This PDF is available from The National Academies Press at http://www.nap.edu/catalog.php?record_id=19001

Industrialization of Biology: A Roadmap to Accelerate the Advanced Manufacturing of Chemicals

ISBN
978-0-309-31652-1

176 pages
6 x 9
PAPERBACK (2015)

Committee on Industrialization of Biology: A Roadmap to Accelerate the Advanced Manufacturing of Chemicals; Board on Chemical Sciences and Technology; Board on Life Sciences; Division on Earth and Life Studies; National Research Council

Add book to cart
Find similar titles
Share this PDF

Visit the National Academies Press online and register for...

- Instant access to free PDF downloads of titles from the
  - NATIONAL ACADEMY OF SCIENCES
  - NATIONAL ACADEMY OF ENGINEERING
  - INSTITUTE OF MEDICINE
  - NATIONAL RESEARCH COUNCIL
- 10% off print titles
- Custom notification of new releases in your field of interest
- Special offers and discounts

Distribution, posting, or copying of this PDF is strictly prohibited without written permission of the National Academies Press.
Unless otherwise indicated, all materials in this PDF are copyrighted by the National Academy of Sciences.
Request reprint permission for this book

Copyright © National Academy of Sciences. All rights reserved.

The National Academies of
SCIENCES · ENGINEERING · MEDICINE

Industrialization of Biology

A Roadmap to Accelerate the Advanced Manufacturing of Chemicals

Committee on Industrialization of Biology:
A Roadmap to Accelerate the Advanced Manufacturing of Chemicals

Board on Chemical Sciences and Technology

Board on Life Sciences

Division on Earth and Life Studies

NATIONAL RESEARCH COUNCIL
OF THE NATIONAL ACADEMIES

THE NATIONAL ACADEMIES PRESS
Washington, D.C.
www.nap.edu

Copyright © National Academy of Sciences. All rights reserved.

THE NATIONAL ACADEMIES PRESS 500 Fifth Street, NW Washington, DC 20001

NOTICE: The project that is the subject of this report was approved by the Governing Board of the National Research Council, whose members are drawn from the councils of the National Academy of Sciences, the National Academy of Engineering, and the Institute of Medicine. The members of the committee responsible for the report were chosen for their special competences and with regard for appropriate balance.

This study was supported by the U.S. Department of Energy under Grant DE-SC0010761 and the National Science Foundation under Grant CBET-1344363.

This report was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor any agency thereof, nor any of their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to a specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or agency thereof.

International Standard Book Number-13: 978-0-309-31652-1
International Standard Book Number-10: 0-309-31652-9
Library of Congress Control Number: 2015937241

Additional copies of the report are available from the National Academies Press, 500 Fifth Street, NW, Keck 360, Washington, DC 20001; (800) 624-6242 or (202) 334-3313; http://www.nap.edu.

Copyright 2015 by the National Academy of Sciences. All rights reserved.

Printed in the United States of America

Copyright © National Academy of Sciences. All rights reserved.

THE NATIONAL ACADEMIES

Advisers to the Nation on Science, Engineering, and Medicine

The **National Academy of Sciences** is a private, nonprofit, self-perpetuating society of distinguished scholars engaged in scientific and engineering research, dedicated to the furtherance of science and technology and to their use for the general welfare. Upon the authority of the charter granted to it by the Congress in 1863, the Academy has a mandate that requires it to advise the federal government on scientific and technical matters. Dr. Ralph J. Cicerone is president of the National Academy of Sciences.

The **National Academy of Engineering** was established in 1964, under the charter of the National Academy of Sciences, as a parallel organization of outstanding engineers. It is autonomous in its administration and in the selection of its members, sharing with the National Academy of Sciences the responsibility for advising the federal government. The National Academy of Engineering also sponsors engineering programs aimed at meeting national needs, encourages education and research, and recognizes the superior achievements of engineers. Dr. C. D. Mote, Jr., is president of the National Academy of Engineering.

The **Institute of Medicine** was established in 1970 by the National Academy of Sciences to secure the services of eminent members of appropriate professions in the examination of policy matters pertaining to the health of the public. The Institute acts under the responsibility given to the National Academy of Sciences by its congressional charter to be an adviser to the federal government and, upon its own initiative, to identify issues of medical care, research, and education. Dr. Victor J. Dzau is president of the Institute of Medicine.

The **National Research Council** was organized by the National Academy of Sciences in 1916 to associate the broad community of science and technology with the Academy's purposes of furthering knowledge and advising the federal government. Functioning in accordance with general policies determined by the Academy, the Council has become the principal operating agency of both the National Academy of Sciences and the National Academy of Engineering in providing services to the government, the public, and the scientific and engineering communities. The Council is administered jointly by both Academies and the Institute of Medicine. Dr. Ralph J. Cicerone and Dr. C. D. Mote, Jr., are chair and vice chair, respectively, of the National Research Council.

www.national-academies.org

Copyright © National Academy of Sciences. All rights reserved.

Copyright © National Academy of Sciences. All rights reserved.

COMMITTEE ON INDUSTRIALIZATION OF BIOLOGY: A ROADMAP TO ACCELERATE THE ADVANCED MANUFACTURING OF CHEMICALS

Members

THOMAS M. CONNELLY, JR. (Chair), E. I. du Pont de Nemours & Company (ret.)
MICHELLE C. CHANG, University of California, Berkeley
LIONEL CLARKE, UK Synthetic Biology Leadership Council
ANDREW D. ELLINGTON, University of Texas at Austin
NATHAN J. HILLSON, Lawrence Berkeley National Laboratory
RICHARD A. JOHNSON, Global Helix LLC
JAY D. KEASLING, University of California, Berkeley
STEPHEN S. LADERMAN, Agilent Technologies, Inc.
PILAR OSSORIO, University of Wisconsin Law School
KRISTALA L. J. PRATHER, Massachusetts Institute of Technology
RESHMA P. SHETTY, Ginkgo Bioworks, Inc.
CHRISTOPHER A. VOIGT, Massachusetts Institute of Technology
HUIMIN ZHAO, University of Illinois, Urbana-Champaign

National Research Council Staff

DOUGLAS FRIEDMAN, Study Director, Board on Chemical Sciences and Technology
INDIA HOOK-BARNARD, Senior Program Officer, Board on Life Sciences
CARL-GUSTAV ANDERSON, Research Associate
ELIZABETH FINKELMAN, Program Coordinator
NAWINA MATSHONA, Senior Program Assistant
JOHN SADOWSKI, Christine Mirzayan Science & Technology Policy Fellow (Winter 2014)

Copyright © National Academy of Sciences. All rights reserved.

BOARD ON CHEMICAL SCIENCES AND TECHNOLOGY

Members

TIMOTHY SWAGER, (Co-Chair), Massachusetts Institute of Technology
DAVID WALT, (Co-Chair), Tufts University
HÉCTOR D. ABRUÑA, Cornell University
JOEL C. BARRISH, Bristol-Myers Squibb
MARK A. BARTEAU, University of Michigan
DAVID BEM, The Dow Chemical Company
ROBERT G. BERGMAN, University of California, Berkeley
JOAN BRENNECKE, University of Notre Dame
HENRY E. BRYNDZA, E. I. du Pont de Nemours & Company
MICHELLE V. BUCHANAN, Oak Ridge National Laboratory
DAVID W. CHRISTIANSON, University of Pennsylvania
RICHARD EISENBERG, University of Rochester
JILL HRUBY, Sandia National Laboratories
FRANCES S. LIGLER, University of North Carolina, Chapel Hill, and North Carolina State University
SANDER G. MILLS, Merck Research Laboratories (ret.)
JOSEPH B. POWELL, Shell
ROBERT E. ROBERTS, Institute for Defense Analyses
PETER J. ROSSKY, Rice University

National Research Council Staff

TERESA FRYBERGER, Director
DOUGLAS FRIEDMAN, Senior Program Officer
KATHRYN HUGHES, Senior Program Officer
CAMLY TRAN, Postdoctoral Fellow
CARL-GUSTAV ANDERSON, Research Associate
ELIZABETH FINKELMAN, Program Coordinator
NAWINA MATSHONA, Senior Program Assistant
COTILYA BROWN, Senior Program Assistant

Copyright © National Academy of Sciences. All rights reserved.

BOARD ON LIFE SCIENCES

Members

JAMES P. COLLINS (*Chair*), Arizona State University
ENRIQUETA C. BOND, Burroughs Wellcome Fund (ret.)
ROGER D. CONE, Vanderbilt University Medical Center
JOSEPH R. ECKER, Salk Institute for Biological Studies
SEAN EDDY, HHMI Janelia Farm Research Campus
SARAH C. R. ELGIN, Washington University in St. Louis
DAVID R. FRANZ, Former Commander USAMRIID; Consultant
STEPHEN FRIEND, Sage Bionetworks
ELIZABETH HEITMAN, Vanderbilt University Medical Center
JOHN G. HILDEBRAND, University of Arizona
RICHARD A. JOHNSON, Global Helix LLC
JUDITH KIMBLE, University of Wisconsin, Madison
MARY E. MAXON, Science Philanthropy Alliance
KAREN E. NELSON, J. Craig Venter Institute
ROBERT M. NEREM, Georgia Institute of Technology
MARY E. POWER, University of California, Berkeley
MARGARET RILEY, University of Massachusetts, Amherst
LANA SKIRBOLL, Sanofi
JANIS C. WEEKS, University of Oregon
MARY WOOLLEY, Research!America

Staff

FRANCES E. SHARPLES, Director
JO L. HUSBANDS, Scholar/Senior Project Director
JAY B. LABOV, Senior Scientist/Program Director for Biology Education
KATHERINE W. BOWMAN, Senior Program Officer
MARILEE K. SHELTON-DAVENPORT, Senior Program Officer
KEEGAN SAWYER, Program Officer
AUDREY THEVENON, Associate Program Officer
BETHELHEM MEKASHA, Financial Associate
ANGELA KOLESNIKOVA, Administrative Assistant
P. KANOKO MAEDA, Senior Project Assistant
JENNA OGILVIE, Senior Project Assistant

Copyright © National Academy of Sciences. All rights reserved.

Copyright © National Academy of Sciences. All rights reserved.

Preface

The efficient production of useful and beneficial goods and services has been the cornerstone of industrial development, driving economic growth for more than two centuries. Throughout this period, the underpinning technologies driving industrialization have evolved in response to new scientific understanding, new technological capabilities, and new market demands. Insights into the chemical nature of matter, reaction mechanisms, and the role of physical and catalytic processes transformed the industrial landscape during the 19th century. By the early 20th century, a new understanding of chemistry transformed crude oil into a feedstock for a vast array of chemical products ranging from plastics and paints to detergents and textiles—transforming nearly every aspect of our lives.

Today, we are at a new inflection point. The tremendous progress in biology over the past half century—from Watson and Crick's elucidation of the structure of DNA to today's astonishing, rapid progress in the field of synthetic biology—has positioned us for the new round of innovation in chemical production. This observation provided the impetus for this study, commissioned by the U.S. Department of Energy and the National Science Foundation. Our committee was charged with understanding how to accelerate biological production of chemicals and also to create a roadmap to that future.

The committee of 13 members (Appendix C) convened from approximately February 2014 through December 2014 and met in person four times. Expertise included synthetic biology, metabolic engineering, molecular biology, microbiology, systems biology, synthetic chemistry, chemical

Copyright © National Academy of Sciences. All rights reserved.

engineering, bioinformatics, systems integration, metrology, chemical manufacturing, and law and bioethics. The committee heard from researchers at the leading edge of microbial biotechnology and from industry leaders, including large, established chemical companies and technology-rich start-ups. We had dialogue with representatives of U.S. government agencies and with nongovernment organizations. In May 2014 the committee held a 2-day workshop (Appendix D), which laid the foundation for the conclusions, recommendations, and roadmap found in this report.

Any roadmap is an ephemeral guide—a snapshot in time. The committee took care to set ambitious goals that emphasize outcomes over individual technologies. As science and technology advance and economic circumstances change, it is often the road-mapping process that can provide lasting value. This observation, coupled with the broad, outcome-oriented goals herein, led the committee to discuss the road-mapping process as a continuing activity that the sponsoring agencies may wish to consider on a regular basis in order to ensure acceleration of this field and maintenance of the roadmap in a living, evergreen process.

As stated in the National Bioeconomy Blueprint released in 2012, "[e]conomic activity that is fueled by research and innovation in the biological science, the 'bioeconomy,' is a large and rapidly growing segment of the world economy that provides substantial public benefit." The picture that emerged through the course of the study was that of a field with tremendous potential for innovation, economic impact, and great discovery—if only we can accelerate its maturity.

Thomas M. Connelly, Jr., *Chair*

Copyright © National Academy of Sciences. All rights reserved.

Acknowledgment of Reviewers

This report has been reviewed in draft form by individuals chosen for their diverse perspectives and technical expertise, in accordance with procedures approved by the National Research Council's Report Review Committee. The purpose of this independent review is to provide candid and critical comments that will assist the institution in making its published report as sound as possible and to ensure that the report meets institutional standards for objectivity, evidence, and responsiveness to the study charge. The review comments and draft manuscript remain confidential to protect the integrity of the deliberative process.

We wish to thank the following individuals for their review of this report:

Scott Baker, Pacific Northwest National Laboratory
Sean Eddy, HHMI Janelia Farm Research Campus
Jennifer Holmgren, LanzaTech
Sang Yup Lee, KAIST
James Liao, University of California, Los Angeles
Richard Murray, California Institute of Technology
Kathie Olsen, ScienceWorks, LLC
Markus Pompejus, BASF Corporation

Although the reviewers listed above have provided many constructive comments and suggestions, they were not asked to endorse the conclusions or recommendations, nor did they see the final draft of the report

Copyright © National Academy of Sciences. All rights reserved.

before its release. The review of this report was overseen by **Klavs Jensen** of the Massachusetts Institute of Technology and **Michael Ladisch** of Purdue University. Appointed by the National Research Council, they were responsible for making certain that an independent examination of this report was carried out in accordance with institutional procedures and that all review comments were carefully considered. Responsibility for the final content of this report rests entirely with the authoring committee and the institution.

Copyright © National Academy of Sciences. All rights reserved.

