportation, recreation, communications, health, housing, food safety and supply, and textiles. Lignocellulosic biomass has the potential to (1) replace petroleum and natural gas as the raw material for producing these products and (2) provide new and improved properties that could enable new products and applications. As commercial-scale production of ethanol derived from lignocellulose is coming online, synthetic biology and metabolic engineering can be applied to convert lignocellulosic biomass into any number of chemical intermediates, building blocks, and final products. By no means exhaustive, this figure represents some examples of chemical intermediates and building blocks that could be gleaned from lignocellulosic biomass to make the same products currently derived from fossil feedstocks.

acrylic acid, or isoprene), a single bioproduct was being targeted. Microbial systems were developed for conversion that led directly to the desired endproduct. In contrast, several chemicals (e.g., succinic acid, isobutanol, and 3-hydroxypropionic acid) were the targeted molecules in other cases because they could serve as platform or intermediate chemicals that could be further converted to a range of bioproducts. For these cases, the microbial system was designed to produce the intermediate compound. Diversification of metabolic biology to access a range of potential intermediate compounds also was proposed (Nikolau et al. 2008).

The emerging technological synergy between bioproducts and specialty fuels became apparent through a growing number of startup companies whose initial focus was on technology development. These technologies not only encompassed second-generation biofuels, but also began to target bioproducts. For example, synthetic biology technology to synthesize farnesene, which originally was developed for biofuels, was leveraged into bioproducts by utilizing farnesene as an intermediate chemical and extending the isoprenoid pathway knowledge to produce the specialty chemical, squalene. Isobutanol development also has been aided by its potential as a platform chemical. In each of these cases, the fermentative product could be subsequently converted to a range of bioproducts using chemical conversions or used directly as an endproduct.

Remaining Biological Research Challenges

Given the broad number of potential bioproducts, the choice of the most appropriate target molecule(s) is challenging. Therefore, technological flexibility that potentially provides viable routes to more than a single bioproduct is desirable. A second challenging aspect when exploring bioproducts from biomass is the choice of feedstock to be utilized. The possible feedstock streams can be placed into two broad categories: (1) underutilized byproduct streams from biofuels production (pre-deconstruction or generated during deconstruction) and (2) sugar streams generated from deconstruction. The challenges for the bioproducts sector are discussed in the context of these two categories of possible feedstock streams.

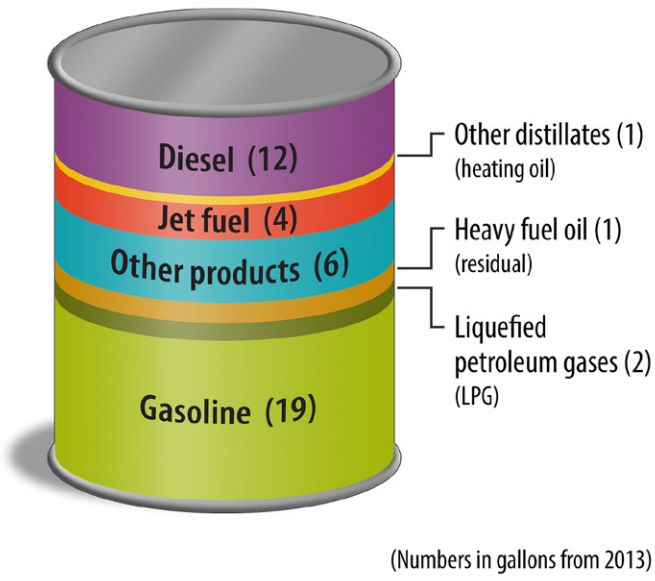


Fig. 11. Products Made from a Barrel of Crude Oil. A 42-gallon (U.S.) barrel of crude oil yields about 45 gallons of petroleum products. Falling under the "other products" category are petrochemicals including naphtha, ethane, ethylene, and other oils. [Image courtesy U.S. Energy Information Administration]

Utilizing Bioproducts Resulting from Underutilized Byproduct Streams from Biofuels Production

Molecules that qualify as coproducts from the biorefinery of the future would include chemicals that can be extracted before the biomass is processed through pretreatment and saccharification, chemicals in the liquid stream from pretreatment, and chemicals in the byproduct streams that remain after saccharification and fermentation.

Pre-deconstruction coproducts

Plants are being utilized as production platforms for vaccines, pharmaceuticals, complex molecules, and industrial enzymes (Howard and Hood 2005). Each of these molecules is manufactured in the plant host best suited to its most efficient production. For example, corn grain is used to produce exogenous industrial cellulases (Hood et al. 2007; Hood et al. 2012), and sugarcane leaves are used to produce polyhydroxybutyrates (PHBs; Petrasovits et al. 2007; Petrasovits et al. 2013). In the case of grain production of enzymes, the coproducts are starch and oil, which are not useful for the cellulosic biomass to biofuels industry. However, if defatted germ is used as the enzyme inoculum, then extra cellulosic sugars can contribute to the biofuel production stream. In the case of PHBs in sugarcane leaves, once the higher-value PHBs are extracted, the leaves can be cycled into the biorefinery as a cellulosic feedstock. Likewise, terpenes from eucalyptus are produced in high concentrations but would need to be extracted and upgraded, while the biomass residue is sent for further cellulosic conversion. Cellulosic fibers are another potential value added—perhaps for new uses such as nanocellulose, in which case the hemicellulosic sugars or lignin would be processed further. In each example, added value is achieved through deliberate production of a nonfuel product. Plant-based biological improvement of these extracted plant material products is highly likely for overall yield and potentially for improved extractability.

Current barriers:

- Cost-effective extraction of the manufactured product or raw material.
- Market development for the plant-produced product.
- Regulatory requirements for genetically modified plants.
- Integration of manufacturing with the fuel production site.

Post-deconstruction coproducts

Biomass is a complex mixture of polymers and chemicals including sugars, phenolics, and amino acids. The primary focus on manufacturing biofuels from plant materials is converting the C6 sugar stream into ethanol or butanol. This focus sidesteps all the other carbon in the streams from the biomass feedstock. Like the petroleum refinery industry, a biorefinery's ultimate goal is to utilize all this carbon. Native plant molecules and polymers that are not currently fermented to alcohols—such as lignin and C5 sugars (e.g., xylose and arabinose), as well as multiple other less abundant compounds—could be targeted for conversion into useful chemicals.

An important challenge is to determine whether the various potential byproduct streams available for recovery and conversion to bioproducts have common chemical features or if the technological approach employed for bioproduct production needs to be tailored to the deconstruction process being used. The various pretreatment processes being studied for biomass to biofuels (beginning on p. 18) will have different impacts on the biomass chemistry and must be chosen carefully to be compatible with the desired product stream.

The product streams from ground plant materials are complex at best, even after an initial division of the lignin stream from the sugar stream and the potential separation of the C5 and C6 streams. Because of this complexity, separating the various streams is a key technology that will require development or adaptation to the biomass biorefinery concept. Separations technology for liquid:liquid and solid:liquid streams will be important. In addition to the separations, high-throughput identification of the chemical species present in the streams will be a critical technology. Identification should be coupled with high-throughput technologies to assist with efficient separations. Initially, molecules that are abundant and easily separated would be important targets to develop the biorefinery concept. Bioproducts from lignin would be particularly attractive. Examples of potential opportunities from lignin macromolecules include carbon fiber, polymer modifiers, adhesives, and resins. A significant technical challenge is that lignins from different biomass sources and isolation processes have significantly differing structures, reactivity, molecular weight distributions, melting points, and polyelectrolyte properties. Appropriate lignin-conditioning process technologies will be necessary to alleviate the complications derived from these basic property and structural differences, and this research ultimately could lead to new materials for the chemical and materials industries (U.S. DOE 2007; Ragauskas et al. 2014). Lignin uses generally have been limited to applications in which the bulk properties such as solubility, surface activity, and solid content are important.

Ideally, generating products from lignin would encompass molecules that are specific products from pure or semipure lignin feedstock streams. There are historical examples in which lignin was used as a feedstock. These examples include phenolics by alkaline hydrolysis, vanillin production by mild oxidation in alkaline conditions, organic acids (e.g., benzoic, toluic, methoxybenzoic, acetic, and formic acids) by strong oxidation, phenols and aromatic hydrocarbons from lignin hydrogenolysis, and dimethyl sulfide or sulfoxide production by reaction with sulfur followed by oxidation. However, because they were not cost competitive with those produced by petroleum-based technology, these processes were discontinued. Current streams from lignin are complex mixtures and will require significant separation technology development to generate useful raw materials for manufacture into specific coproducts. In the future, the ability to create specific lignin structures in the original biomass could lead to easier utilization of the lignin stream.

Significant progress has been made in co-utilizing C5 sugars with C6 sugars in fermentative processes to produce alcohols. Consequently, the need to find alternative bioproducts derived from C5 sugars is not as critical for carbon utilization in a biorefinery as are bioproducts from the lignin-derived stream. However, the efficiency of the fermentative conversion of C5 sugars generally is less than for the C6 sugars. Additionally, the C5 sugar streams generated from hemicellulose are commonly more complex than the C6 sugar stream, which is primarily from cellulose. The number of chemical species (monomers) present in hemicelluloses is higher than in cellulose, and there are more fermentation inhibitory species in the hemicellulose-derived stream due to its

more facile conversion of C5 sugars to degradation products during biomass deconstruction than with glucose.

The challenges associated with fermenting the C5 stream to alcohols will not be addressed by merely changing to a different fermentation product. Therefore, the objective of utilizing this stream for value-added bioproducts must be predicated on either finding (1) a higher-value fermentation product that can justify the added difficulty associated with the stream or (2) an alternative conversion approach to fermentation, such as a chemical conversion process.

Current barriers:

- No systematic approach for identifying target bioproducts.
- Inconsistent potential byproduct stream compositions—no deconstruction process is the "standard."
- Inadequate compositional analysis of potential byproduct streams.
- Need for cost-effective byproduct stream conversion processes.
- Lack of effective or generalizable separation technology for the byproduct streams.
- Need for integrating separation and conversion process technologies.
- Lack of methods to depolymerize lignin.
- Frequent degradation of biomass residues after conversion.

Developing Bioproducts from Sugar Streams Generated During Deconstruction

As discussed in the Specialty Fuels section (see p. 29), a tremendous challenge for bioproducts research is to identify the "right" molecules to target. The target molecule could be exactly the same molecule currently produced through a petrochemical process (direct replacement) or could be a new molecule with properties that allow replacement of a current petrochemical (functional replacement). Both direct and functional replacement molecules continue to be a focus of bioproducts research.

Single-target molecule fermentation processes have been successfully developed such as for PDO. However, few bioproduct markets are sufficiently large to justify development of a unique organism to make a specific bioproduct. The most desirable attributes of a target bioproduct would be high value and high volume. Unfortunately, these two attributes are almost always at odds. One approach to addressing the dichotomy would be to produce platform molecules (U.S. DOE 2004). A platform molecule is an intermediate molecule that could subsequently be converted into a range of bioproducts. The bioproducts produced from a common intermediate could range from high volume and lower value to low volume and higher value but, when taken together, would require a high production volume for the intermediate. This approach potentially could bring together bioproducts and specialty fuels, an example of which is farnesene.

Whether targeting a single bioproduct or an intermediate molecule that could be used to generate a family of bioproducts, conversion systems and the catalysts within them will need to be developed. Conceptually, the technological requirements and approaches needed to develop biological conversion processes for bioproducts will be the same as those discussed in the Specialty Fuels section (p. 29), so the technological barriers for these types of processes will not be readdressed in this section. However, in addition to biological conversion processes, bioproduct production also could involve the use of chemical conversion processes. While deep knowledge of chemical catalysis has been developed in the petrochemical industry, the catalysis needs for conversion of biomass-derived molecules are significantly different and not as well known. In biomass systems, the chemical catalyst will need to function and be stable in the condensed phase, which typically is aqueous. Also, the focus of bioproduct processes will be the selective removal of functionality rather than the selective incorporation of functionality that is required in petrochemical processes (Shanks 2010). An additional challenge will be to develop chemical catalysts that can tolerate the unique impurities in biomass-derived streams, which are different from petrochemical processes.

One coproduct with potentially elastic market demand is protein—for subsequent use in animal feed or in other fermentations. The highest protein contents are in the green material, which presents an interesting underexplored research challenge in balancing protein recovery with sustainability. Protein also is easily converted by microbes—either intentionally during fermentation or unintentionally during storage.

Like the isolation of bioproducts from byproduct streams, the production of bioproducts from sugar streams also will require the development of new separation strategies. Efficient purification of fermentation products will be needed to achieve the chemical purity specifications typically required for chemical products.

Current barriers:

- No systematic identification of target bioproducts or intermediates.
- Lack of efficient, stable, and scalable biocatalysts for desired bioproducts.
- Inadequate tools for the rapid development of viable biocatalysts (for an in-depth discussion, see Specialty Fuels, p. 29).
- Limited knowledge of chemical catalyst materials that are stable and impurity tolerant while providing selective conversion of biomass-derived molecules.
- Insufficient general separation strategies for attaining high-purity bioproducts.

Key Research Opportunities

The topic of bioproducts from biomass conversion is quite broad and, therefore, difficult to navigate. Substantial research has focused on biomass production, logistics, pretreatment, and deconstruction into sugars for fermentation into fuels. If bioproducts are to be part of the newly conceptualized biorefinery that utilizes all the biomass feedstock components, then significant research is needed to determine how to gain the most value from these components.

Short Term

In the near term, research needs deemed critical to the success of bioproduct integration into the biorefinery include exploring and refining separations technology of the sugar and lignin streams resulting from biomass deconstruction. These streams contain numerous chemical species that vary depending on the deconstruction approach. Being able to separate the components and enrich those that are of highest value or in greatest abundance will provide the most initial value. Moreover, in addition to separating the components in the streams, high-throughput analytical techniques will be required to understand and characterize the pre- and post-separation streams in more detail. Specific research needs include:

- Development of novel separations technologies.
- Utilization of current byproduct streams.
- Development of generalizable post-processing technology.
- Identification of microbial and chemical pathways to promising intermediates.
- Identification and improvement of plants for higher extractible levels of desired bioproducts or intermediates.
- Increased run times (>5 fold) so that current fermentation and reaction processes are viable.
- Development of high-throughput analytical methodologies, as well as high-throughput biological and chemical catalyst synthesis technologies.
- Definition of viable target molecules.
- Identification of high-value bioproducts in existing biomass streams.
- Identification of atom-economical pathways to intermediate and final bioproducts from biomass.

Long Term

In the longer term, research needs deemed critical to the success of bioproduct integration into the biorefinery include conversion technologies and high-tech separations, as well as enzymatic applications for biomass breakdown into components and building blocks that can be utilized for bioproducts. In addition, the important need for continuous processes was identified to foster the economic viability of a biorefinery. Specific research needs include:

- Development of efficient and economical continuous processes.
- Lignin streams that are homogenous and consistent.
- Flexible process technology to accommodate multiple feedstock sources.
- Integration of the diverse carbon sources that result from a single biomass source (e.g., hexoses, pentoses, and lignin).
- Development of synthetic biological and chemical chassis that require only minor modifications for a range of bioproducts.
- Development of real time, *in situ* analytical capabilities.
- Methodology for efficiently identifying target molecules.
- Development of a suite of atom-economical bioproducts that are less toxic and more environmentally benign compared to current feedstocks and products.

In summary, the openness of the bioproducts field is both an advantage and a challenge. The focus on utilizing all the carbon streams from biomass conversion for either fuels (other than alcohols) or chemicals to replace petrochemicals could have a significant impact on the successful deployment of a biorefinery. Separations and analytical tools remain at the base of the research paradigm. Harnessing lignin also will be of extreme value; however, the bottom line will be integrating separation, analysis, and synthesis into a cost model that will enable moving these products into the marketplace.