Contents

SUMMARY **1**
The Potential of Industrial Biotechnology, 1
Why Now?, 3
The Science Is Advancing, 3
Industry Is Ready, 3
The Gaps Have Been Characterized, 5
A Vision of the Future, 6
Technical Conclusions, Recommendations, and Roadmap Goals, 6
Nontechnical Insights and Societal Concerns, 8
How Do We Get There?, 12

1 INTRODUCTION AND CONTEXT **13**
Charge to the Committee and Interpretation of Scope, 15
Definitions, 16
Chemical Manufacturing, 18
Tools and Technologies, 20
Societal Factors, 21
Organization of the Report, 21

2 INDUSTRIAL BIOTECHNOLOGY: PAST AND PRESENT **25**
The Bioeconomy and Global Challenges, 25
Bio-based Markets Already Are Significant and Thriving, 26
The Innovative Power of Industrial Biotechnology, 27

Copyright © National Academy of Sciences. All rights reserved.

Enabling the Industrialization of Biology and the Development of an Ambitious Roadmap for Accelerating the Advanced Manufacturing of Chemicals, 30
Convergence, 30
Societal Benefits in Addressing Global Grand Challenges, 32
Energy, 33
Climate Change and Environmental Sustainability, 33
Agriculture, 34
Competitiveness and Innovation, 34
The Time Is Right: Current State and Advances in Science and Industry, 35
Opportunities Arising from DNA Technologies, Systems Biology, Metagenomics, and Synthetic Biology, 35
New High-Value Chemical Products Unobtainable by Traditional Chemical Synthesis, 37
Implementing Computation in Cells, 38
Industry Is Ready, 39
A Few Examples, 42
Artemisinin, 42
Biofuels: Moving to Commercial, 44
1,4-Butanediol (BDO), 45
Industrial Enzymes, 47
Governance Framework, 47

3 VISION OF THE FUTURE: WHAT NEW CHEMICALS COULD BE MADE? 53
What Chemicals Could Be Made?, 53
Natural Products, 56
Genes to Products, 56
Natural Product Analogs, 57
Tapping New Structural Diversity, 57
Advanced Molecules, 57
Engineering the Production of Complex Building Blocks, 58
Engineering the Stereo- and Regioselective Transformation of Synthetic Building Blocks, 59
Catalysis with Key Functional Groups and for New C-C Bond-Making Chemistry, 59
Polymers, 60
Existing Monomers, 60
New Monomers, 61
Polymerases, 63

Copyright © National Academy of Sciences. All rights reserved.

Polymers for Templating the Formation of Inorganic Materials, 63
Business Models for Future Industrial Biotechnology, 64

4 HOW DO WE GET THERE? 67
Overview of Issues, 67
Feedstocks, 68
New Sources of Carbon, 68
Multiple Generations of Feedstocks, 69
Grain-Derived Sugars, 69
Lignocellulosic Biomass, 71
C1 Feedstocks, 73
Enabling Transformations, 74
Fermentation and Processing, 74
Fermentation, 75
Scaling, 77
Enzyme-Mediated Reactions, 77
Cell-Free Processing, 78
Additional Bioprocessing Operations, 79
Organism, 80
Introduction: The Design-Build-Test-Learn Loop, 82
Fully Integrated Design Toolchain, 82
Design, 85
Pathway Design, 85
Enzyme Design, 86
Systems Biology Design, 86
Bioprocess Design, 88
Build, 88
Pathways, 88
Chassis, 91
Test and Measurement, 95

5 WHAT IS SUCCESS AND HOW TO GET THERE: RECOMMENDATIONS 101
How Do We Get There?, 102
Technical Needs and Roadmap, 104
Nontechnical Insights and Societal Concerns, 106
Economic, 106
Education and Workforce, 106
Governance, 108
Concluding Remarks, 109

REFERENCES 111

Copyright © National Academy of Sciences. All rights reserved.

APPENDIXES

A Glossary 121
B The Current Regulatory Framework 125
C Committee Member and Staff Biographies 133
D Workshop Agenda and Attendees 141

Copyright © National Academy of Sciences. All rights reserved.

Summary

In response to a request from the U.S. Department of Energy and the National Science Foundation, the National Research Council convened an ad hoc committee to create a roadmap for accelerating the advanced manufacturing of chemicals using biological systems. The committee was charged to "develop a roadmap of necessary advances in basic science and engineering capabilities, including knowledge, tools and skills," while "working at the interface of synthetic chemistry, metabolic engineering, molecular biology and synthetic biology" and "considering when and how to integrate non-technological insights and societal concerns into the pursuit of the technical challenges." The full statement of task can be found in Box 1-1. While the central focus of this report and roadmap is on industrial biotechnology, many of the roadmap goals, conclusions, and recommendations herein will also benefit other sectors, including health, energy, and agriculture.

THE POTENTIAL OF INDUSTRIAL BIOTECHNOLOGY

In its 2012 National Bioeconomy Blueprint, the Obama Administration defined the bioeconomy simply as "one based on the use of research and innovation in the biological sciences to create economic activity and public benefit." It went on to observe that "[t]he U.S. bioeconomy is all around us," with new bio-based chemicals, improved public health through improved drugs and diagnostics, and biofuels that reduce our dependency on oil.[1]

Copyright © National Academy of Sciences. All rights reserved.

Bio-based product markets are already significant in the United States—representing more than 2.2 percent of gross domestic product, or more than $353 billion in economic activity in 2012.[2] While biotechnology has had its greatest economic impact, to date, in human health and in agriculture, bio-based chemicals are neither entirely new, nor are they a historic artifact. Current global bio-based chemical and polymer production is already estimated to be about 50 million tons each year, and bioprocessing techniques (such as fermentation, baking, and tanning) have been used throughout much of human industrial history.

Agilent Technologies estimates that U.S. business-to-business revenues from industrial biotechnology alone reached at least $125 billion in 2012.[2b] Bio-based chemical applications accounted for about $66 billion of that activity with biofuels adding another $30 billion. Lux Research estimates that industrial chemicals made through synthetic biology currently represent a $1.5 billion market and that this likely will expand at a 15 to 25 percent annual growth rate for the foreseeable future.[3] Based on a 2009 Organisation for Economic Co-operation and Development (OECD) analysis, a recent U.S. Department of Agriculture (USDA) report indicates that, this year, bio-based chemicals will comprise greater than 10 percent of the chemical market.[4]

Despite this impressive recent and projected growth, the manufacturing of chemicals using biological synthesis and engineering could expand even faster. Today, many of the chemicals being produced are selected, in part, because well-established chemical syntheses toward them already exist. In many cases, bio-based routes are often not even considered. Yet the addition of bio-based routes to chemicals could open the door to making and marketing chemicals that cannot presently be made at scale or may allow the use of new classes of feedstocks. This report examines the technical, economic, and societal factors that limit the adoption of bioprocessing in the chemical industry today and that, if surmounted, would markedly accelerate the advanced manufacturing of chemicals via industrial biotechnology and the benefits that would accrue.

The advanced manufacturing of chemicals through biology can help address global challenges related to energy, climate change, sustainable and more productive agriculture, and environmental sustainability. For example, these processes may help reduce toxic by-products, greenhouse gas emissions, and fossil fuel consumption in chemical production. Lowered costs, increases in production speed, flexibility of manufacturing plants, and increased production capacity are among the many potential benefits that the increased industrialization of biology may bring to producers and consumers of chemical products that have not been previously available at scale.

Copyright © National Academy of Sciences. All rights reserved.

WHY NOW?

The Science Is Advancing

The genetics underlying the natural world are being illuminated by DNA sequencing, the cost of which is declining rapidly.[5] The first human genome (3.2 billion base pairs [bp]) was sequenced in 2001 at a cost of $2.7 billion.[6] Nine years later 1,000 human genomes (3.2 trillion bp) were sequenced, and in 2014 the company Illumina released the HiSeq X, promising a $1,000 human genome.[7] Databases of sequences have rapidly grown; as of 2013, there were 160 million sequences from 300,000 organisms.[8] This growth has built an enormous potential catalogue of natural "parts"—functional units of DNA—from which high-value chemical pathways can be discovered or created.

The past decade has seen an explosion in the technologies to compose, read, write, and debug DNA. This has rapidly increased the scale and sophistication of genetic engineering projects, and in the near term this will lead to more complex chemical structures and composite nanomaterials, which require precise control over dozens of genes. Examples of this include mining drug candidates from the human microbiome, pesticides from environmental samples, and the production of metal nanoparticles for electronics and medical devices. In the longer term, one can imagine organisms designed from the ground up for consolidated bioprocessing and automated product assembly that requires multiple steps to synthesize relevant industrial chemicals.

The ability to compose, or decide the sequence of, DNA has lagged behind our ability to read and write it. The most valuable functions require many genes and complex regulatory control over how much, when, and where they are turned on. Synthetic biologists pursue the creation of important tools to solve this problem, including genetic circuits, precision gene regulation parts, and computer-aided design to systematically recode multigene systems. Although it is possible to synthesize entire genomes, we are far from being able to write them from scratch from the bottom up. The current state of the art is the top-down "editing" of existing genomes using technologies such as MAGE[9] and CRISPR/Cas9[10] to introduce incremental changes in an otherwise natural genome. Similarly, genome-scale design tools have begun to emerge to control flux through metabolic pathways.

Industry Is Ready

The applications of synthetic biology in human health and agriculture have advanced more quickly than the manufacturing of chemicals. As a result, groundwork has been laid for the manipulation of genes and

Copyright © National Academy of Sciences. All rights reserved.

proteins to beneficial purposes and for the scaling of bioprocesses to large volumes. For human health applications, therapeutic proteins are more structurally complex than the small molecules that make up most important industrial chemicals. Their synthesis, however, is directly related to the DNA chosen for expression; simple overexpression in the right host of as little as a single gene produces the product of interest.

Agricultural applications of biotechnology involve the introduction and regulation of a small number of genes. Typically, one or two genes are introduced to confer each desired property (e.g., herbicide tolerance, insect resistance, or disease resistance). Agricultural uses are complicated by the need to express the genes in the tissues of a plant, without adverse phenotypic responses such as slower growth or reduced yield. That transgenic plants are grown in an open environment increases the scope of regulatory controls.

In contrast to health and agriculture applications, synthesis of a chemical product requires the coordination of the expressions of many genes. Biologically produced chemicals are the result of a series of enzyme-catalyzed reactions, with each enzyme encoded by at least one gene. In total, the expression of as many as dozens of genes must be regulated to affect a chemical synthesis. This complexity of the pathways involved creates a systems-level challenge that requires systems-oriented solutions. Biological engineering seeks to take advantage of the tools of recombinant DNA technology while applying systems and network analyses to the challenge of engineering more productive host organisms. These principles have already been successfully applied to generate highly efficient and productive fermentation processes for a number of products. Early successes include, for example, the production of industrial enzymes, artemisinin, lactic acid, 1,3-propanediol, isoprenoids, and alcohol-based biofuels.

Based on these early successes, and powered by the rapidly developing science, use of industrial biology to produce a broad range of chemical products is likely to continue to accelerate. The growth of this field will enable the use of biology to produce high-valued chemical products that cannot be produced at high purity and high yield through traditional chemical synthesis. The future may also include a large number of high-volume chemicals, where biology represents a better synthetic pathway (cheaper and greener) than the conventional chemical synthesis.

In the future production of chemicals, industrial chemical synthesis will frequently take advantage of both biosynthesis and traditional chemical synthetic steps, employing each so as to optimize the overall synthetic pathway.

Copyright © National Academy of Sciences. All rights reserved.

The Gaps Have Been Characterized

Achieving a future where biosynthesis and traditional chemical synthesis are equally viable candidates in the industrial production of chemicals requires closing several scientific, technical, and societal gaps. This report identifies feedstock design and use, fermentation and processing, enabling chemical transformations, and governance and societal factors as critical areas in its roadmap and recommendations. Scientific and engineering challenges remain, particularly in the areas of feedstocks, enabling transformations, and the development of an integrated design toolchain.

Today, the feedstock for biomanufacturing chemicals is fermentable sugars from starch. The starch, in turn, derives from grains such as corn. The continued expansion of biomanufacturing chemicals will require additional feedstocks from nongrain sources. Cellulosic biomass holds great promise as a feedstock, but there are still many challenges associated with using recalcitrant cellulosic material in industrial biotechnology. While much current attention is focused on different forms of biomass, there is also significant active work in facilitating the use of syngas, methane, and carbon dioxide in manufacturing.

One of the major engineering considerations is related to fermentation and processing that is required for production of biological systems. Fermentation can be facilitated in many ways, but it typically represents a large capital expense that must be overcome in order to begin production. To mitigate this capital expense, the ability to scale up processes is a critical step. While fermentation is typically conducted batchwise or in "fed batch mode," developments such as continuous fermentation, continuous product removal, and cell-free processing are needed for rapid improvement.

Further research and development is needed to facilitate chemical transformations. The dramatic advances in synthetic biology are at the heart of chemical manufacturing via biological synthesis and engineering. Continued progress is needed in both the organismal "chassis" and the metabolic pathways of the microorganisms used in chemical manufacturing. In addition, the number and range of microorganisms "domesticated" for industrial use will need to increase with the diversity of products manufactured.

A number of governance and societal factors will also influence the rate of industrialization of biology. Governance starts with the establishment of industry norms and standards that are needed for industrial biology value chains to be established and for economic exchange to occur. Such standards are needed in areas such as (1) read/write accuracy for DNA; (2) DNA "part" performance specifications; (3) data and machine standards across "-omics" technologies; and (4) organism performance in terms of production rates, titers, and yields.

Copyright © National Academy of Sciences. All rights reserved.

Beyond standards, an updated regulatory regime is needed to speed the safe commercialization of new host organisms, new metabolic pathways, and new chemical products. Such regimes must be harmonized across national boundaries, enabling rapid, safe, and global access to new technologies and products. It must be recognized that ultimately it is society that confers the right to operate new technologies. Efforts are needed to inform the public of the nature of industrial biotechnology and of its societal benefits, and to make sure that public concerns are communicated effectively.

Finally, a roadmap should be an evergreen document. A mechanism is needed to maintain this roadmap, to sustain the momentum, and to ensure that the complex network of technical, economic, and societal factors is progressing in harmony as we build the industrial biology ecosystem.

A VISION OF THE FUTURE

The vision of the future put forth herein is one where biological synthesis and engineering and chemical synthesis and engineering are on par with one another for chemical manufacturing. The current capabilities of traditional chemical manufacturing are vast, but limit the types of chemicals that can be produced at scale (see Chapter 3). Furthermore, the core petroleum-based feedstock is a limited resource and diversification of feedstocks will provide even greater opportunity for the chemical manufacturing industry.

The recommendations and roadmap goals outlined throughout this report were all conceived in the context of this vision and are designed with the understanding that, in order for the industrialization of biology to be fully realized, the use of biological and chemical routes must be thought of as equals. That does not imply that each would be used interchangeably, but rather that biological options would be considered in the same way individual chemical reactions are considered when developing a synthetic route. The following conclusions, recommendations, and roadmap goals given in Tables S-1 and S-2 are aligned to help achieve this major goal.

Technical Conclusions, Recommendations, and Roadmap Goals

There are many areas of science and engineering that must be advanced to accelerate the industrialization of biology. The roadmap items and categories are all in the context of the core technical conclusion: *Biomanufacturing of chemicals is already a significant element of the national economy and is poised for rapid growth during the next decade. Both the scale*

Copyright © National Academy of Sciences. All rights reserved.

and scope of biomanufacturing of chemicals will expand and will involve both high-value and high-volume chemicals. Progress in the areas identified in this report will play a major role in achieving the challenge of increasing the contribution of biotechnology to the national economy. While the roadmap is clearly designed to push forward industrial biotechnology, there are many aspects of fundamental research that are needed, and described in this report, that can be applied broadly to other fields, such as health, energy, and agriculture.

The technical roadmap is broken down into six main categories that follow along the production model outlined in the chemical manufacturing flowchart (Figure 1-1). They are:

1. Feedstocks and Pre-Processing;
2. Fermentation and Processing;
3. Design Toolchain;
4. Organism: Chassis;
5. Organism: Pathways; and
6. Test and Measurement.

Each category contains a set of conclusions (Table S-1) leading to Roadmap Goals (Figure S-1) that would represent a step change in the field. It is important to note that not all roadmap goals are geared toward all manufacturing sectors. For example, the roadmap goals for feedstocks assume that feedstock cost is a major component of overall production costs, as it is for fuels or other high-volume chemicals. In order to be competitive with current manufacturing costs, the cost of feedstocks needs to be reduced and choice of feedstocks diversified. Similarly, reducing the quantity of process water used in bioprocessing will not only reduce costs, but also serve to create a more environmentally friendly production process. This too, will focus largely on high-volume materials.

By contrast, the roadmap goals for organism (chassis and pathways) and design toolchain will benefit lower-volume, higher-value chemicals, including pharmaceuticals, where one may have to rely on developing newer pathways to generate higher value. Much of the basic research that will be invested herein not only will be applicable to industrial biotechnology, but also will have implications for health, energy, and agriculture as well.

The following recommendation is central to the success of the proposed roadmap: **In order to transform the pace of industrial biotechnology by enabling commercial entities to develop new biomanufacturing processes, the committee recommends that the National Science Foundation, U.S. Department of Energy, National Institutes of Health, National Institute of Standards and Technology, U.S. Department of Defense,**

Copyright © National Academy of Sciences. All rights reserved.

Industrialization of Biology: A Roadmap to Accelerate the Advanced Manufacturing of Chemicals

and other relevant agencies support the scientific research and foundational technologies required to advance and to integrate the areas of feedstocks, organismal chassis and pathway development, fermentation, and processing as outlined in the roadmap goals.

Nontechnical Insights and Societal Concerns

In addition to the technical roadmap, recommendations, and conclusions, a number of nontechnical insights and societal concerns are important to ensuring the success of this roadmap. In light of this issue and to better enable implementation of the technical goals set forth, a series of recommendations relating to Economic, Education and Workforce, and Governance issues are shown in Table S-2. As an example, this and many other reports discuss the bioeconomy and its contribution to the overall economy on several occasions; however, the term "bioeconomy" is poorly defined and can lead to confusion. A formal, quantitative measure of the bioeconomy would allow all stakeholders to speak on the same terms and focus on enabling technical solutions. It would also provide a benchmark for measuring improvement in the industrial biotechnology sector.

TABLE S-1
Technical Conclusions

Feedstocks and Pre-Processing
• Improvements in availability of economically feasible and environmentally sustainable feedstocks are necessary to accelerate the production of fuels and high-volume chemicals via bioprocessing.
• Improvements in the availability, reliability, and sustainability of biofeedstocks, including — cellulosic feedstocks from plants, including plants engineered for biomanufacturing with special attention to low-cost saccharification; — full use of lignin co-product from feedstocks; — utilization of dilute sugar streams; — ability to convert complex feedstocks into clean, fungible, usable intermediates via biological pathways; — dramatic lowering of environmental impact; — utilization of methane, methane derivatives, carbon dioxide, and formate as feedstocks; and — use of noncarbon feedstocks (e.g., metals, silicon) would increase the range of economically viable products, provide more predictive levels and quality of feedstock, and lower barriers to entry into the biological production of chemicals.
• Improving the basic understanding of C1-based fermentation, including both host organism and fermentation processes, would enable greater feedstock diversity in light of the increased availability of natural gas in the United States.

Copyright © National Academy of Sciences. All rights reserved.

TABLE S-1 Continued

Fermentation and Processing
• Aerobic, fed-batch, monoculture fermentation has been the dominant process for bioproduction of chemicals for many decades. Successful improvement efforts have focused on more productive host organisms. Little research has been conducted to improve the productivity of the fermentation process, by means of enhanced mass and heat transfer, continuous product removal, and more extensive use of co-cultures, co-products, and co-substrates. • The development of predictive computational tools based on small-scale experimental models that realistically predict performance at scale would accelerate the development of new products and processes for the production of chemicals via industrial biotechnology. • Unlike many traditional chemical processes, industrial biotechnology generates large aqueous process streams that require efficient mechanisms for product isolation and for efficient water reuse.
Design Toolchain
• The development and use of a robust integrated design toolchain across all scales of the process—individual cells, cells inside reactor, and the fermentation reactor itself—is an important step in bringing biomanufacturing onto the same level as traditional chemical manufacturing. • The development of predictive modeling tools within and for integration across all scales of the process—individual cells, cells inside reactor, and the fermentation reactor itself—would accelerate the development of new products and processes for the production of chemicals via industrial biotechnology.
Organism: Pathways
• Improvements in the ability to rapidly design enzymes with respect to catalytic activity and specific activity and engineer their biophysical and catalytic properties would significantly reduce the costs associated with biomanufacturing and scale-up.
Organism: Chassis
• Continued development of fundamental science and enabling technologies is required for the rapid and efficient development of organismal chassis and pathways. • Expanding the palette of domesticated microbial and cell-free platforms for biomanufacturing is critical to expanding the repertoire of feedstocks and chemicals accessible via bio-based manufacturing. • The design, creation, and cultivation of robust strains that remain genetically stable and retain performance stability over time in the presence of diverse feedstocks and products will reduce the costs involved in the use and scaling of biological production.
Test and Measurement
• The ability to rapidly, routinely, and reproducibly measure pathway function and cellular physiology will drive the development of novel enzymes and pathways, which are needed to increase the array of efficient and low-cost chemical transformations available for use in biomanufacturing. • The fall in cost and increase in throughput of measurement technologies should track that of strain engineering technologies and vice versa.

Copyright © National Academy of Sciences. All rights reserved.

Industrialization of Biology: A Roadmap to Accelerate the Advanced Manufacturing of Chemicals

TABLE S-2
Nontechnical Insights and Societal Concerns

Recommendation: Economic
• The U.S. government should perform a regular quantitative measure of the contribution of bio-based production processes to the U.S. economy to develop a capacity for forecasting and assessing economic impact.
Recommendations: Education and Workforce
• Industrial biotechnology firms individually, and especially through industry groups, should strengthen their partnerships with all levels of academia, from community colleges, undergraduate institutions, and graduate institutions, to communicate changing needs and practices in industry in order to inform academic instruction. • Federal agencies, academia, and industry should devise and support innovative approaches toward expanding the exposure of student trainees to design-build-test-learn paradigms in a high-throughput fashion and at industrial scale.
Recommendations: Governance
• The administration should ensure that the Environmental Protection Agency (EPA), U.S. Department of Commerce, USDA, Food and Drug Administration (FDA), Occupational Safety and Health Administration (OSHA), National Institute of Standards and Technology (NIST), and other relevant agencies work together to broadly assess, and regularly reassess, the adequacy of existing governance, including but not limited to regulation, and to identify places where industry, academia, and the public can contribute to or participate in governance. • Science funding agencies and science policy offices should ensure outreach efforts that facilitate responsible innovation by enabling the extension of existing relevant regulatory practices, concordance across countries, and increased public engagement. • Government agencies, including EPA, USDA, FDA, and NIST, should establish programs for both the development of fact-based standards and metrology for risk assessment in industrial biotechnology and programs for the use of these fact-based assessments in evaluating and updating the governance regime.

Consideration of the educational and workforce needs as the bio-economy expands the needs of industry and academia will change as well. It is important that the broader stakeholder community come together to determine future needs and strengthen partnerships broadly. Finally, as with any growing field, a series of governance challenges have emerged. First, engagement with the public will be a key factor in the acceptance of the technology and the conveying industries right to operate, as has been started with many groups in the United Kingdom and United States. Secondly, key government stakeholders will have to address and ensure that governance needs are being met, and continually assess whether the correct stance is being taken. Finally, in order for the community to work

Copyright © National Academy of Sciences. All rights reserved.

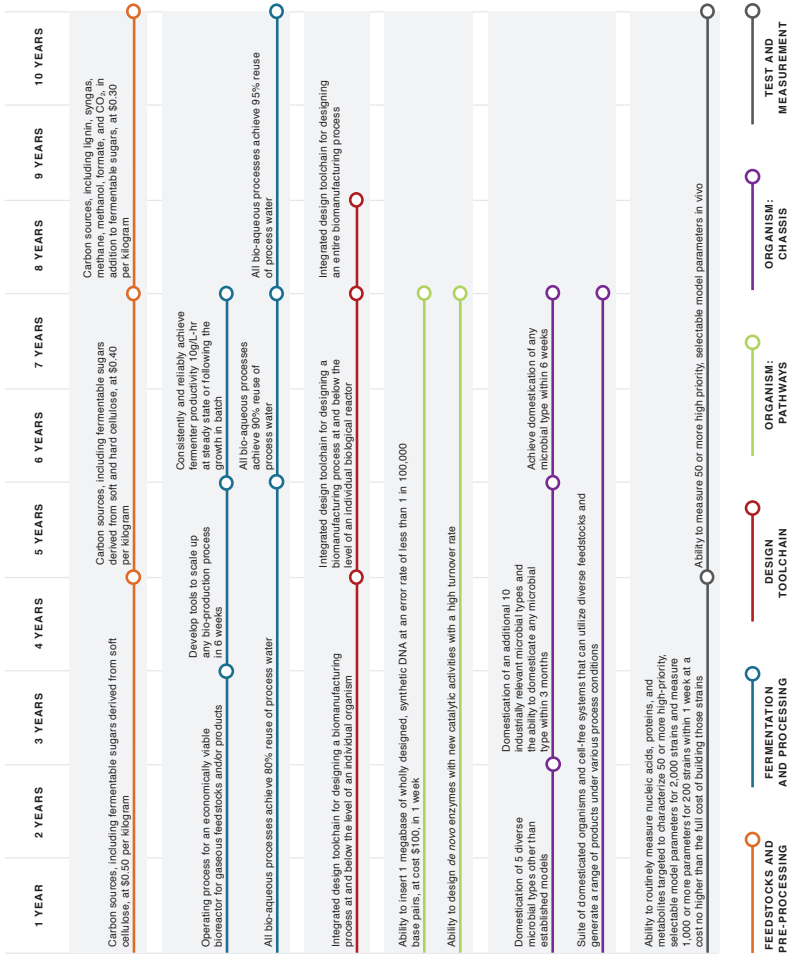


FIGURE S-1 Technical roadmap to enable the industrialization of biology.
NOTE: A larger version of this roadmap can be found as a foldout at the end of this book.

Copyright © National Academy of Sciences. All rights reserved.

together, the development of fact-based standards will be an important step forward.

HOW DO WE GET THERE?

Biomanufacturing of chemicals is already a significant element of the national economy, and it is poised for rapid growth during the next decade. Both the scale and scope of biomanufacturing of chemicals will expand and will involve both high-value and high-volume chemicals. High-value chemicals will benefit from the specificity of biological synthesis, leading to high-purity products, produced at high yield via pathways that minimize by-product formation. Large-volume chemicals must be produced in a cost efficient manner, taking advantage of cheap, abundant carbon sources, while minimizing the capital costs for the production facilities.

However, the realization of the promise of the industrialization of biology for chemical manufacturing can only be achieved through a sustained effort among multiple stakeholders. The next decade will be critical to the realization of the promise. Therefore**, the Committee recommends that the relevant government agencies consider establishment of an ongoing road-mapping mechanism to provide direction to technology development, translation, and commercialization at scale.**

As outlined in Chapter 5, a road-mapping activity, maintained in an evergreen fashion, could serve as a catalyst for many of the roadmap goals and recommendations in this report and could foster productive collaborations among diverse stakeholder groups. Examples are provided illustrating how this approach could be applied.

Copyright © National Academy of Sciences. All rights reserved.

1

Introduction and Context

The efficient production of useful and beneficial goods and services has been the cornerstone of industrial development, driving economic growth for more than two centuries.[11] Throughout this period, the underpinning technologies driving industrialization have evolved in response to new scientific understanding, new technological capabilities, and new market demands. Insights into the chemical nature of matter, reaction mechanisms, and the role of physical and catalytic processes transformed the industrial landscape during the 19th century. By 1882, dyes such as indigo, previously extracted from natural substances and relying on significant manual labor, could now be synthesized and made affordable. Medicines such as aspirin were similarly isolated and synthesized, making them affordable and widely available. By the early 20th century, a new understanding of chemistry transformed crude oil into a feedstock for a vast array of chemical products ranging from plastics and paints to detergents and textiles. Discoveries in physics early in the 20th century also entered the industrial landscape, leading to electronics, computers, satellites, and mobile communications, transforming economies, cultures, and the global community.

The human use and improvement of biological processes is an ancient and vital contributor to human progress, from the earliest periods of domestication of crops and animals through the agricultural revolution to the contemporary world of life sciences. Until recently, however, it has remained an essentially empirical pursuit because of the seemingly impenetrable complexity of biological systems.

Much of the underlying science for the establishment of industrial

Copyright © National Academy of Sciences. All rights reserved.

biotechnology began to emerge in the middle of the 20th century, particularly stemming from the discovery in 1953 of the structure of DNA by Crick and Watson and the realization that DNA's double-helix structure provided a unique mechanism for encoding information. In the decades following this discovery, significant progress has been made in understanding the relationship between these underlying biological building blocks and the functional performance of biological systems. The arrival of increasingly rapid computers and massive data handling capacity at the beginning of the 21st century facilitated the translation of data derived from high-throughput screening methods into more robust and predictive design techniques. The convergence of the life sciences with chemistry, chemical engineering, computer science, and other disciplines has increased the potential for industrialization of the biological sciences for chemical manufacturing.