Summary and Conclusions

This workshop report describes the current state of the science required to facilitate the emergence of a robust U.S. lignocellulosic biofuels and bioproducts industry. Substantial gains have been made toward this goal in the past decade, a period that has witnessed a significant increase in biofuels research funding and activity, including the launch of the first commercial cellulosic ethanol plants in the United States. Clearly, however, much remains to be done to accelerate the emergence of this industry.

Research advances are needed along the entire development pipeline, beginning with the biomass sources that will serve as inputs. Given the great variation of conditions under which biofuel crops will be cultivated, there will be not a single feedstock but rather a collection of feedstock options that share common attributes such as high sugar yields, reduced recalcitrance, nutrient reallocation upon senescence, and other properties that this report discusses in detail. This biomass must be effectively deconstructed into its constituent sugars and other components, preferably in a manner that minimizes deleterious effects to the conversion processes that occur downstream. Ideally, deconstruction processes will be developed that are agnostic to the input biomass; however, in the near term, multiple options still will be needed to maximize fermentable sugar yield as a function of biomass source. Given the advances in microbial pathway engineering, the time is now appropriate to think not only of a cellulosic *ethanol* industry, but rather of a cellulosic *liquid fuels* industry capable of producing renewably sourced products for domestic transportation fuel needs. This expanded array of fuel compounds likely will employ recombinant microorganisms for sugar conversion; however, thermochemical processing, either of sugars directly or fermentation-derived intermediates, also may contribute to this expanded product suite.

This report also considers the barriers inhibiting the production of value-added bioproducts from biomass. Renewable chemical production can provide benefits similar to those of renewable fuels, including more benign processing configurations, reduced environmental impacts, and security of supply, while in many cases providing economic incentives to bypass a petroleum feedstock. Bioproducts may be considered in the context of a biorefinery, in which process fractions such as lignin are upgraded to commercially viable molecules, but also important is the recognition that bioproducts have value in their own right, independent of biofuel production, for the reasons outlined previously. Many of the technological challenges that must be overcome in facilitating the accelerated emergence of cellulosic fuels also apply to biomass-derived chemicals, and the economic benefits could support a biofuels industry. Thus, considering bioproducts along with biofuels is both synergistic and efficient, not only in basic research but in commercial application.

Although the specific needs to be addressed vary as the different components of the development pipeline are considered, some common themes have emerged. First is the need for analytical tools, especially for chemical characterization. In

some cases, these technologies will help to better understand the structure of plant cell walls (Biomass Development, p. 9); in others, the tools will aid in determining the exact composition of a process stream with the aim of understanding potential inhibitory interactions (Bioproduct Development from Biomass, p. 39). In all cases, the need for high-throughput, low-volume, and high-sensitivity chemical detection and quantitation is evident.

Second, the need for mathematical and computational models emerges repeatedly. For example, there is a need for structure-function prediction of hydrolytic enzymes (Lignocellulose Deconstruction, p. 17) and the ability to more precisely determine the feasibility of certain biological and thermochemical pathways for fuel and bioproduct production (Specialty Fuels, p. 29, and Bioproduct Development from Biomass, p. 39). This encompassing need can be most succinctly stated as a desire for model-driven design and predictability throughout the research and development pipeline. Just as these tools have utility in multiple areas, research objectives and outcomes clearly are heavily influenced by common themes among the four workshop discussion areas. For example, advances in structure-function prediction of hydrolytic enzymes could facilitate advances in these same types of predictions for metabolic enzymes involved in biosynthetic pathways. Tools for engineering new microbial hosts are needed to enable high-level expression of highly active hydrolytic enzymes; these same tools would be useful in facilitating the engineering of novel microbes with beneficial phenotypes for fuel and chemical production. Consolidated bioprocessing (CBP) aims to use a single organism to achieve cellulose and hemicellulose hydrolysis and subsequent conversion of the resulting sugars to fuel, without the need for added enzymes. CBP also requires an ability to efficiently engineer diverse microorganisms. Lignocellulose deconstruction results in a complex mixture, which includes compounds that may be toxic to microbial catalysts. Likewise, there may be compounds that impair or destroy homogenous or heterogeneous catalysts designed for thermochemical conversion for either fuel upgrading or bioproduct formation. Processes to produce bioproducts that derive directly from byproduct streams of biomass deconstruction (e.g., protein or lignin fractions) are necessarily dependent on the outcome of the deconstruction step. Active collaboration among these various research areas is thus necessary to achieve significant advances in production of biomass-derived compounds. Connections such as those described herein make research goals, which at first seem disconnected, ripe for synergistic exploration and further illustrate the need to continue support of integrated, multidisciplinary research.

Finally, worth noting is that although this report focused on the scientific challenges that must be overcome for the lignocellulosic fuels and biochemicals industry to truly thrive, the engineering barriers also must ultimately be addressed. New processes arise as the result of advances in both science and engineering. In the case of, for example, biological conversion of feedstocks to fuels and chemicals, these two aspects often must be considered in tandem to achieve significant advances. One specific example is the critical role that engineering and technoeconomic analysis can play in defining those areas of research that may have the largest impact on the commercial feasibility of

cellulosic fuels plants. Such an analysis would reveal those aspects of the process that contribute the highest fraction to operating costs and, accordingly, indicate research needs whose success can provide the largest reductions in those costs. Likewise, a process engineering analysis may suggest alternative plant configurations whose implementation may be facilitated by advances in specific areas of basic research. Research priorities established in the absence of such analyses may indeed bring significant advances in basic understanding of and ability to modify a cellulosic fuels facility, but they ultimately could have limited influence on the acceleration of this industry. A continued, interdisciplinary approach toward cellulosic fuels research should be encouraged—one that fosters constant and deep communication among all aspects of the development pipeline to maximize the probability of success.

APPENDICES

Appendix 1. Workshop Agenda..... 54

Appendix 2. Workshop Scope..... 56

Appendix 3. Workshop Tasks, Charge Questions 57

Appendix 4. Workshop Participants 59

Appendix 5. Bibliography..... 61

Appendix 1. Workshop Agenda

June 23–24, 2014

Monday

8:30 a.m.	**Opening and Introductions**
9:00 a.m.	**Speaker Session 1: Defining the main immediate challenges**
9:00 a.m.	Maureen McCann: *Redefining recalcitrance: Insights gained from dissecting cell wall architecture with chemical catalysts*
9:30 a.m.	Greg Stephanopoulos: *Biomass to bioproducts: Potential, reality check, and major challenges*
10:00 a.m.	Chris Somerville: *From biomass production to downstream engineering challenges*
10:30 a.m.	**Break**
11:00 a.m.	**Breakout Session 1: Immediate challenges** A. Biomass Development: Tom Brutnell and David Braun, leaders B. Cell Wall Deconstruction: Michael Ladisch and Birgitte Ahring, leaders C. Bioproduct Development from Biomass: Beth Hood and Brent Shanks, leaders D. Specialty Fuels: Ryan Gill and Terry Papoutsakis, leaders
	Working Lunch
1:00 p.m.	**Breakout Session Presentations (5 min. each)**
1:30 p.m.	**Speaker Session 2: Paths toward solutions to immediate challenges**
1:30 p.m.	Ken Keegstra: *Sustainable biomass production: How many Ps are needed?*
2:00 p.m.	Lee Lynd: *Solving recalcitrance*
2:30 p.m.	Jay Keasling: *Production of advanced fuels from sugars using engineered microorganisms*
3:00 p.m.	**Group discussion of solutions to immediate challenges**
3:30 p.m.	**Break**
4:00 p.m.	**Breakout Session 2: Existing solutions to immediate challenges** A. Biomass Development B. Cell Wall Deconstruction C. Bioproduct Development from Biomass D. Specialty Fuels
5:30 p.m.	**Breakout Session Presentations (5 min. each)**
6:00 p.m.	**Summary of the day; tasks for next day**

Tuesday

8:30 a.m.	**Speaker Session 3: Bioenergy after 2025**
8:30 a.m.	Michael Martin: *Breeding new bioenergy crops: Not just the genetics*
9:00 a.m.	**Discussion**
9:45 a.m.	**Break**
10:15 a.m.	**Breakout Session 3: Revisiting the challenges and possible solutions with opportunities for process integration** A. Phase II Breakout Group A B. Phase II Breakout Group B C. Phase II Breakout Group C D. Phase II Breakout Group D
11:30 a.m.	**Breakout Session Presentations (5 min. each)**
	Working lunch
12:00 p.m.	**Breakout Session 4: Identifying main gaps in our knowledge and solutions in the 5-, 10-, and 20-year windows** A. Biomass Development B. Cell Wall Deconstruction C. Bioproduct Development from Biomass D. Specialty Fuels
2:00 p.m.	**Breakout Session Presentations (5 min. presentation; 5 min. discussion)**
2:40 p.m.	**Revisit Task Questions, Discuss, and Complete**
4:30 p.m.	**Adjourn**

Appendix 2. Workshop Scope

DOE and Bioenergy: Background

Long-term projected increases in U.S. and global energy needs require that renewable resources be considered as part of the energy supply landscape. Bioenergy derived from biomass offers several consistent advantages in energy security, decreased greenhouse gas emissions, and improved economic considerations including new revenue streams for farmers and new biotechnology for a biobased economy. Recent advances in genomic science and biotechnology have demonstrated that cellulose and associated plant components from nonfood crops can be converted into a range of biofuel and chemical compounds, thereby replacing products currently derived from petroleum. Nonlignocellulosic phototrophic biomass such as micro- and macroalgae also may serve as inputs for biological or thermochemical conversion to products of interest. Insight gained from coordinated basic research at the U.S. Department of Energy (DOE) Bioenergy Research Centers and many other institutions on bioenergy feedstock development, cell wall deconstruction, and microbial conversion strategies provides a scientific foundation from which to build a renewable biofuels and bioproducts industry. Yet, despite significant research advancements, commercial production of cellulosic biofuels and bioproducts remains constrained by production costs. Continued basic research is warranted to address remaining bottlenecks in bioenergy production from plant biomass to increase efficiencies and lower production costs. To explore the remaining key scientific and technological gaps in bioenergy production, DOE's Office of Biological Environmental Research (BER) will organize a workshop to assess the current state of the science with regard to production of advanced biofuels and bioproducts from biomass and identify remaining barriers and new opportunities for sustainable biomass production and conversion. This workshop will be informed by the experiences and technical realities of existing commercial enterprises seeking to develop advanced biofuels; however, it will not seek to identify research areas intended to address short-term industrial problems.

Focus of Workshop

- Articulate the current state of the science in feedstock production and biomass deconstruction and conversion from academic, governmental, and industrial laboratories.
- Identify areas of understanding of the structural properties of plant cell walls that will inform engineering and breeding strategies to reduce resistance to deconstruction and conversion to biofuels without negatively impacting plant vitality.
- Identify strategies to reduce the high cost associated with currently available physical, chemical, enzymatic, and microbial treatments for the breakdown of lignocellulosic and nonlignocellulosic biomass and conversion to biofuel and biochemical compounds.
- Identify gaps in fundamental understanding of plant biology with the potential to serve as biomass feedstock crops.
- Identify opportunities for engineering microorganisms to mediate deconstruction of complex plant biomass, synthesis of advanced biofuel compounds and bioproducts, and photosynthetic capture of atmospheric carbon dioxide (CO_2).
- Identify strategies to overcome limitations in genome-scale engineering and targeted biodesign aimed to improve biomass feedstock crop yields or microbial deconstruction, microbial conversion of biomass, and synthesis of advanced biofuels compatible with existing engines and value-added biochemicals.

Expected Outcomes

The output of this workshop will be a report that assesses the current state of the science with regard to the production of advanced cellulosic biofuels and bioproducts and identifies remaining scientific and technical barriers to the establishment of a sustainable next-generation biofuels and bioproducts commercial sector at the national level. This report will be used to directly inform DOE BER programs in bioenergy, defining goals and describing key research approaches needed to generate results and conclusions.

Appendix 3. Workshop Tasks, Charge Questions

Breakout Session Tasks and Charge Questions – Phase I

Biomass Development

Tasks

1. Articulate the current state of the science in feedstock production from academic, governmental, and industrial laboratories.
2. Identify areas of understanding concerning structural properties of plant cell walls that will inform engineering and breeding strategies to reduce resistance to deconstruction and conversion to biofuels without negatively impacting biomass production.
3. Identify gaps in the fundamental understanding of the biology of plants that have the potential to serve as biomass feedstock crops.
4. Identify strategies to overcome limitations in genome-scale engineering and targeted biodesign aimed at improving biomass feedstock crop quality and yield.
5. Articulate the technological barriers that impede economically feasible implementation of biomass as a feedstock.

Questions

1. What are the most promising crop species to achieve high yields?
   a. Limitations rated by geography and transportation limitation.
   b. Potential for domestic and global use.
2. Should feedstock development be tiered? For example, research would first exploit existing infrastructure for harvesting and shipping of lignocellulosics using stover yielded from maize or bagasse from cane and then move to more dedicated feedstocks as technologies develop.
3. What are the unique features of C4 versus C3 cell walls and woody trees that may limit single-stream processing? How might they be modified to allow this type of processing?
4. What model systems should be developed to accelerate discovery?
5. What traits should be stacked (e.g., enhanced pest resistance and reduced input requirements)?
6. What is the breadth of lignocellulosic production domestically and abroad? How does this affect feedstock economics? What are competing uses? What are feedstock costs at process inlet versus feedstock production site? What are costs of storage? Can biomass feedstock crops be designed to be more amenable to storage?

Cell Wall Deconstruction

Tasks

1. Articulate the current state of the science in biomass deconstruction from academic, governmental, and industrial laboratories.
2. Identify areas of understanding concerning structural properties of plant cell walls that will inform engineering and breeding strategies to reduce resistance to deconstruction and conversion to biofuels without negatively impacting plant vitality.
3. Identify strategies to reduce the high cost associated with currently available physical, chemical, enzymatic, and microbial treatments for the breakdown of lignocellulosic and nonlignocellulosic biomass and conversion to biofuel and biochemical compounds.
4. Identify strategies to overcome limitations in genome-scale engineering and targeted biodesign aimed at improving microbial and enzymatic deconstruction.

Questions

1. What amount of cell wall and storage carbohydrates can be made in a soluble form to make "sweet biofuel crops"?
2. What is the ideal balance between hydrolytic enzyme expression *in planta*, pretreatment, and subsequent catalysis of carbohydrate breakdown for different biofuel crops? What is the potential of identifying and engineering the accumulation of cell wall inhibitors of microbial hydrolytic enzymes?
3. What other high-value products can be recovered during the breakdown of lignocellulosic feedstocks?
4. What are the real monetary and energy costs of cell wall deconstruction?
5. Given that, in nature, organic material is recycled by microbial communities, can mixed cultures be used to improve engineered system performance and economics, and, if so, how?

Bioproduct Development from Biomass

Tasks

1. Articulate the current state of the science and lessons learned in biobased feedstock conversion to bioproducts from academic, governmental, and industrial laboratories.

2. Identify strategies, including process engineering approaches, to reduce the high costs associated with currently available physical, chemical, enzymatic, and microbial treatments for the conversion of lignocellulosic and nonlignocellulosic biomass and conversion to biofuel and biochemical compounds.
3. Identify opportunities for engineering microorganisms to mediate synthesis of advanced bioproducts.
4. Identify opportunities for using mixed cultures with a broader spectrum of capabilities for deconstruction and synthesis.
5. Identify opportunities for using thermochemical and integrated biothermoconversion methods (e.g., development of novel chemistries and catalysts) to produce new bioproducts from biomass.

Questions

1. What products should be made from biomass?
2. Given the hundreds of potential bioproducts, how can basic research efforts be connected to more than a single bioproduct?
3. What new paradigms can be designed to produce high-value coproducts to improve the overall economics of fuels and chemicals in a biorefinery?
4. How many different feedstocks can be combined for sugar production? Can each of them contribute a coproduct in a hub and spokes model where the hub uses pre-extracted feedstocks for commodity fuel production?
5. If coproducts from abundant biomass sources should saturate the market for each coproduct, could having multiple coproducts using multiple feedstocks resolve that problem? What issues does that raise for converting the residual biomass into sugars for commodity fuels?
6. How can lignin be processed efficiently into a coproduct for conversion into high-value endproducts? What future is there in using plant phenolics for BTX production?