The Organisation for Economic Co-operation and Development (OECD) first defined the bioeconomy as linking renewable biological resources and bioprocesses through industrial-scale biotechnologies and manufacturing to produce sustainable products, jobs, and income.[12] In its 2012 National Bioeconomy Blueprint, the Obama Administration redefined the bioeconomy simply as "one based on the use of research and innovation in the biological sciences to create economic activity and public benefit."[1] It went on to observe that the U.S. bioeconomy is "all around us" with new bio-based chemicals, improved public health through improved drugs and diagnostics, and biofuels that reduce our dependency on oil.

A proactive strategy—implemented through the development of a technical roadmap similar to those that enabled sustained growth in the semiconductor industry and our explorations of space—is needed if we are to realize the widespread benefits of accelerating the industrialization of biology.

A confluence of overlapping developments has created the conditions for making this achievable: the proliferation of emerging tools, technologies, and computational models; new investment opportunities and financial instruments; exciting new insights from scientific convergence and transdisciplinary research; innovative business models and entrepreneurial enterprises (large and small); new platforms for designing biological systems for next-generation American manufacturing; and novel opportunities to enhance competitiveness and create well-paying jobs. These trends, in turn, will transform existing chemical production, create new chemical and other sectors enabled by the industrialization of biology, and open a range of new markets for bio-based products resulting from advanced chemical manufacturing.

The roadmap proposed in this report underscores the widely held view that 21st-century innovation increasingly will rely on biology and, in

Copyright © National Academy of Sciences. All rights reserved.

particular, the convergence of biology with engineering and physical sciences such as chemistry. As a National Research Council report predicted: "Discoveries at all levels of biology will reverberate throughout science and provide the transformational insights that will lead to practical solutions in seemingly unrelated research areas."[13] A roadmap for accelerating advanced chemical manufacturing through the industrialization of biology begins to operationalize President Obama's 2011 observation that "[t]he world is shifting to an innovation economy and nobody does innovation better than America."[14]

CHARGE TO THE COMMITTEE AND INTERPRETATION OF SCOPE

At the request of the National Science Foundation and the U.S. Department of Energy, the National Research Council appointed an ad hoc committee with a broad range of expertise to identify key technical milestones for chemical manufacturing through biological routes. The committee's task included several key components: (1) the identification of the core scientific and technical challenges; (2) the identification of and timeline for the development of tools, measurement techniques, databases, and computational techniques needed to serve as the building blocks for research and applications; (3) a discussion of how to develop, share, and diffuse common interoperable standards, languages, and measurements; and (4) when and how to integrate nontechnological insights and societal concerns into the pursuit of the technical challenges (Box 1-1).

To address this task, the committee held a 2-day workshop in Washington, DC, to gather input from a range of experts and stakeholders. Speakers provided perspectives from the chemical industry process and experiences of scaling up (or out) production; insight into challenges in biosafety and biocontainment; and technical discussions of synthesis and genome-scale engineering, measurement, computer-aided design, and advanced molecules. Insights from this workshop served as a groundwork for the committee's deliberations, with additional data gathering occurring throughout the study process.

The committee identified three dimensions that will require progress to ensure the acceleration of the industrialization of biology:

1. the selection of the right chemical, material, and fuel targets, based on technical and economic criteria;
2. continued progress in the rapidly developing science and technology that support industrialization of biology; and
3. engagement with significant societal factors impacted by the acceleration of this industry.

Copyright © National Academy of Sciences. All rights reserved.

BOX 1-1
Statement of Task

In order to realize the full benefit of research investments intended to enable the advanced manufacturing of chemicals using biological systems, an ad hoc committee will develop a roadmap of necessary advances in basic science and engineering capabilities, including knowledge, tools, and skills. Working at the interface of synthetic chemistry, metabolic engineering, molecular biology, and synthetic biology, the committee will identify key technical goals for this next-generation chemical manufacturing, then identify the gaps in knowledge, tools, techniques, and systems required to meet those goals, and targets and timelines for achieving them. It will also consider the skills necessary to accomplish the roadmap goals, and what training opportunities are required to produce the cadre of skilled scientists and engineers needed. While focused on industrial manufacturing of chemicals, the roadmap challenges identified here will also be relevant to applications in health, energy, environment, and agriculture by advancing the tools and techniques required for new development in these areas.

Essential elements of the roadmap that the committee will consider in the study and in its report, include the following:

- identification of the core scientific and technical challenges that must be overcome;
- identification of and timeline for tools, measurement techniques, databases, and computational techniques needed to serve as the building blocks for research and applications;
- how to develop, share, and diffuse common interoperable standards, languages, and measurements; and
- when and how to integrate nontechnological insights and societal concerns into the pursuit of the technical challenges.

The report will provide guidance to both the research and research funding communities regarding key challenges, knowledge, tools, and systems needed to advance the science and engineering required for advanced manufacturing of chemicals using biological systems and to develop the workforce required to realize these advances. The report will not include recommendations related to funding, government organization, or policy issues.

Definitions

A number of key terms are used throughout this report. Because many of these terms do not necessarily have universally agreed-upon definitions, we define the following for the purposes of this report:

The **bioeconomy** refers to the portion of the economy that is derived from biological processes and manufacturing. With reference to Figure 1-1, **feedstock** refers to the starting material used in the manufacturing process. This may be a form of biomass, a crude or refined petroleum hydrocarbon product, or a material that has already been chemically modified

Copyright © National Academy of Sciences. All rights reserved.

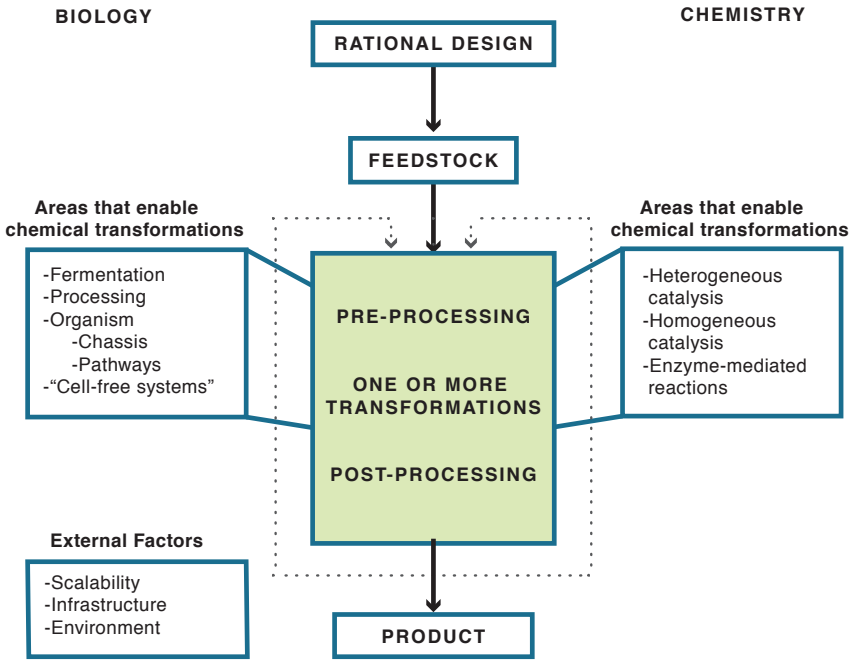


FIGURE 1-1 Chemical manufacturing flowchart showing the report's conceptual schema of the chemical manufacturing process, from rational design to product.

in some way. Likewise, **product** refers to the material after it has been subject to a change in its chemical structure. Finally, **transformation** refers to a change in chemical structure. This could be via *traditional* chemical synthesis, biological routes, or both.

Biotechnology involves the "use of living cells, bacteria, etc., to make useful products."[15] **Genetic engineering** encompasses the cutting and joining of recombinant DNA and its incorporation into an organism in order to change its characteristics,[16] for example to make a new product or enhance its production. Genetic engineering is made up of a variety of technologies. **Protein engineering** seeks to modify the properties of an individual protein, for example to improve its stability or catalyze a new reaction. **Metabolic engineering** encompasses the purposeful modification of metabolic, gene regulatory, and signaling networks to achieve enhanced production of desired chemicals.

Synthetic biology is a newer discipline that seeks to deliver greater speed, cost-effectiveness, and predictability to the design of biological systems. The field applies engineering principles to reduce genetics into DNA "parts" and understand how they can be combined to build desired

Copyright © National Academy of Sciences. All rights reserved.

functions in living cells. This has been driven in advances to build long stretches of DNA and "edit" the genomes of natural organisms. The UK Synthetic Biology Roadmap Coordination Group defined synthetic biology as "the design and engineering of biologically-based parts, novel devices and systems as well as the redesign of existing, natural biological systems."[17] Synthetic biology is a toolbox, not an end in itself. Advances in synthetic biology accelerate the industrialization of biology.

CHEMICAL MANUFACTURING

Human health, energy, the environment, and agriculture are important domains for the application of biotechnology. Tremendous progress has been made in these areas, and these developments will be greatly accelerated by advances toward the scientific and technological milestones discussed in this report. While this committee highlights the relevance of the industrialization of biology to health, energy, the environment, and food, the focus of this report is the production of chemicals through industrialization of biology.

The arrival of this emerging capability in biological science comes at a time when new approaches are eagerly sought. As the processing of petrochemicals has become an increasingly mature industry, new global challenges have begun to emerge. To sustain the needs of an increasing global population, the provision of goods and services must in turn become more sustainable, making more efficient use of fossil feedstocks and enabling the greater use of renewable feedstocks.

As a way to frame the discussions throughout this report, Figure 1-1 provides a conceptual framework for the chemical manufacturing process, including both biological and traditional chemical routes to chemical transformations. At its most basic level, the chemical manufacturing process has four basic waypoints. After conception of the product, or properties of a product, to be made, the rational process design is considered. This includes considering the capabilities available in science and engineering, as well as beginning to consider possible chemical transformations that will lead to the product of choice. As part of the design process, a feedstock is selected. In the case of traditional chemical manufacturing, this feedstock may be crude oil. In the case of using a biological transformation, this may be a plant-derived material (e.g., switch grass, corn stover) or a crude hydrocarbon mixture. The green box represents the core components of the chemical transformation or transformations that will occur. In this case, feedstocks typically undergo some initial processing before being subject to the one or more chemical transformations that are required to generate a product. It is important to note here that either chemical or biological means may be used to enable this transformation.

Copyright © National Academy of Sciences. All rights reserved.

Finally, some sort of post processing (e.g., separation of the product from a fermentation liquid) will yield a final product for sale, or an intermediate product that may undergo further transformations. This report addresses technical and societal challenges relevant to each aspect of this figure, including rational design, selection and development of feedstocks, pre-processing and process design, and various methods of chemical transformation. In addition, the report discusses many of the external factors that affect the entire production process, including scalability, infrastructure, the environment, and even legal and business frameworks.

This report concerns the use of biology in the production of chemicals for industrial and consumer use. As described in Figure 1-2, these materials include large-volume chemicals—the majority of which are produced through chemical routes today—and specialty chemicals, which may be uniquely suited to production via industrial biology. Large-volume (bulk) chemicals include final products such as fuels, and bulk chemical intermediates such as ethylene and butadiene, among others. Some specialty chemicals will be natural products that can be adapted to industrial biology, while other materials may be modifications of natural products, such as enzymes and polypeptides used for industrial purposes, as catalysts or additives.

Although many of the materials described in this report derive from renewable feedstocks—starch or cellulose-derived sugars—this report concerns chemicals and materials produced by the use of biological sciences, regardless of the specific starting materials used. Production processes that include both chemical transformations and biological processing are germane to this report.

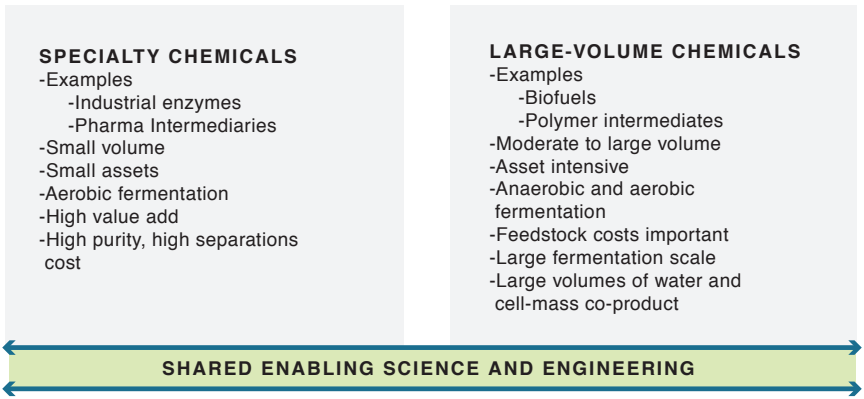


FIGURE 1-2 Comparison of the technical, economic, and production differences between small-volume specialty chemical production and large-volume commodity chemical production.

Copyright © National Academy of Sciences. All rights reserved.

TOOLS AND TECHNOLOGIES

The core technologies enabling the industrialization of biology are those that enable microbial biotechnology. The production of chemicals through biological processes may entail fermentation using living host organisms, "cell-free" bioprocessing, or simply enzyme-mediated syntheses. Synthetic biology is therefore at the heart of the ongoing industrialization of biology. Synthetic biology takes advantage of the science of recombinant DNA and the ability to read, write, and edit the DNA of microorganisms, allowing the design and construction of new, more efficient metabolic pathways.

These technologies generally, but not exclusively, affect our ability to perform chemical transformations through biological processes.

Several areas of science support the speed, efficiency, and cost of development of these technologies. Some important areas among these are the following:

1. Advances in DNA sequencing and DNA synthesis have dramatically reduced the costs associated with synthetic biology. Proteomics and metabolomics continue to provide insights into the biochemistry of the cell. High-throughput techniques have accelerated the pace of metabolic engineering and reduced the time and expense associated with constructing metabolic pathways in host organisms.
2. The tools of bioinformatics and cell profiling enable an ever more detailed understanding of gene expression and cell metabolism including the ability to collect, manipulate, interrogate, and share the large data sets associated with synthetic biology.
3. Early efforts have been taken to increase the number and range of DNA "parts" that are available to engineer new functions into cells. This includes large enzyme lists that are being gleaned from the sequence databases, synthesized, and characterized for function. Also, efforts have been taken in the characterization of regulatory parts to better control expression with greater precision.[18]
4. Modeling and visualization tools are critical to protein engineering. Predictive modeling is important at the level of proteins, the metabolic pathway, and whole-cell metabolism. Modeling is equally important at the macro level, from predicting how whole cells function as a population and interact with their environment to the design and operation of bioprocessing plants.
5. The commercial-scale production of chemicals via bioprocessing requires the design and operation of large-scale facilities capable of economic production and purification of the chemical products. The science and technology needed for design and "scale-up" of new bioprocesses is important to the industrialization of biology.