Specialty Fuels

Tasks

1. Articulate the current state of the science in specialty fuels biosynthesis from academic, governmental, and industrial laboratories.
2. Identify opportunities for engineering microorganisms to mediate deconstruction of complex plant biomass, synthesis of advanced biofuel compounds, and photosynthetic capture of atmospheric CO_2.
3. Identify strategies to reduce the high cost associated with currently available physical, chemical, enzymatic, and microbial treatments for the breakdown of lignocellulosic and nonlignocellulosic biomass and conversion to biofuel and biochemical compounds.
4. Identify strategies to overcome limitations in genome-scale engineering and targeted biodesign aimed at improving microbial conversion of biomass and synthesis of advanced biofuels compatible with existing engines and value-added biochemicals.
5. Identify strategies for using mixed-culture engineering for operation of consolidated bioprocesses.

Questions

1. What is the potential for hybrid biological-chemical conversion to upgrade precursor molecules for specialty fuels?
2. What balance of fuels to products should be the goal?
3. What smaller-volume but higher-value fuel additives, or fuel-upgrading molecules, are viable targets?
4. How will the ability to rapidly design and construct biofuel-producing organisms expressing a range of desired traits affect the selection of specialty fuel targets and associated processes?
5. Because oil provides a source for most fuels, making shipping and storing easy, and specialty fuels are generated as need arises, should a similar strategy be considered for biofuels? If so, what would that be?
6. What are the major limitations (and promises) related to CO_2 capture and conversion for high-density specialty fuels?

Breakout Session Tasks and Charge Questions – Phase II

1. What are the opportunities for process integration?
2. What can be done upstream to facilitate later process steps?

Appendix 4. Workshop Participants

Name	Institution	Role
Birgitte Ahring	Washington State University	Breakout Lead - Deconstruction
Ana Alonso	Ohio State University	Participant
David Braun	University of Missouri	Breakout Lead - Biomass
Stevens Brumbley	University of North Texas	Participant
Thomas Brutnell	Danforth Center	Breakout Lead - Biomass
Bruce Dale	Michigan State University, GLBRC	Participant
Tim Donohue	University of Wisconsin, GLBRC	Participant
John Frost	Michigan State University	Participant
Ryan Gill	University of Colorado, Boulder	Breakout Lead - Specialty Fuels
Paul Gilna	BioEnergy Science Center, ORNL	Participant
Louise Glass	UC Berkeley	Participant
Erich Grotewold	Ohio State University	Co-Chair
Michael Himmel	National Renewable Energy Laboratory	Participant
Elizabeth Hood	Arkansas State University	Breakout Lead - Bioproducts
Jay Keasling	Joint BioEnergy Institute, LBNL	Speaker
Ken Keegstra	Michigan State University, GLBRC	Speaker
Michael Ladisch	Purdue University	Breakout Lead - Deconstruction
Yebo Li	Ohio State University	Participant
Dominique Loque	Joint BioEnergy Institute, LBNL	Participant
Lee Lynd	Dartmouth College	Speaker
Michael Martin	Monsanto Corporation	Speaker
Maureen McCann	Purdue University, C3Bio	Speaker
James McMillan	National Renewable Energy Laboratory	Participant
Dennis Miller	Michigan State University	Participant
Ray Ming	University of Illinois	Participant
Debra Mohnen	University of Georgia; BioEnergy Science Center	Participant
John Morgan	Purdue University	Participant
Basil Nikolau	Iowa State University	Participant
Michelle O'Malley	University of California at Santa Barbara	Participant
Terry Papoutsakis	University of Delaware	Breakout Lead - Specialty Fuels
John Perkins	POET-DSM Advanced Biofuels	Participant
Brian Pfleger	University of Wisconsin	Participant
Steve Picataggio	DuPont Inc.	Participant
Kris Jones Prather	Massachusetts Institute of Technology	Co-Chair
Yuriy Roman	Massachusetts Institute of Technology	Participant
Richard Sayre	Los Alamos National Laboratory	Participant

Continued on next page

Continued from previous page

Name	Institution	Role
Andreas Schirmer	Renewable Energy Group, Inc.	Participant
Susannah Scott	University of California at Santa Barbara	Participant
Brent Shanks	Iowa State University	Breakout Lead - Bioproducts
Blake Simmons	Joint BioEnergy Institute, SNL	Participant
Chris Somerville	UC Berkeley, Energy Biosciences Institute	Speaker
Greg Stephanopoulos	Massachusetts Institute of Technology	Speaker
Bob Tabita	Ohio State University	Participant
Christian Tobias	USDA Agricultural Research Service, Albany	Participant
Wilfred Vermirris	University of Florida	Participant

Table Acronyms

GLBRC: Great Lakes Bioenergy Research Center
LBNL: Lawrence Berkeley National Laboratory
ORNL: Oak Ridge National Laboratory
SNL: Sandia National Laboratories
UC Berkeley: University of California at Berkeley
USDA: U.S. Department of Agriculture

Report preparation: Biological and Environmental Research Information System group at Oak Ridge National Laboratory (Benjamin Allen, Kris Christen, Holly Haun, Brett Hopwood, Betty Mansfield, Sheryl Martin, Marissa Mills, and Judy Wyrick)

Appendix 5. Bibliography

Atsumi, S., Z. Li, and J. C. Liao. 2009. "Acetolactate Synthase from *Bacillus subtilis* Serves as a 2-Ketoisovalerate Decarboxylase for Isobutanol Biosynthesis in *Escherichia coli*," *Applied and Environmental Microbiology* **75**(19), 6306–6311. DOI: 10.1128/AEM.01160-09.

Baxter, H. L., et al. 2014. "Two-Year Field Analysis of Reduced Recalcitrance Transgenic Switchgrass," *Plant Biotechnology Journal*, 1–11. DOI: 10.1111/pbi.12195.

Bihmidine, S., et al. 2013. "Regulation of Assimilate Import into Sink Organs: Update on Molecular Drivers of Sink Strength," *Frontiers in Plant Science* **4**, 177. DOI: 10.3389/fpls.2013.00177.

Blanch, H. W. 2012. "Bioprocessing for Biofuels," *Current Opinion in Biotechnology* **23**(3), 390–95. DOI: 10.1016/j.copbio.2011.10.002.

Bommarius, A. S., J. K. Blum, and M. J. Abrahamson. 2011. "Status of Protein Engineering for Biocatalysts: How to Design an Industrially Useful Biocatalyst," *Current Opinion in Chemical Biology* **15**(2), 194–200. DOI: 10.1016/j.cbpa.2010.11.011.

Bonawitz, N. D., et al. 2014. "Disruption of Mediator Rescues the Stunted Growth of a Lignin-Deficient *Aradopsis* Mutant," *Nature* **509**(7500), 376–80. DOI: 10.1038/nature13084.

Boyle, N. R., et al. 2013. "Recombineering to Homogeneity: Extension of Multiplex Recombineering to Large-Scale Genome Editing," *Biotechnology Journal* **8**(5), 515–22. DOI: 10.1002/biot.201200237.

Brandt, A., et al. 2013. "Deconstruction of Lignocellulosic Biomass with Ionic Liquids," *Green Chemistry* **15**, 550–83. DOI: 10.1039/C2GC36364J.

Brunecky, R., et al. 2013. "Revealing Nature's Cellulase Diversity: The Digestion Mechanism of *Caldicellulosiruptor bescii* CelA," *Science* **342**(6165), 1513–16. DOI: 10.1126/science.1244273.

Burton, R. A., and G. B. Fincher. 2014. "Plant Cell Wall Engineering: Applications in Biofuel Production and Improved Human Health," *Current Opinion in Biotechnology* **26**, 79–84. DOI: 10.1016/j.copbio.2013.10.007.

Carter, S. A., M. Rodemeyer, M. S. Garfinkel, and R. M. Friedman. 2014. *Synthetic Biology and the U.S. Biotechnology Regulatory System: Challenges and Options*. J. Craig Ventor Institute. www.jcvi.org/cms/research/projects/synthetic-biology-and-the-us-biotechnology-regulatory-system/overview/.

Chen, X., et al. 2013. "Comparison of Different Mechanical Refining Technologies on the Enzymatic Digestibility of Low Severity Acid Pretreated Corn Stover," *Bioresource Technology* **147**, 401–08. DOI: 1016/j.biotech.2013.07.109.

Chundawat, S. P. S., and G. T. Beckham. 2011. "Deconstruction of Lignocellulosic Biomass to Fuels and Chemicals," *Annual Review of Chemical and Biomolecular Engineering* **2**, 121–45. DOI: 10.1146/annurev-chembioeng-061010-114205.

Clark, T. A., et al. 2012. "Characterization of DNA Methyltransferase Specificities Using Single-Molecule, Real-Time DNA Sequencing," *Nucleic Acids Research* **40**(4), e29. DOI: 10.1093/nar/gkr1146.

Clomburg, J. M., and R. Gonzalez. 2013. "Anaerobic Fermentation of Glycerol: A Platform for Renewable Fuels and Chemicals," *Trends in Biotechnology* **31**(1), 20–28. DOI: 10.1016/j.tibtech.2012.10.006.

Dale, B. E., and R. G. Ong. 2012. "Energy, Wealth, and Human Development: Why and How Biomass Pretreatment Research Must Improve," *Biotechnology Progress* **28**(4), 893–98. DOI: 10.1002/btpr.1575.

Dellomonaco, C., et al. 2011. "Engineered Reversal of the ß-Oxidation Cycle for the Synthesis of Fuels and Chemicals," *Nature* **476**(7360), 355–59. DOI: 10.1038/nature10333.

Donohoe, B. S., et al. 2011. "Surface and Ultrastructural Characterization of Raw and Pretreated Switchgrass," *Bioresource Technology* **102**(24), 11,097–104. DOI: 10.1016/j.biortech.2011.03.092.

Fischer, C. R., D. Klein-Marcuschamer, and G. Stephanopoulos. 2008. "Selection and Optimization of Microbial Hosts for Biofuels Production," *Metabolic Engineering* **10**(6), 295–304. DOI: 10.1016/j.ymben.2008.06.009.

Gibson, D. G., et al. 2010. "Creation of a Bacterial Cell Controlled by a Chemically Synthesized Genome," *Science* **329**(5987), 52–56. DOI: 10.1126/science.1190719.

Gibson, D. G., et al. 2009. "Enzymatic Assembly of DNA Molecules Up to Several Hundred Kilobases," *Nature Methods* **6**(5), 343–45. DOI: 10.1038/nmeth.1318.

Gibson, D. G., et al. 2008. "One-Step Assembly in Yeast of 25 Overlapping DNA Fragments to Form a Complete Synthetic *Mycoplasma genitalium* Genome," *Proceedings of the National Academy of Sciences USA* **105**(51), 20404–09. DOI: 10.1073/pnas.0811011106.

Gonzalez, D., et al. 2014. "The Functions of DNA Methylation by CcrM in *Caulobacter crescentus*: A Global Approach," *Nucleic Acids Research* **42**(6), 3720–35. DOI: 10.1093/nar/gkt1352.

Gray, J., D. Caparrós-Ruiz, and E. Grotewold. 2012. "Grass Phenylpropanoids: Regulate Before Using!" *Plant Science* **184**, 112–20. DOI: 10.1016/j.plantsci.2011.12.008.

Green, E. M. 2011. "Fermentative Production of Butanol—the Industrial Perspective," *Current Opinion in Biotechnology* **22**(3), 337–43. DOI: 10.1016/j.copbio.2011.02.004.

Gronenberg, L. S., R. J. Marcheschi, and J. C. Liao. 2013. "Next Generation Biofuel Engineering in Prokaryotes," *Current Opinion in Chemical Biology* **17**(3), 462–71. DOI: 10.1016/j.cbpa.2013.03.037.

Hatzimanikatis, V., et al. 2005. "Exploring the Diversity of Complex Metabolic Networks," *Bioinformatics* **21**(8), 1603–09. DOI: 10.1093/bioinformatics/bti213.

Hazelwood, L. A., et al. 2008. "The Ehrlich Pathway for Fusel Alcohol Production: A Century of Research on *Saccharomyces cerevisiae* Metabolism," *Applied and Environmental Microbiology* **74**(8), 2259–66. DOI: 10.1128/AEM.02625-07.

Hood, E. E., et al. 2012. "Manipulating Corn Germplasm to Increase Recombinant Protein Accumulation," *Plant Biotechnology Journal* **10**(1), 20–30. DOI: 10.1111/j.1467-7652.2011.00627.x.

Hood, E. E., et al. 2007. "Subcellular Targeting is a Key Condition for High-Level Accumulation of Cellulase Protein in Transgenic Maize Seed," *Plant Biotechnology Journal* **5**(6), 709–19. DOI: 10.111/j.1467-7652.2007.00275.x.

Howard, J. A., and E. Hood. 2005. "Bioindustrial and Biopharmaceutical Products Produced in Plants," *Advances in Agronomy* **85**, 91–124. DOI: 10.1016/S0065-2113(04)85002-8.

Jung, H. J., D. A. Samac, and G. Sarath. 2012. "Modifying Crops to Increase Cell Wall Digestibility," *Plant Science* **185–86**, 65–67. DOI: 10.1016/j.plantsci.2011.10.014.

Juturu, V., and J. C. Wu. 2012. "Microbial Xylanases: Engineering, Production, and Industrial Applications," *Biotechnology Advances* **30**(6), 1219–27. DOI: 10.1016/j.biotechadv.2011.11.006.

Kaul, P., and Y. Asano. 2012. "Strategies for Discovery and Improvement of Enzyme Function: State of the Art and Opportunities," *Microbial Biotechnology* **5**(1), 18–33. DOI: 10.1111/j.1751-7915.2011.00280.x.

Kim, M., and D. Day. 2013. "Enhancement of the Enzymatic Digestibility and Ethanol Production from Sugarcane Bagasse by Moderate Temperature-Dilute Ammonia Treatment," *Applied Biochemistry and Biotechnology*," **171**(5), 1108–17. DOI: 10.1007/s12010-013-0327-7.

Kim, Y., et al. 2013. "Fractionation of Cellulase and Fermentation Inhibitors from Steam Pretreated Mixed Hardwood," *Bioresource Technology* **135**, 30–38. DOI: 10.1016/j.biotech.2012.10.130.

Kim, Y., et al. 2011. "Soluble Inhibitors/Deactivators of Cellulase Enzymes from Lignocellulosic Biomass," *Enzyme and Microbial Technology* **48**(4–5), 408–15. DOI: 10.1016/j.enzmictec.2011.01.007.

Klein-Marcuschamer, D., et al. 2010. "Technoeconomic Analysis of Biofuels: A Wiki-Based Platform for Lignocellulosic Biorefineries," *Biomass and Bioenergy* **34**(12), 1914–21. DOI: 10.1016/j.biombioe.2010.07.033.

Koch, K. E. 1996. "Carbohydrate-Modulated Gene Expression in Plants," *Annual Review of Plant Physiology and Plant Molecular Biology* **47**, 509–40. DOI: 10.1146/annurev.arplant.r7.1.509.

Kung, Y., W. Runguphan, and J. D. Keasling. 2012. "From Fields to Fuels: Recent Advances in the Microbial Production of Biofuels," *American Chemical Society Synthetic Biology* **1**(11), 498–513. DOI: 10.1021/sb300074k.

Ladisch, M. R., et al. 2010. "Converting Cellulose to Biofuels," *Chemical Engineering Progress* **106**(3)56–63. ISSN: 03607275.

Lajoie, M. J., et al. 2013. "Genomically Recoded Organisms Expand Biological Functions," *Science* **342**(6156), 357–60. DOI: 10.1126/science.1241459.

Lennen, R. M., and B. F. Pfleger. 2013. "Microbial Production of Fatty Acid–Derived Fuels and Chemicals," *Current Opinion in Biotechnology* **24**(6), 1044–53. DOI: 10.1016/j.copbio.2013.02.028.