Copyright © National Academy of Sciences. All rights reserved.

SOCIETAL FACTORS

Acceleration of the industrialization of biology will require the convergence of several societal factors, including a properly trained workforce; appropriate legal frameworks; and physical infrastructure and standard operating procedures for safely containing, working with, and disposing of organisms used in bioprocessing. Public acceptance and endorsement of bioprocessed chemicals will be an important consideration in their commercial viability. In addition, international harmonization of policies would make the economic and governance environment more conducive to advances in industrial biology. The challenge for policy makers is to find the right mix of governance tools to promote innovation while also respecting a diversity of values and supporting effective oversight.

The workforce required for industrial biology to reach its potential will need a multidisciplinary education, with expertise in the biological sciences, chemistry, engineering, and computing. Expertise in environmental science will also be crucial for some industrial applications. Robust development of industrial biology will require a workforce with the expertise to create and safely operate complex organisms.

Industrial biotechnology will need a governance framework that balances important social goals and manifests important values (Figure 1-3). Governance involves deployment of a variety of policy tools by which an industry's behavior can be shaped, including education of industry actors, industry self-governance through standard setting, accreditation, government standard setting and regulation, public engagement and public scrutiny, tort liability, and other mechanisms for developing safety standards and controls. For governing the industrialization of biology, key goals pertain to safety (risk identification and mitigation) and sustainability. For industrial biotechnology to deliver widespread benefits it must have low environmental impacts, use biological feedstocks sustainably, and operate according to high safety standards with respect to humans, animals, and the environment.

Ultimately, the right to operate bioprocesses is conferred by society. Regulatory frameworks to promote safety are necessary but not sufficient. The public must have sufficient understanding and acceptance of the science and technology involved to ensure comfort with and acceptance of goods produced through new techniques.

ORGANIZATION OF THE REPORT

The remainder of this report addresses the elements of the statement of task outlined above.

Chapter 2 examines the unique opportunity for the acceleration of the industrialization of biology presented by the convergence of biology,

Copyright © National Academy of Sciences. All rights reserved.

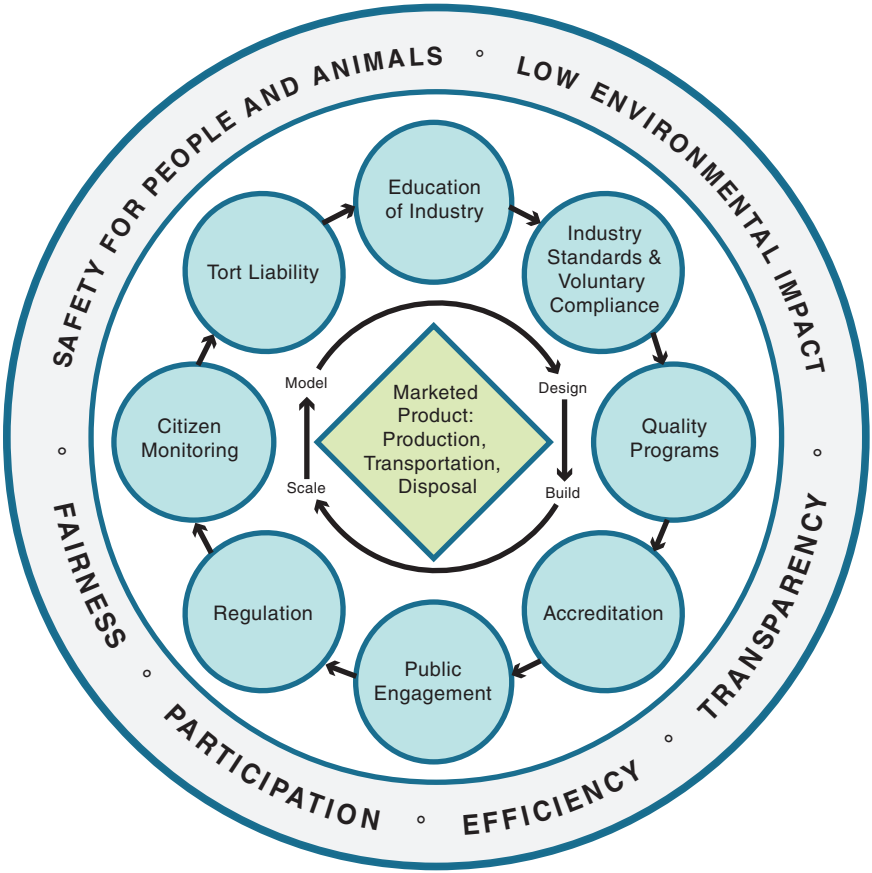


FIGURE 1-3 Tools of governance. This figure indicates the relationships between a specific production process (represented by the central quadrilateral figure), the various tools of governance (represented by the blue circles), and the concepts that a governance structure for chemical manufacturing should value (the outer circle).

chemistry, chemical engineering, and other critical fields; the development of new tools and methods; and the current economic success of chemicals produced through biological routes. The chapter discusses the core drivers of the industrialization of biology and identifies the societal challenges that the industrialization of biology is poised to address.

Chapter 3 develops a vision of a future in which the industrialization of biology is ubiquitous. It explores what materials might plausibly be produced through biological routes, what kind of economic infrastructure might be developed, and how these changes might affect society. This

Copyright © National Academy of Sciences. All rights reserved.

chapter also identifies those societal questions that should be addressed as chemical manufacturing through biological routes matures and discusses potential governance mechanisms that might be used to address those questions.

Chapter 4 presents the committee's technical roadmap for the industrialization of biology, including specific roadmap goals and timelines for feedstock utilization and development, chemical transformations, and deeper understanding of organisms. This chapter discusses critical aspects of each of these major technical areas and provides specific recommendations for the rapid achievement of the roadmap goals and the necessity of viewing the roadmap and the process that generated it as an evergreen process.

Chapter 5 distills the committee's analysis and assessment of both technical and nontechnical issues into a set of specific recommendations to stakeholders involved in the industrialization of biology.

Copyright © National Academy of Sciences. All rights reserved.

Copyright © National Academy of Sciences. All rights reserved.

2

Industrial Biotechnology: Past and Present

THE BIOECONOMY AND GLOBAL CHALLENGES

The industrialization of biology offers far-reaching benefits at both the global and the national scale

1. by driving the innovation economy and sustainable economic growth;
2. by potentially contributing to the solutions to some of the societal grand challenges of our time, such as helping to deliver clean, affordable, and sustainable energy;
3. by enabling sustainable, next-generation manufacturing; and
4. by creating new skills and jobs to benefit today's and tomorrow's generations.

Accelerating advanced chemical manufacturing by industrializing biology can drive the rapid growth of an innovative U.S. bioeconomy. A substantial share of economic output will be increasingly related to the development and use of biological materials and bio-based processes for both chemical production and the development of new materials. The industrialization of biology creates social, environmental, and financial advantages that combine economic growth with public benefits and better lives for our citizens.

Copyright © National Academy of Sciences. All rights reserved.

BIO-BASED MARKETS ALREADY ARE SIGNIFICANT AND THRIVING

As demonstrated throughout this report, the future economic and societal benefits from the industrialization of biology are compelling. Bio-based markets already are significant in the United States, representing more than 2.2 percent of the gross domestic product (GDP) in 2012, or more than $353 billion in economic activity in 2012.[2a] The European Commission estimates that the European bioeconomy (excluding health applications) already is worth more than €2 trillion annually and employs more than 21.5 million people.[19]

Carlson has constructed a genetically modified domestic product (GMDP) metric to compare bio-based markets and biotechnology with the economy as a whole. His current data reveal that "the U.S. economy, and in particular annual U.S. GDP growth, is becoming increasingly dependent on biotechnology."[2a] The Carlson GMDP-to-GDP comparison shows that bio-based markets have grown rapidly as a percentage of American GDP and that, by 2012, they constituted 5.4 percent of annual GDP growth.[2a]

Bio-based chemicals also are not entirely new, nor are they a historic artifact. Current global bio-based chemical and polymer production already is estimated to be about 50 million tons each year.[20] Bioprocessing techniques such as fermentation, baking, and tanning have been used throughout much of human history. In recent history, we have witnessed major advances made possible by techniques such as genetic engineering and the development of the biotechnology industry.

According to several Organisation for Economic Co-operation and Development (OECD) analyses, "[i]ndustrial biotechnology has rapidly matured, and has produced some tangible products, including a large number of bio-based chemicals and bioplastics."[21] The OECD predicted in 2009 that bio-based products would constitute at least 2.7 percent of GDP among the OECD member countries by 2030.[22] The rapid advances in scientific research and technological developments in only the past 5 years have led the OECD to revise that projection significantly. In Denmark, it is estimated that about 40 percent of manufacturing already takes places in a "cell factory."[23]

Other recent studies confirm the rapid growth in the thriving markets for bio-based products. Data from Agilent Technologies show that U.S. business-to-business revenues from industrial biotechnology alone reached at least $125 billion in 2012.[2b] Bio-based chemical applications accounted for about $66 billion of that U.S. economic activity, while biofuels added another $30 billion. Lux Research estimates that industrial chemicals made through synthetic biology currently represent a $1.5 billion market and that this likely will expand at 15 to 25 percent annual growth rates for the foreseeable future.[3] A recent U.S. Department of

Copyright © National Academy of Sciences. All rights reserved.

Agriculture (USDA) report indicates that, by 2015, bio-based chemicals will make up greater than 10 percent of the chemical market.

The markets for bio-based chemicals and industrial biotechnology for chemical manufacturing processes are growing roughly twice as fast as those in biomedicine or agriculture (Figure 2-1). They also reflect new structural market shifts as decentralized production processes, innovative new value chains, and collaborative ventures both compete with and complement the vertically integrated chemical manufacturing facilities that have marked the past century.

New chemicals and biochemical materials are being developed through the rapid emergence of novel bio-based technologies and processes. Biology is being used to develop innovative and resource-conserving solutions to difficult problems.

Rapid growth already is occurring in three interrelated segments of bio-based advanced chemical manufacturing: enabling technologies (tools and platforms needed for the development of advanced chemicals); core technologies (processes and inputs used to make chemical products); and enabled products (chemical products on the market). According to BCC Research, all three segments in synthetic biology are growing at more than 70 percent per year, with significant markets for enabling technologies ($653 million), core technologies ($699 million), and enabled chemical products ($2.8 billion) by 2016.[24]

THE INNOVATIVE POWER OF INDUSTRIAL BIOTECHNOLOGY

The biggest bioeconomy market opportunities from accelerating advanced manufacturing of chemicals through the industrialization of biology have yet to be realized. As the Obama Administration posited in its 2012 National Bioeconomy Blueprint, the rapid development of new bio-based chemicals and materials in the U.S. bioeconomy can "allow Americans to live longer, healthier lives, reduce our dependence on oil, address key environmental challenges, transform manufacturing processes, and increase the productivity and scope of the agricultural sector while growing new jobs and industries."[1]

A 2013 Milken Institute report underscored the huge potential opportunities when it noted that "[n]inety-six percent of all U.S. manufactured goods use some sort of chemical product, and businesses depend[e]nt on the chemical industry account for nearly $3.6 trillion in U.S. GDP."[25] The global market for enzymes used in consumer products and industrial production processes—and a prime target for the industrialization of biology—alone is expected to reach $8 billion by 2015.

The OECD has projected that industrial biotechnology and bio-based chemical manufacturing likely will accelerate and lead the development

Copyright © National Academy of Sciences. All rights reserved.

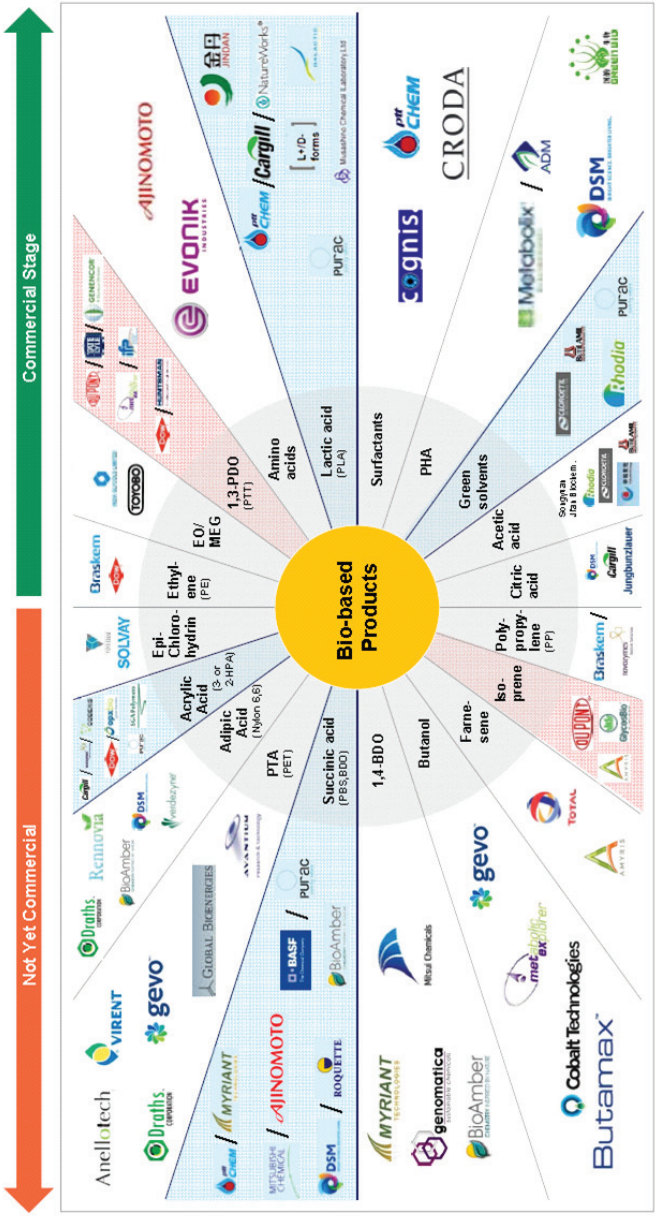


FIGURE 2-1 Biomaterials competitive landscape. For biomaterials, a robust and highly competitive industrial ecosystem already is beginning to emerge. This figure provides key examples of successful chemical manufacturing through biological routes.