Leonowicz, A., et al. 2001. "Fungal Laccase: Properties and Activity on Lignin," *Journal of Basic Microbiology* **41**(3–4), 185–227. DOI: 10.1002/1521-4028(200107)41:3/4<185::AID-JOBM185>3.0.CO;2-T.

Li, X., et al. 2010. "Lignin Monomer Composition Affects *Arabidopsis* Cell-Wall Degradability After Liquid Hot Water Pretreatment," *Biotechnology for Biofuels* **3**(27). DOI: 10.1186/1754-6834-3-27.

Lin, P. P., et al. 2014. "Isobutanol Production at Elevated Temperatures in Thermophilic *Geobacillus thermoglucosidasius*," *Metabolic Engineering* **24**, 1–8. DOI: 10.1016/j.ymben.2014.03.006.

Lloyd, T. A., and C. E. Wyman. 2005. "Combined Sugar Yields for Dilute Sulfuric Acid Pretreatment of Corn Stover Followed by Enzymatic Hydrolysis of the Remaining Solids," *Bioresource Technology* **96**(18), 1967–77. DOI: 10.1016/j.biotech.2005.01.011.

Mali, P., K. M. Esvelt, and G. M. Church. 2013. "Cas9 as a Versatile Tool for Engineering Biology," *Nature Methods* **10**, 957–63. DOI: 10.1038/nmeth.2649.

Mandels, M., L. Hontz, and J. Nystrom. 1974. "Enzymatic Hydrolysis of Waste Cellulose," *Biotechnology and Bioengineering* **16**(11), 1471–93. Online: Feb. 18, 2004. DOI: 10.1002/bit.260161105.

Mandels, M., R. Andreotti, and C. Roche. 1976. "Measurement of Saccharifying Cellulase," *Biotechnology & Bioengineering Symposium* **6**: Wiley, New York, 21–33. No text available. www.ncbi.nlm.nih.gov/pubmed/1000065.

Mazumder, R., et al. 2000. "Low-Substrate Regulated Microaerophilic Behavior as a Stress Response of Aquatic and Soil Bacteria," *Current Microbiology* **41**(2), 79–83. DOI: 10.1007/s0028-40010097.

NAE. 2010. *The Power of Renewables: Opportunities and Challenges for China and the United States*. National Academy of Engineering, National Research Council, Chinese Academy of Sciences, and Chinese Academy of Engineering. ISBN: 13-978-0-309-16000-1. www.nae.edu/Publications/Reports/47328.aspx.

Nikolaou, S., S. M. Gaida, and E. T. Papoutsakis. 2010. "A Comparative View of Metabolite and Substrate Stress and Tolerance in Microbial Bioprocessing: From Biofuels and Chemicals, to Biocatalysis and Bioremediation," *Metabolic Engineering* **12**(4), 307–31. DOI: 10.1016/j.ymben.2010.03.004.

Nikolau, B. J., et al. 2008. "Platform Biochemicals for a Biorenewable Chemical Industry," *The Plant Journal* **54**(4), 536–45. DOI: 10.1111/j.1365-313X.20008.03484.x.

Nivón, L. G., et al. 2014. "Automating Human Intuition for Protein Design," *Proteins: Structure, Function, and Bioinformatics* **82**(5), 858–66. DOI: 10.1002/prot.24463.

OSTP. 2012. *National Bioeconomy Blueprint*. White House Office of Science and Technology Policy (April 2012). www.whitehouse.gov/sites/default/files/microsites/ostp/national_bioeconomy_blueprint_april_2012.pdf.

Paredes, C. J., K. V. Alsaker, and E. T. Papoutsakis. 2005. "A Comparative Genomic View of Clostridial Sporulation and Physiology," *Nature Reviews Microbiology* **3**, 969–78. DOI: 10.1038/nrmicro1288.

Peplow, M. 2014. "Cellulosic Ethanol Fights for Life," *Nature* **507**(7491), 152–53. DOI: 10.1038/507152a.

Peralta-Yahya, P. P., and J. D. Keasling. 2010. "Advanced Biofuel Production in Microbes," *Biotechnology Journal* **5**(2), 147–62. DOI: 10.1002/biot.200900220.

Petrasovits, L. A., et al. 2013. "Chemical Inhibition of Acetyl Coenzyme A Carboxylase as a Strategy to Increase Polyhydroxybutyrate Yields in Transgenic Sugarcane," *Plant Biotechnology Journal* **11**(9), 1146–51. DOI: 10.1111/pbi.12109.

Petrasovits, L. A., et al. 2007. "Production of Polyhydroxybutyrate in Sugarcane," *Plant Biotechnology Journal* **5**(1), 162–72. DOI: 10.1111/j.1467-7652.2006.00229.x.

Poteete, A. R. 2001. "What Makes the Bacteriophage λ Red System Useful for Genetic Engineering: Molecular Mechanism and Biological Function," *FEMS Microbiology Letters* **201**(1), 9–14. DOI: 10.111/j.1574-6968.2001.tb10725.x.

Ragauskas, A. J., et al. 2014. "Lignin Valorization: Improving Lignin Processing in the Biorefinery," *Science* **344**(6185), 1246843(1–10). DOI: 10.1126/science.1246843.

Ragauskas, A. J., et al. 2006. "The Path Forward for Biofuels and Biomaterials," *Science* **311**(5760), 484–89. DOI: 10.1126/science.1114736.

Resch, M. G., et al. 2014. "Clean Fractionation Pretreatment Reduces Enzyme Loadings for Biomass Saccharification and Reveals the Mechanism of Free and Cellulosomal Enzyme Synergy," *ACS Sustainable Chemistry and Engineering* **2**(6), 1377–87. DOI: 10.1021/sc500210w.

Roitsch, T. 1999. "Source-Sink Regulation by Sugar and Stress," *Current Opinion in Plant Biology* **2**(3), 198–206. DOI: 10.1016/S1369-5266(99)80036-3.

Rolland, F., E. Baena-Gonzalez, and J. Sheen. 2006. "Sugar Sensing and Signaling in Plants: Conserved and Novel Mechanisms," *Annual Review of Plant Biology* **57**, 675–709. DOI: 10.1146/annurev.arplant.57.032905.105441.

Sapp, M. 2013. "Mascoma Passes 1 Billion Gallon Mark for Its MGT Yeast Technology," *BiofuelsDigest*. www.biofuelsdigest.com/bdigest/2013/12/16/mascoma-procures-1-billion-gallons-from-its-bioprocessing-technology/.

Schirmer, A., et al. 2010. "Microbial Biosynthesis of Alkanes," *Science* **329**(5991), 559–62. DOI: 10.1126/science.1187936.

Shanks, B. H. 2010. "Conversion of Biorenewable Feedstocks: New Challenges in Heterogeneous Catalysis," *Industrial Engineering Chemical Research* **49**(21), 10212–17. DOI: 10.1021/ie100487r.

Smith, M. A., et al. 2013. "*Hypocrea jecorina* Cellobiohydrolase I Stabilizing Mutations Identified Using Noncontiguous Recombination," *ACS Synthetic Biology* **2**(12), 690–96. DOI: 10.1021/sb400010m.

Socha, A. M., et al. 2014. "Efficient Biomass Pretreatment Using Ionic Liquids Derived from Lignin and Hemicellulose," *Proceedings of the National Academy of Sciences USA* **111**(35), E3587–95. DOI: 10.1073/pnas.1405685111.

Sommer, P., T. Georgieva, and B. K. Ahring. 2004. "Potential for Using Thermophilic Anaerobic Bacteria for Bioethanol Production from Hemicellulose," *Biochemical Society Transactions* **32**(2), 283–89. DOI: 10.1042/BST0320283.

Tao, L., et al. 2014. "Techno-Economic Analysis and Life-Cycle Assessment of Lignocellulosic Biomass to Sugars Using Various Pretreatment Technologies," in *Biological Conversion of Biomass for Fuels and Chemicals: Explorations from Natural Utilization Systems. RSC Energy and Environment Series* **10**, 358–80. DOI: 10.1039/9781849734738-00358.

Thomason, L., et al. 2007. "Recombineering: Genetic Engineering in Bacteria Using Homologous Recombination." In *Current Protocols in Molecular Biology*. Hoboken, N.J.: John Wiley & Sons, Inc., Chapter 1, Unit 1.16, pp. 11–24 (1.16.1–1.16.24). DOI: 10.1002/0471142727.mb0116s106. Full text: http://redrecombineering.ncifcrf.gov/References_files/CurrentProtocols.pdf.

Uppugundla, L., et al. 2014. "A Comparative Study of Ethanol Production Using Dilute Acid, Ionic Liquid, and AFEX™ Pretreated Corn Stover," *Biotechnology for Biofuels* **7**(72). DOI: 10.1186/1754-6834-7-72.

USDA. 2014. *Renewable Chemicals and Materials Opportunity Assessment: Major Job Creation and Agricultural Sector Engine*. Prepared by Nexant Inc. for the U.S. Department of Agriculture (January 2014). www.usda.gov/oce/reports/energy/USDA_RenewChems_Jan2014.pdf.

U.S. DOE. 2014. *Research for Sustainable Bioenergy: Linking Genomic and Ecosystem Sciences, Workshop Report*, DOE/SC-0167. U.S. Department of Energy Office of Science. genomicscience.energy.gov/sustainability/.

U.S. DOE. 2012. *Biosystems Design: Report from the July 2011 Workshop*, DOE/SC-0141. U.S. Department of Energy Office of Science. genomicscience.energy.gov/biosystemsdesign/report/biosystemsdesignreport2012.pdf.

U.S. DOE. 2011. *U.S. Billion-Ton Update: Biomass Supply for a Bioenergy and Bioproducts Industry*. R. D. Perlack and B. J. Stokes (Leads), ORNL/TM-2011/224. Oak Ridge National Laboratory, Oak Ridge, TN. 227p. www1.eere.energy.gov/bioenergy/pdfs/billion_ton_update.pdf.

U.S. DOE 2010. *National Algal Biofuels Technology Roadmap*. U.S. Department of Energy, Office of Energy Efficiency and Renewable Energy, Biomass Program. www1.eere.energy.gov/bioenergy/pdfs/algal_biofuels_roadmap.pdf.

U.S. DOE. 2007. *Top Value Added Chemicals from Biomass. Volume II—Results of Screening for Potential Candidates from Biorefinery Lignin*, PNNL-16983. Prepared by Pacific Northwest National Laboratory and National Renewable Energy Laboratory for the U.S. Department of Energy Office of Energy Efficiency and Renewable Energy (October 2007). www.osti.gov/scitech/biblio/921839.

U.S. DOE. 2006. *Breaking the Biological Barriers to Cellulosic Ethanol: A Joint Research Agenda*, DOE/SC-0095. Prepared by Oak Ridge National Laboratory for the U.S. Department of Energy Office of Science and Office of Energy Efficiency and Renewable Energy (June 2006). genomicscience.energy.gov/biofuels/b2bworkshop.shtml.

U.S. DOE. 2004. *Top Value Added Chemicals from Biomass. Volume I—Results of Screening for Potential Candidates from Sugars and Synthesis Gas*, DOE/GO-102004-1992. Prepared by Pacific Northwest National Laboratory and National Renewable Energy Laboratory for the U.S. Department of Energy Office of Energy Efficiency and Renewable Energy (August 2004). www.osti.gov/scitech/biblio/926125.

van der Oost, J., et al. 2014. "Unravelling the Structural and Mechanistic Basis of CRISPR–Cas Systems," *Nature Reviews Microbiology* **12**, 479–92. DOI: 10.1038/nrmicro3279.

Vasilakoglou, I., et al. 2011. "Sweet Sorghum Productivity for Biofuels Under Increased Soil Salinity and Reduced Irrigation," *Field Crops Research* **120**(1), 38–46. DOI: 10.1016/j.fcr.2010.08.011.

Vink, E. T. H., et al. 2004. "The Sustainability of NatureWorks™ Polylactide Polymers and Ingeo™ Polylactide Fibers: An Update of the Future," *Macromolecular Bioscience* **4**(6) 551–64. DOI: 10.1002/mabi.200400023.

Wilkerson, C. G., et al. 2014. "Monolignol Ferulate Transferase Introduces Chemically Labile Linkages into the Lignin Backbone," *Science* **344**(6179), 90–93. DOI: 10.1126/science.1250161.

Ximenes, E., et al. 2011. "Deactivation of Cellulases by Phenols," *Enzyme and Microbial Technology* **48**(1), 54–60. DOI: 10.1016/j.enzmictec.2010.09.006.

Ximenes, E., et al. 2010. "Inhibition of Cellulases by Phenols," *Enzyme and Microbial Technology* **46**(3–4), 170–76. DOI: 10.1016/j.enzmictec.2009.11.001.

Yang, F., et al. 2013. "Engineering Secondary Cell Wall Deposition in Plants," *Plant Biotechnology Journal* **11**, 325–35. DOI: 10.1111/pbi.12016.

Yee, K. L., et al. 2014. "Consolidated Bioprocessing of Transgenic Switchgrass by an Engineered and Evolved *Clostridium thermocellum* Strain," *Biotechnology for Biofuels* **7**, 75. DOI: 10.1186/1754-6834-7-75.

Zakzeski, J., et al. 2010. "The Catalytic Valorization of Lignin for the Production of Renewable Chemicals," *Chemical Reviews* **110**(6), 3552–99. DOI: 10.1021/cr900354u.

Zalesny, R. S., Jr., et al. 2009. "Biomass and Genotype × Environment Interactions of *Populus* Energy Crops in the Midwestern United States," *BioEnergy Research* **2**(3), 106–22. DOI: 10.1007/s12155-009-9039-9.

Zhou, C.-H., et al. 2008. "Chemoselective Catalytic Conversion of Glycerol as a Biorenewable Source to Valuable Commodity Chemicals," *Chemical Society Reviews* **37**(3), 527–49. DOI: 10.1039/B707343G.

Zingaro, K. A., S. A. Nicolau, and E. T. Papoutsakis. 2013. "Dissecting the Assays to Assess Microbial Tolerance to Toxic Chemicals in Bioprocessing," *Trends in Biotechnology* **31**(11), 643–53. DOI: 10.1016/j.tibtech.2013.08.005.

Acronyms and Abbreviations

ABE	acetone-butanol-ethanol
AFEX™	ammonia fiber expansion
BER	DOE Office of Biological and Environmental Research
BNICE	Biochemical Network Integrated Computational Explorer
BRC	DOE Bioenergy Research Center
Cas9	CRISPR-associated protein 9
CBP	consolidated bioprocessing
CO	carbon monoxide
CO_2	carbon dioxide
CoA	coenzyme A
CRISPR	clustered regularly interspaced short palindromic repeats
DMAP	dimethyl-allyl pyrophosphate
DOE	U.S. Department of Energy
DXP	1-deoxy-D-xylulose-5-phosphate pathway
E-AFEX™	extractive AFEX™
EC	Enzyme Commission system (initiated by the Nomenclature Committee of the International Union of Biochemistry and Molecular Biology 1955)
EISA	Energy Independence and Security Act
EMSL	DOE Environmental Molecular Sciences Laboratory
GH	glycoside hydrolase
GWAS	genome-wide association study
H_2	hydrogen
HCl	hydrogen chloride
IL	ionic liquid
IPP	isopentenyl-pyrophosphate
JGI	DOE Joint Genome Institute
KBase	DOE Systems Biology Knowledgebase
MALDI	matrix-assisted laser desorption/ionization
MEV	mevalonate pathway
MEP	2-C-methyl-D-erythritol-4-phosphate pathway
NaOH	sodium hydroxide
NH_3	anhydrous ammonia
O_2	dioxygen
PDO	1,3-propanediol
PHB	polyhydroxybutyrate
RFS	Renewable Fuel Standard (see EISA)
TALEN	transcription activator-like effector nuclease
USDA	U.S. Department of Agriculture
WLP	Wood-Ljungdahl pathway
ZNF	zinc finger

February 2015