Copyright © National Academy of Sciences. All rights reserved.

of a robust, global bioeconomy. The potential economic and societal benefits predicted by the OECD and others become clear when we realize that while there are more than $4 trillion of products made by chemical transformations globally, only about 5 percent of these potentially "addressable markets" have been addressed biologically. A study from BCC Research suggests that synthetic biology markets for chemicals will grow to $11 billion by 2016,[24] and a broader review from McKinsey Global Institute estimates that synthetic biology and the industrialization of biology will provide a disruptive set of technologies with an economic impact of at least $100 billion by 2025.[26]

As a result, the broad applications of advanced chemical manufacturing for multiple uses in energy, health, advanced consumer products, agriculture and food, cosmetics, and environmental technologies are expected to produce trillions of dollars in addressable global market opportunities. Several recent studies estimate that at least 20 percent of today's petrochemical production can be replaced by the industrialization of biology in chemical manufacturing over the next decade.[25]

Aside from the large size of the chemical markets that can be addressed biologically, making biology easier to engineer and developing new chemical manufacturing capabilities based on synthetic biology will have a broad range of other economic benefits. They include opening up the potential for innovative products and processes, decentralizing production process and value chains, creating incentives for new entrants (including high-growth small- and medium-sized enterprises), creating new jobs and skills, and incentivizing new business models.

In addition to substitution or displacement effects, the industrialization of biology will lead to the production of new molecules for chemical, fuel, and material applications, which are not currently possible from fossil fuel sources or traditional manufacturing. The potential for new innovation and market creation remains considerable. Major new advances in measurement tools, computer-aided design, and design-build-test-learn cycles—from the most fundamental level to engineer living material to that needed for making complex chemicals at the commercially competitive industrial scale—not only provide new tools but also create new bioeconomy markets and investment opportunities in themselves (Box 2-1).

As a result, a robust and disruptive new industrial ecosystem is emerging. Even at this early stage, the number of American synthetic biology commercial companies has grown from 54 to 131 between 2009 and 2013.[27] A large number of new startups have been created, and a number of them have successfully gone public with initial public offerings. But these figures understate the full economic impact of synthetic biology because of the rapid uptake and interest by larger firms in a broad range of sectors.

Copyright © National Academy of Sciences. All rights reserved.

BOX 2-1
Living Foundries

For example, the Defense Advanced Research Projects Agency's Living Foundries program is working with many companies, national laboratories, and universities to develop new tools to enable rapid engineering of biology. It is tackling "impossible today" industrial projects that could become "possible" if we enable, scale, and rapidly prototype genetic designs and operating systems never before accessible for industrial production. And its most recent large-scale initiative, the 1,000 Molecules Project, seeks nothing short of a fundamental disruption of traditional chemicals and materials industries and processes by developing 1,000 new chemical building blocks for entirely new materials at the molecular scale and nanoscale in the next 3-5 years.

Established chemical companies are making significant investments internally, in partnership with startups, or both. New business models are proliferating, as are innovative collaborations, driven by advances in synthetic biology. University-industry linkages span the continuum from high-risk basic research to late-stage prototype development projects that can scale and compete at market-driven price and performance points.

Enabling the Industrialization of Biology and the Development of an Ambitious Roadmap for Accelerating the Advanced Manufacturing of Chemicals

Six core drivers—the "6 C's"—are intersecting to drive the industrialization of biology and to accelerate the development of advanced manufacturing of chemicals: (1) convergence of biology and engineering; (2) challenges for society in addressing global grand challenges related to energy, climate change, the environment, agriculture and food, and health; (3) competitiveness; (4) human capital and capacity; (5) the confluence of new enabling tools, platforms, data, and processes competitiveness; and (6) the current state and readiness of both the science and industry.

Convergence

The industrialization of biology in chemical manufacturing is enabled by the convergence of biology with chemistry and engineering in transformative new ways. A major part of the U.S. strategy for advanced manufacturing of chemicals is the expectation that the next industrial revolution will involve making things with greater precision, at ever

Copyright © National Academy of Sciences. All rights reserved.

higher speeds, and at lower costs, and more sustainably by focusing on the biological processes.

Convergence includes not only transdisciplinary research and development and the integration of science but also the intersection of previously distinct industrial sectors such as chemical synthesis, industrial biotechnology and bioenergy, information technologies, and enabling tools and platforms from a number of business sectors. The growing convergence of transdisciplinary science, technology, engineering, and mathematics, along with overlapping markets and innovative business models, enables novel solutions to many previously intractable societal challenges.

Four influential reports in recent years highlight this trend and underscore the importance of convergence as a core driver. *A New Biology for the Twenty-first Century* was issued in 2009 and received extensive attention.[13] It concluded that biology would be the key, new driver for innovation in the 21st century, much as the physical sciences had led to the information and communications technology revolution and other major breakthroughs in the 20th century. It also showed that biology increasingly would intersect with previously disparate disciplines, including chemistry. This new integration would provide the basis for not only new economic growth but also the tools and platforms for addressing many of the major global grand challenges of this century.

In 2011, a Massachusetts Institute of Technology (MIT) faculty foresight review concluded that we were entering a Third Revolution in the life sciences.[28] The first had been the DNA, genetics, and molecular biology revolution that provided the basis for today's modern biotechnology industries and approaches. The second was the genomics revolution made possible by the Human Genome Project. The Third Revolution is based on convergence and will transform next-generation manufacturing and production by merging biology and engineering in completely new ways.

The third major report was the ARISE II report from the American Academy of Arts and Sciences (AAAS), titled *Advancing Research in Science and Engineering: Unleashing America's Research & Innovation Enterprise*. It noted that many of the historical distinctions between the modes of thought and organizational principles in the life sciences and the physical sciences were converging around common challenges and opportunities, but that outdated structures and approaches impeded collaborations, communication, and the translation of research into new products and services While the physical sciences have linked basic and applied research as "an interwoven continuum," the life sciences have tended to make sharp distinctions between basic and applied sciences in its disciplines.[29]

The ARISE II report made two principal recommendations. First, America's research and innovation enterprise must embrace a new

Copyright © National Academy of Sciences. All rights reserved.

transdisciplinary organizing principle. The AAAS called for providing incentives to ensure that "tools and expertise developed within discrete disciplines are shared and combined to enable a deep conceptual and functional integration across the disciplines."[29] It also recommended creating an interdependent ecosystem for linking academic, government, and private sectors throughout the discovery and development process.

In May 2014, a follow-up National Research Council report, *Convergence: Facilitating Transdisciplinary Integration of the Life Sciences, Physical Sciences, Engineering, and Beyond*, sought to capture two dimensions: the convergence of the subsets of expertise necessary to address a set of research problems and the formation of the web of partnerships involved in supporting such scientific investigations. These two dimensions further enable the resulting advances to be translated into new forms of innovation and new bio-based products and services.

In this frame, convergence represents a major cultural and organizational shift for academic organizations and government science or technology ministries that have been traditionally organized around discipline-based departments. The overall ecosystem needed for convergence draws on not only academic contributors but also a much broader cross-fertilization of ideas from national laboratories, industry, citizen scientists, and funding bodies, as well as from new insights provided by economics and the social sciences. The process of convergence is applicable to basic science discovery as well as translational applications in industry. Because it is commonly focused on achieving an outcome to a challenge at the frontiers of knowledge and new markets, many convergence efforts include a major entrepreneurship component that leads to the development of new webs and ecosystems of startup companies and economic innovation.

An emerging metaphor from convergence is that of the "cell as tomorrow's factory." As Neri Oxman from the MIT Media Lab observed, "The biological world is displacing the machine as a general model of design."[30] In short, the industrialization of biology and synthetic biology will be as important for the next 50 years as semiconductors and related information and communication devices have been to economic growth over the past 50 years.

SOCIETAL BENEFITS IN ADDRESSING GLOBAL GRAND CHALLENGES

When compared to traditional manufacturing, advanced manufacturing of chemicals through biology might produce social benefits while requiring fewer trade-offs between growth and sustainability. In addition to the economic benefits generated through innovation, productivity

Copyright © National Academy of Sciences. All rights reserved.

increases, and new sources of sustainable economic growth, the advanced manufacturing of chemicals through biology can address 21st-century grand challenges related to energy, climate change, sustainable and more productive agriculture, environmental sustainability, and inclusive growth.

Energy

The industrialization of biology can serve to enhance the United States' energy independence. Advanced chemical manufacturing based on biological sources such as plants, algae, bacteria, yeast, filamentous fungi, and other organisms can replace many chemicals now derived from petroleum or other fossil fuels. If properly designed, bio-based production processes, including new bio-based inputs, can improve energy efficiency and, in some cases, reduce energy costs.

Over time, a growing part of the demand for chemical products and processes likely will come from the increasing economic activity of emerging and developing countries. Given the burgeoning demand for oil and other scarce natural resources in many emerging markets, sustainable sources of new and advanced bio-based chemicals may be the only viable way to meet the needs of their populations.

A driver for the transition to the bioeconomy and novel advanced chemical manufacturing is the anticipation by some energy experts, such as the International Energy Agency, that oil, gas, and coal "will reach peak production in the not too distant future and that prices will climb." The OECD recently demonstrated that the scope and platforms for the bio-based production of chemicals and fuels increased significantly in 2013. Its analysis concludes that these developments "may open the door to greater replacement of the oil barrel."[31]

Climate Change and Environmental Sustainability

The advanced manufacturing of bio-based chemicals could provide numerous environmental benefits. Many producers of consumer products are now committed to "green growth," which will require new enzymes and other chemical inputs, together with more sustainable production processes. Bio-based production, properly designed and managed, has the potential for generating fewer toxic by-products and less waste than traditional chemical manufacturing.

New approaches to the advanced manufacturing of chemicals are well aligned with American efforts to mitigate the adverse effects of greenhouse gases and to enable the United States to meet its global climate change commitments. By using biomass as a feedstock and through advanced manufacturing techniques that industrialize biology, the manu-

Copyright © National Academy of Sciences. All rights reserved.

facturing of advanced chemicals can achieve significant savings in greenhouse gas emissions compared to production from oil or other fossil fuels.

The development of advanced manufacturing for bio-based chemicals also means increasing the number of products that are carbon neutral in terms of not producing any net increase in carbon dioxide or other greenhouse gases over their entire life cycle: from design and production through disposal. At the same time, significant waste reduction may be achieved through bio-based production processes and the resulting products' life cycle, including manufacturing the advanced chemicals used to produce them. The milder bioprocess conditions—such as generally lower temperature and pressure—used for bio-based manufacturing compared to fossil fuels also contribute to sustainability. Additional environmental benefits will be related to using synthetic biology and related techniques for bioremediation that can bring contaminated soil back into productive use. The OECD notes that the world's soil is being lost 18-30 times faster than it is formed, and that new methods such as synthetic biology are important for limiting soil destruction and for growing crops more efficiently.

Agriculture

A roadmap focused on manufacturing at commercially competitive scale will create new opportunities for American agriculture and provide new value chains that do not require costly trade-offs with land. As a Milken Institute study concluded, "[b]io-based chemicals offer the prospect of new cash crops like switchgrass, new demand for the cellulosic fiber in traditional crops, and new jobs in bio-chemical production and process."[25]

The increased use of biomass as a feedstock for the production both of high-value, low-volume, bio-based chemicals and bioplastics and of low-value, high-volume, bulk biofuels and commodity chemicals provides new opportunities for innovation in sustainable agriculture. Integrated production facilities that offer the ability to produce not only biofuels but also bio-based chemicals and bioplastics are becoming increasingly technologically feasible and economically viable. Advanced feedstocks will also permit farmers to produce larger yields on smaller amounts of land to feed a growing population.

Competitiveness and Innovation

The promise and importance of the industrialization of biology has not gone unnoticed around the world. China is investing huge amounts in synthetic biology, and it has made this set of technologies a priority in its current 15-year Science and Technology Plan.[32] The United Kingdom

Copyright © National Academy of Sciences. All rights reserved.

developed a list of "The Eight Great Technologies" for the future of Britain, and synthetic biology is considered number two of the top technologies for the United Kingdom's future. A number of countries now are developing national strategies or plans related to synthetic biology, the industrialization of biology, and the future bioeconomy, including many emerging markets such as South Africa, Brazil, and Mexico.

THE TIME IS RIGHT: CURRENT STATE AND ADVANCES IN SCIENCE AND INDUSTRY

Opportunities Arising from DNA Technologies, Systems Biology, Metagenomics, and Synthetic Biology

Biology has the potential to build intricate material and chemical structures with atomic precision. Biotechnology has only begun to harness this capability, and leading-edge products in development have simple structures, such as butanediol, isobutanol, farnesene, and lactic acid. Biology excels at producing more complex molecules and mixtures of molecules. Access to this chemical complexity via biotechnology has been limited by the investment needed to engineer multistep biological transformations.

The past decade has seen an explosion in the technologies to read, write, compose, and debug DNA. This is rapidly increasing the scale and sophistication of genetic engineering projects. In the near term, this will lead to more complex chemical structures and composite nanomaterials, which require precise control over dozens of genes. Examples of this include mining drugs from the human microbiome, obtaining pesticides from environmental samples, and producing metal nanoparticles for electronics and medical devices. In the longer term, one can imagine organisms designed from the ground up for consolidated bioprocessing and automatic assembly of a product that requires multiple steps.

The genetics underlying the natural world are rapidly being illuminated by DNA sequencing, the cost of which is declining faster than Moore's law. The first human genome (3.2 billion base pairs [bp]) was sequenced in 2001 at a cost of $2.7 billion. Nine years later 1,000 human genomes (3.2 trillion bp) were sequenced, and in 2014, Illumina released the HiSeq X, promising a $1,000 genome. Beyond human genetics, this technology has been applied to sequence communities of organisms (metagenomics) populating niches in the environment or associated with hosts, such as the human gut. Databases of sequences have been rapidly grown with information; as of 2013, there were 160 million sequences from 240,000 organisms. This has built an enormous potential catalogue of natural parts, from which high-value chemical pathways can be discovered or created.

Copyright © National Academy of Sciences. All rights reserved.

Accessing these chemicals requires more than the sequence information. Historically, collaboration required the physical transfer of DNA materials, such as genes, between labs. The rise of DNA synthesis has moved biology toward an information science where the DNA can be reconstructed from the sequence information alone, thus eliminating the need for physical transfer and enabling the direct access to biological functions encoded in the sequence databases.[33] DNA synthesis has been applied to build entire 1 MB bacterial genomes and yeast chromosomes.[34] Synthesis provides the genetic designer with full operational control over the identity of each bp of a large design, as opposed to previously, where DNA was stitched together from existing pieces. There is still significant room for improvement; while a large DNA sequencing center can sequence >4 trillion bp/day, the top industrial synthesis companies only produce ~300,000 bp/day.

The ability to compose DNA has lagged behind our ability to read and write it. The most valuable functions require many genes and complex regulatory control over how much, when, and where they are turned on. Synthetic biology offers some tools to tackle this challenge, including genetic circuits, precision regulatory parts, and computer-aided design to systematically recode multigene systems.[35] While it is possible to synthesize entire genomes, we are far from being able to write them from scratch from the bottom up. The current state of the art is the top-down "editing" of existing genomes using technologies such as MAGE[9] and CRISPR/Cas9[10] to introduce incremental changes in an otherwise natural genome.[9, 36] Similarly, genome-scale design tools have emerged to control flux through metabolic pathways (e.g., COBRA and Optknock), but the output of these are predictions of the impact of the top-down knockout of enzymes in a defined host.[37]

Genome-scale engineering, where designs are composed of thousands of genes assembled from the bottom up, will become the norm. This will require computational tools that merge the simulation tools from systems biology with biophysical methods that can convert a desired feature into a specific DNA sequence. New design paradigms are also needed in order to manage a large project and integrate across different cellular systems.[38] The ability to build synthetic regulation (sensors and circuits) needs to be combined with metabolic engineering and the ability to control cellular functions (e.g., protein secretion) and stress responses.[39] As synthetic systems become larger, it will be more important to be quantitative in understanding the distribution of resources and their load on cellular growth and maintenance. Increasingly, it will be important to develop methods to insulate a synthetic system from the background processes of the host.[40]

Genome-scale design will require genome-scale debugging. It is currently impossible to get a snapshot of how a change in the genetic design

Copyright © National Academy of Sciences. All rights reserved.

impacts all of the processes in a cell. Advances in -omics technologies make it possible to characterize the mRNA, proteins, and small molecules in a cell. However, each of these requires specialized expertise and instrumentation and is cost prohibitive to perform on failed designs. The form of the data is nonstandard, and integrating information across transcriptomics, proteomics, and metabolomics is difficult. Furthermore, it is difficult to convert the results into actionable design changes to optimize a system.

Integrated national-level infrastructure can help accelerate the transition to genome-scale designs. Strain databanks and sequencing centers provide surveillance of the natural world and could populate parts libraries with billions of natural enzymes and pathways. BioFABs provide the substrate for large designs by providing high-quality genetic parts through large-scale engineering and characterization.[41] Foundries work to systematically produce products by pushing the scale of genetic designs and integrating DNA manufacturing with cellular analytics.[38] Finally, metrology institutes (e.g., run by the National Institute of Standards and Technology) provide the standards for characterizing genetic parts, reporting construction precision, and software integrating -omics data.[42]

New High-Value Chemical Products Unobtainable by Traditional Chemical Synthesis

Organic synthesis is a mature discipline where nearly any target molecule can be made through a logical combination of reaction steps. A similar capability has not been realized in biotechnology, the products of which have been somewhat limited to chemicals naturally made by cells. Only small numbers of enzymes have been characterized relative to the potential of chemical space, and there is a lack of standards in reporting enzyme activity and specificity. However, we are at an inflection point in enzymology, where sequence databases have been populated with tens of millions of enzymes and access to this resource *en masse* is enabled by DNA synthesis. This will lead to a revolution in pathway design, where obtaining a non-natural target molecule by combining enzymes will be intellectually analogous to the logic of organic chemistry.

In the early 2000s, access to even a single enzyme required the physical transfer of DNA materials between labs. The decline in DNA synthesis has made it routine to build all of the genes from a sequence database predicted to have a desired function. Such "part mining" typically involves constructing the genes for hundreds of enzymes and has been very successful at identifying variants with desired properties. This often results in the combination of genes from diverse organisms in building a pathway.

This mining will soon grow from individual enzymes to entire pathways.[33a] Potentially high-value products are encoded by 10- to 150-kb

Copyright © National Academy of Sciences. All rights reserved.

gene clusters that encode multienzyme cellular factories. Many natural products from these pathways are produced industrially by the native organisms; for example, rapamycin is a pharmaceutical that may have an effect in treating cancer and spinosyn is a biological pesticide. While synthetic routes to complex natural products can be found, there are few labs capable of this and the routes are often low yield and not viable industrially. It is clear that only the surface has been scratched in identifying potential products. Just recently, bioinformatics algorithms have been developed that enable the enumeration of clusters in genome databases. To date, 40,000 have been found in the National Center for Biotechnology Information, and metagenomic samples, including the human microbiome, yield tens of thousands of new clusters.[43] The Agency for Science, Technology and Research of Singapore's collection consisting of 120,000 strains is estimated to contain several million novel pathways. Fully accessing their chemical products will require a further reduction in synthesis costs and the application of synthetic biology to control multigene systems and functionally transfer activity to new organisms.

Collectively, the mining efforts are yielding a deluge of new enzyme data across entire families that encompass activity and specificity information. In addition to the mapping of natural diversity, advances in engineering enzymes have yielded transformations that are chemically difficult, for example, the use of cytochrome P450s for C-C and C-N chemistry. The expansion of the enzyme toolbox will feed new computational methods whose input is a desired chemical structure and whose output is a combination of enzymes predicted to build the molecule. Examples of this include the Biochemical Network Integrated Computational Explorer (BNICE)[44] and the ACT ontology, which has been used to build a synthetic pathway to N-acetyl-p-aminophenol (Tylenol).[45]

Implementing Computation in Cells

Living cells have an incredible capacity to sense and interpret their surroundings. They achieve this using gene regulation that functions as sensors to respond to environmental stimuli and circuits that process this information and commit cells to a response.[39, 46] To date, this has not been utilized as part of chemical production in biotechnology. However, the potential impact is enormous. Even simple operations would be valuable, such as turning on different pathways at various times during fermentation and implementing feedback regulation to avoid the accumulation of toxic pathway intermediates.[47] Sophisticated circuitry would enable new options in cellular design. For example, entire process control algorithms could be implemented that optimize the uptake of feedstock and control of flux through metabolic pathways. Consolidated bioprocessing could

Copyright © National Academy of Sciences. All rights reserved.

be implemented, where the degradation of biomass and construction of a complex chemical is performed via a preprogrammed order of events. Finally, the composite materials that are possible via biology require precise timing and spatial location of gene expression.[48]

INDUSTRY IS READY

The use of biological organisms to transform precursor molecules into targeted molecular end points dates back to the earliest days of recorded human history in the form of fermentation to produce beer, cheese, and bread. Jokichi Takamine's work on the "koji" process—resulting in an 1894 patent on a microbial enzyme in the United States—marked a turn toward the industrialization of biological processes.[49] The modern era of industrial biotechnology began more recently, with its birth typically tied to large-scale fermentation of penicillin. The antibiotic, first isolated by Sir Alexander Fleming in 1928, was not advanced to large-scale production until World War II, driven by the need for an alternative to sulfa drugs to treat bacterial infections.[50] Although initial titers in 1939 were estimated to be on the order of 0.001 g/L, by the end of the war U.S. capacity for penicillin production was sufficient for ~100,000 patients per year. Penicillin was not the only major fermentation process that attained industrial prominence during this time. The "ABE" process, in which *Clostridium acetobutylicum* ferments sugars to a mixture of acetone, butanol, and ethanol, was first developed in the early 20th century and was a primary source of acetone during World War I. During World War II, the ABE process was the primary source of butanol, and fermentation remained popular as a source for this commodity until lower-cost petroleum-derived products emerged in the 1960s.

It is notable that these achievements preceded the elucidation of the structure of DNA by nearly a decade. The status of penicillin, acetone, and butanol as natural products derived from (reasonably) productive host organisms enabled the use of methods such as classical mutagenesis that required neither complete understanding of the underlying biochemical pathways nor knowledge of the genes encoding the constituent enzymes in order to enhance productivity. In the three decades that followed, fermentation processes were developed for large-scale commercial production of several additional products including citric acid, vitamin B12, glutamic acid, and lysine. Even in the absence of detailed genetics, understanding of pathway regulation facilitated the use of classical mutagenesis to select improved strains. For example, the observation that lysine inhibits upstream enzymes in the pathway for its synthesis led to the use of the lysine analog S-2-aminoethyl cysteine as a selective inhibitor to identify feedback-resistant mutants. Such methods could be routinely used to

Copyright © National Academy of Sciences. All rights reserved.

improve natural product fermentations, but the utility of biology for chemical manufacturing remained limited to a small subset of molecules.

In 1973, Herbert Boyer of the University of California, San Francisco, and Stanley Cohen of Stanford University and colleagues published a manuscript titled *Construction of Biologically Functional Bacterial Plasmids In Vitro*.[51] This publication described what is widely considered to be the first genetic engineering experiment and signaled the start of the Biotechnology Age. However, despite the decades-old history of the industrialization of biology for small-molecule production, "biotechnology" became nearly synonymous with "biopharmaceuticals" in the aftermath of this breakthrough. Less than 10 years later, recombinant human insulin was approved by the Food and Drug Administration and numerous protein drugs followed.

Why did biopharmaceuticals eclipse biochemicals as the first and most obvious beneficiaries of recombinant DNA technology? This is perhaps best answered in consideration of the complexity of the final products and the reactions that encode them.

Therapeutic proteins are structurally more complex than small molecules, yet their synthesis is directly related to the heterologous DNA chosen for expression. Simple overexpression in the right host of as little as a single gene produces the product of interest.

Conversely, biochemicals are the result of a series of enzyme-catalyzed reactions, with each enzyme encoded by at least one gene. This complexity of *pathways* over *products* creates a systems-level challenge that requires more systems-oriented solutions. This challenge and the discipline that emerged to address it was first codified in *Toward a Science of Metabolic Engineering* by James E. Bailey.[52] Metabolic engineering sought to take advantage of the tools of recombinant DNA technology while applying systems and network analyses to the challenge of engineering more productive strains.[53] These principles were successfully applied to generate highly efficient and productive fermentation processes for a number of products, including, for example, 1,3-propanediol and lysine from an engineered strain of *E. coli* with extraordinary maximum production rates of 8 g/L/hr.

In addition, unlike large-volume chemicals produced to replace or supplement an already existing commodity (e.g., ethanol), novel biopharmaceuticals are highly specialized products for which no other production route is possible or economically realistic. As Figure 2-2 indicates, biopharmaceuticals and other specialized products (such as industrial enzymes) operate at a much higher gross margin than products in large-volume commodity markets. The biology and chemistry of potential products helps to explain why biopharmaceuticals were some of the first

Copyright © National Academy of Sciences. All rights reserved.

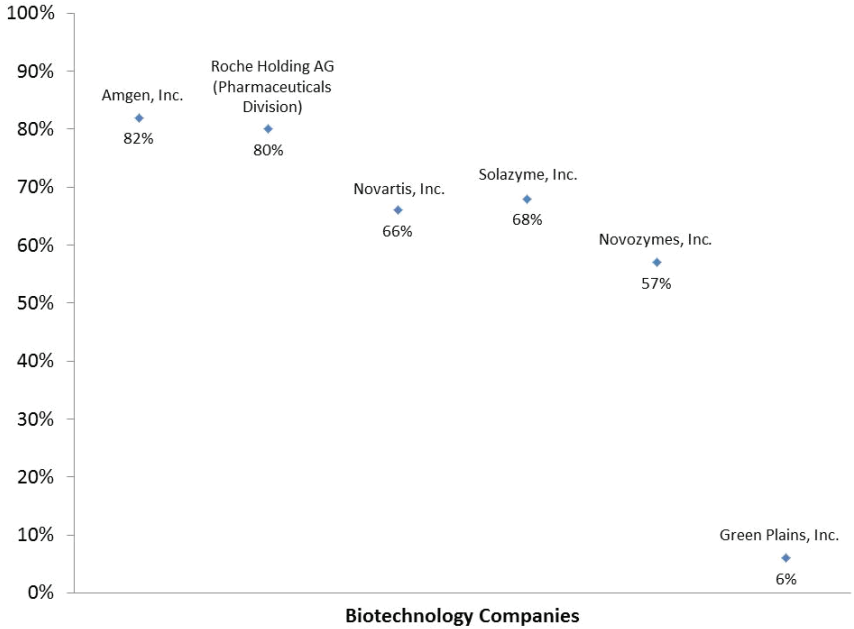


FIGURE 2-2 This figure displays the gross margin of several representative companies producing biopharmaceuticals, industrial enzymes, high-value oils, and ethanol
SOURCES: Amgen, Inc. (2014, February 24) 10-K. Retrieved from http://www.sec.gov; Solazyme, Inc. (2014, March 14) 10-K. Retrieved from http://www.sec.gov; Novartis, Inc. (2014) Novartis Annual Report 2014. Retrieved from http://www.novartis.com; Roche Holding AG (2014) Roche Financial Review 2014. Retrieved from http://www.roche.com [note: gross margin is exclusive to Roche pharmaceuticals division]; Novozymes, Inc. (2014) 2014 Annual Report. Retrieved from http://www.novozymes.com; Green Plains, Inc. (2014, February 10) 10-K. Retrieved from http://www.sec.gov

products of the industrialization of biology, but the economic factors are also critical.

While bio-based chemical production from whole-cell organisms was being advanced, the tools of biotechnology were also being applied in other arenas. Of note is the use of genetic engineering in agriculture. Engineered crops have led to such favorable properties as increased yields, decreased pesticide use, and reduced greenhouse gas emissions in field operations. Additionally, the ability to clone and heterologously express recombinant proteins has greatly facilitated the use of biology to perform single-step reactions *in vitro*. Purified enzymes have been used to produce

Copyright © National Academy of Sciences. All rights reserved.

Industrialization of Biology: A Roadmap to Accelerate the Advanced Manufacturing of Chemicals

a number of products, especially for pharmaceutical synthesis, but the power to extend this lies in the advanced tools of biotechnology. Protein engineering, especially through directed evolution, has generated enzyme variants with activities on previously naïve substrates and enabled the synthesis of numerous small molecules, particularly chiral ones, of pharmaceutical significance.[54] The use of biocatalysis for chemical synthesis can not only access structures with higher atom economy but also significantly reduce the environmental footprint of the associated process. In the case of sitagliptin, a molecule produced through collaboration between Merck and Codexis, the optimized use of a biocatalytic step to replace an analogous chemical one in the process resulted in reduction in the ratio of total mass of materials used to mass of isolated product from 37 to 6.[55]

The key drivers for a company using bio-based methods for chemical production can vary greatly based on the products being manufactured. Under ideal circumstances, an analysis of the margin on a series of products would allow us to understand the economic drivers that are most relevant to their production processes. Margin, however, is not typically released by most corporations. In the absence of specific margin, gross margin can be calculated based on annual statements of public companies. An analysis of gross margin for 6 companies that have products limited to an individual sector can give some general insight into the considerations of margin that they will face. As shown in Figure 2-2, Amgen, Roche (Pharmaceuticals), and Novartis have relatively high gross margins ranging from 66 percent to 82 percent. This is not unexpected for a pharmaceutical manufacturer, because they typically develop low-volume, high-cost chemicals. By contrast, Green Plains, Inc., a biofuels manufacturer, is selling a true commodity chemical and is able to make only a small margin (6 percent). Each of these entities will have different economic factors to consider, for example, feedstock costs will be the dominant driver for a biofuels manufacturer with a small margin. Feedstocks are likely to be less of a concern for the others shown in Figure 2-2.

Today, the use of biology in industry is both strong and increasing. Applications range from traditional areas such as food to newer markets such as renewable energy. In all cases, the potential for increased use of biology will be enhanced by advances in the scale at which biological systems can be engineered.

A FEW EXAMPLES

Artemisinin

Malaria is a global health problem that threatens 300 to 500 million people and kills more than 1 million people annually.[56] Disease control is

Copyright © National Academy of Sciences. All rights reserved.

hampered by the occurrence of multi-drug-resistant strains of the malaria parasite *Plasmodium falciparum*.[57] Artemisinin, a sesquiterpene lactone endoperoxide extracted from *Artemisia annua* L (family Asteraceae; commonly known as sweet wormwood), is highly effective against multi-drug-resistant *Plasmodium* spp. but is in short supply and unaffordable to most malaria sufferers.[58] Although total synthesis of artemisinin is difficult and costly,[59] the semisynthesis of artemisinin from *E. coli*–sourced artemisinic acid, its immediate precursor, could be a cost-effective, environmentally friendly, high-quality, and reliable source of artemisinin.[60]

The production of semisynthetic artemisinin is one of the first success stories for the combined use of metabolic engineering and synthetic biology in the production of a pharmaceutical at industrial scale (see Figure 2-3). As semisynthetic artemisinin is functionally equivalent to the plant-derived drug,[61] it has now been approved by the WHO (World Heath Organization) for the preparation of approved artemisinin deriva-

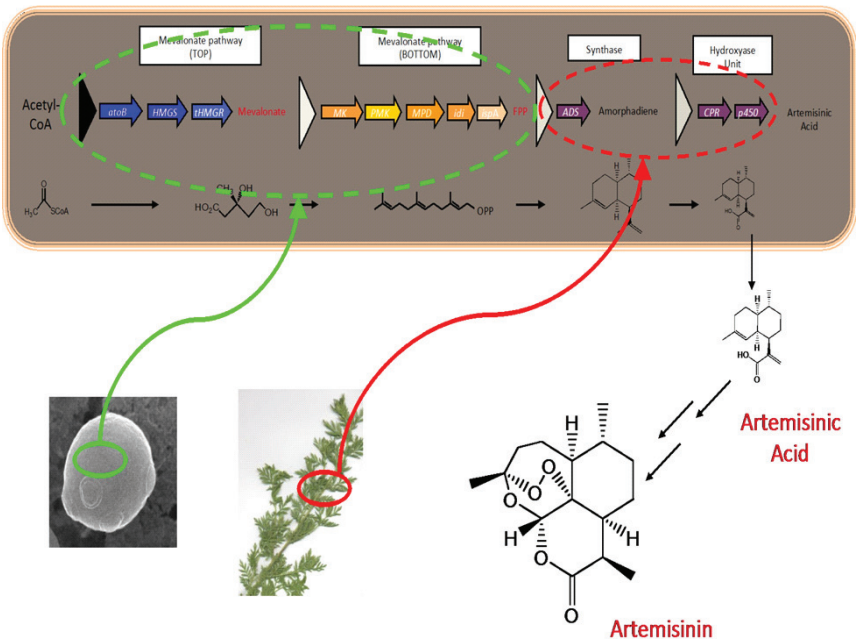


FIGURE 2-3 Overall scheme for engineering artemisinic acid production into *Escherichia coli*. Genes from *Saccharomyces cerevisiae* and *Artemisia annua* were expressed in *E. coli* to transform acetyl-CoA, energy, and reducing equivalents into artemisinic acid, which was then chemically transformed into artemisinin.
SOURCE: Keasling, J. (May, 2012) Synthetic biology and the development of tools for metabolic engineering. *Metabolic Engineering* 14(3), 189-195.

Copyright © National Academy of Sciences. All rights reserved.

tives (such as artesunate) for incorporation into artemisinin-combination therapies.[62] Despite concerns over the amount of semisynthetic artemisinin made available to people in countries affected by malaria,[63] semisynthetic artemisinin is a major success of industrial biotechnology.

Biofuels: Moving to Commercial

Realization of the sustainable benefit of the innovations described above requires commercialization on a massive scale. Advanced biofuels must be economically competitive with existing products, overcoming the primary economic drivers of feedstock price and overall process productivity and yield. While many biofuel candidates have desirable performance characteristics, the potential yield for each product is limited by the theoretical yield of the production pathway,[64] and this limit sets a bottom price that a product can achieve for a given feedstock. Commercialization requires advancing lab-scale processes to yields and productivities that approach the maximum theoretical output of that process (generally >85 percent) and scaling them to reactors that will be >600 cubic meters in volume (~10^6-fold scale-up from most lab-scale fermentations). The engineering of catalysts to reach the yields and productivities to meet economic targets and the scaling of these processes without losing performance are the greatest challenges to commercialization and represent an expensive purgatory for many of the advanced biofuels in the pipeline. To date, only a few promising technologies for advanced biofuels have reached the later phases of commercialization, and these are for the higher alcohols, butanol and isobutanol. Both are anaerobic processes that can leverage existing ethanol facilities, and whose products are either natural or engineered pathways linked to microbial fermentative growth,[65] While tremendous investments have gone into isoprenoid and fatty acid–derived products, commercial success is yet to be demonstrated. Interestingly, as these technologies become more defined they appear to differentiate in regard to industrial host, process and product trade-offs, and feedstock choice. Indeed, as processes move to commercialization, choice of host becomes a critical decision.

Yeast is an attractive host organism because of its robustness, extensive fermentation knowledge, availability of genetic tools, tolerance to industrial conditions and solvents (butanol tolerance >20 g/L), low media pH, and lack of susceptibility to bacteriophage.[20] Yeasts' main limitations are an inability to digest C5 sugars, such as xylose and arabinose, which are present in lignocellulosic biomass; a natural ability to produce ethanol, which may hinder metabolic engineering efforts to produce advanced biofuels; limited synthetic biology tools for pathway optimization; and

Copyright © National Academy of Sciences. All rights reserved.

reduced protein expression levels when compared to *E. coli*, which may limit the flux through biofuel-producing pathways.

Yeast is currently the preferred organism for the production of butanol and isoprenoid-based biofuels.[66] According to publicly available information, at least three large enterprises are pursuing higher alcohol production in yeast using slightly different strategies. Gevo's strategy has been to link the production of isobutanol to anaerobic growth and to select for strains that approach theoretical yields.[65a] This in combination with a stripping of the isobutanol from the broth through continuous flash evaporation has led to a process with over 90 percent theoretical yield and for which a commercial-scale plant is now in place. One strategy being pursued by Butamax is to construct many different metabolic pathways leading to butanoyl.[67] Finally, Butalco, which has strains that metabolize C5 sugars, proposes to use only endogenous genes to improve isobutanol production.[68]

Technologies exploiting fatty acid metabolism are pursuing a variety of host organisms. *E. coli* is being exploited for the production of fatty acid–derived compounds, such as fatty acid methyl esters, fatty alcohols, alkanes, and olefins, directly from carbohydrate in a single-step fermentation. *E. coli* is believed to be preferred for this application because of its exceptionally high rate of fatty acid biosynthesis (3g/L/hr per gram of dry cell weight, based on 30-minute doubling time and 9.7 percent lipid content of cell mass), its natural ability to secrete these products, its natural ability to consume both C5 and C6 sugars, its extensive industrial precedent in the commercial production of metabolically engineered small molecules (1,3-propanediol, lysine, phenylalanine, etc.), and its ease of engineering. *E. coli* does have limitations, such as preference for neutral pH and susceptibility to bacteriophage. Cyanobacteria are also being pursued, but for the production of these compounds from CO_2 in photobioreactors. Technologies for the production of triacyl glyceride (TAG), which is the precursor for biodiesel and renewable diesel (hydrodeoxygenated TAG),[69] have gravitated to oleaginous algae, Generally Recognized As Safe (GRAS) organisms that naturally produce high levels of intracellular oil in both photobioreactors and heterotrophic fermentations.[70] Until any of the promising technologies described above begin producing fuels at prices competitive with existing products, the jury shall remain out as to how effective they can be to meet the goal of widespread adoption.

1,4-Butanediol (BDO)

Genomatica is a San Diego–based startup company that has been active for 15 years. Its focus has been the biological production of chemicals. Its initial focus has been polymer intermediates—specialty monomers

Copyright © National Academy of Sciences. All rights reserved.

that are used as ingredients for plastics and fibers. Genomatica[71] has faced all the challenges that will be faced by others entering this field. Bringing new-to-world technology to a mature industry is difficult. It is important to start with a thoroughly validated value proposition. Product targets must be chosen carefully. Process economics are always a challenge. Feedstock costs and feedstock flexibility are critically important. Conventional sugars are subject to price variability and represent a viable feedstock only in Brazil, India, the United States, and Thailand. Concerns about the use of sugar for fuel (and feedstocks) versus food are also an issue. At this stage biomass-based sugars are challenged to meet the cost and quality needed to produce polymer intermediates. C1 feedstocks look attractive, but today face major challenges for engineering production strains and for production process technology.

As a startup company, Genomatica has faced significant challenges that derive from the development timelines (5-8 years) needed to bring a product to market. The long development time is a contributing factor to the elevated costs to develop a new product platform, on the order of $100 million. Genomatica is focused on reducing both the timeline and costs associated with future product platform developments. Many elements in the product development are beyond the capabilities of any single company. Genomatica has relied on a network of key partnerships to deliver its initial products. Partners have assisted with feedstocks, scale-up, engineering, commercial-scale production, and market-related activities.

In 2008, Genomatica announced its novel bioprocess for the production of 1,4-butanediol (renewable BDO), an important intermediate used in the production of poly(butylene terephthalate) and polyethers. Petrochemical-based BDO is a high-volume chemical, produced in large scale and largely depreciated assets. Current estimates put worldwide production of fossil-based BDO at > 106 T/year. With an installed competitive base, economics for the bioprocess were a paramount consideration.

By 2012, this process had been demonstrated at commercial scale, and in 2012/2013 both Novamont and BASF entered into a license of this GENO BDO technology from Genomatica. Both licensees have begun customer sampling and have communicated their intent to build commercial-scale facilities to produce renewable BDO.

Genomatica constructed the BDO pathway in an industrial *E. coli* production host.[72] Making use of the naturally occurring sugar metabolism via glycolysis and the TCA cycle, Genomatica researchers added the genes necessary to convert Succinyl-CoA to 4-hydroxy-butyrate and on to 1,4-butanediol. As is typical for bioproduction of chemicals, once the basic metabolic pathway is constructed, Genomatica tuned the pathway to enhance production rate and titer, and to improve yield by eliminat-

Copyright © National Academy of Sciences. All rights reserved.

ing metabolites produced via competing pathways. These improvements were critical to achieving the market-competitive process economics.

Following the development of the production host, process engineering, both fermentation and a multistep separation and purification process was needed to produce polymer-grade BDO.

Industrial Enzymes

The industrial enzyme industry rapidly expanded in the 1960s. Most of the early products were produced through fermentation processes, using wild-type hosts, either bacterial or fungal. These hosts are still predominant in today's industry.

Starting from the 1970s, the use of recombinant DNA techniques, combined with protein engineering, had a profound impact on the industry, allowing for more efficient enzyme production, reducing enzyme production costs, and leading to growth of markets and applications. The development of deep-tank, fed-batch aerobic fermentation, replacing the earlier Koji process of culturing of microorganisms on semisolid media, enabled improved efficiency and reduced costs.

Today, the market for industrial enzymes exceeds $5 billion, worldwide.[73] Enzymes enable industries ranging from high-fructose corn syrup and citric acid as food ingredients, to fuel ethanol, to enzyme-based "stone washing" of blue jeans, to more efficient stain removal by detergents (Box 2-2).

GOVERNANCE FRAMEWORK

Industrial biotechnology will need a governance framework[i] that balances important social goals and manifests important values. Governance involves deployment of a variety of policy tools by which an industry's behavior can be shaped, including education of industry actors, industry self-governance through standard setting, accreditation, government standard setting and regulation, public engagement and public scrutiny, tort liability, and other mechanisms for developing safety standards and controls. For governing the industrialization of biology, key goals pertain to safety (risk identification and mitigation) and sustainability. For industrial biotechnology to deliver widespread benefits, it must have low

[i] Governance framework refers to the process of governing the industrial biotechnology, including industry norms, government regulations, and trade associations, among other methods. A more detailed background of U.S. government regulations can be found in Appendix C and at Carter S. R., Rodemeyer M., Garfinkel M.S., and Friedman R.M. *Synthetic Biology and the U.S. Biotechnology Regulatory System: Challenges and Options,* J. Craig Venter Institute: Rockville, MD, 2014.

Copyright © National Academy of Sciences. All rights reserved.

BOX 2-2
Cold-Water Protease Enzyme

Each day, Americans do 123 million loads of laundry. Most of us choose to set our dials to warm or hot water to ensure that stains are removed. We now have a new choice—one that is better for our clothes, lowers our costs, and protects the environment. A new cold-water protease enzyme makes this possible.

Enzymes have been used to improve the cleaning efficiency of detergents for decades. Detergent enzymes account for about 30 percent of industrial enzyme production and are one of the most successful applications of industrial bioscience. Proteases, in particular, are the most widely used enzymes in detergents and facilitate the removal of proteinaceous deposits and stains. Subtilisins are the prototypical group of bacterial alkaline serine proteases that are extensively used in detergents. Historically, they were not particularly stable, or active, in the high-surfactant, alkaline washing processes.

Recently, new protein engineering technology has enabled the tailoring of new protease enzymes that meet the trend toward cold-water washing with uncompromised cleaning performance. This breakthrough came from the use of synthetic genes and massively parallel predictive screens to map key properties such as enzyme activity and stability. In parallel, computation and modeling of physicochemical and structural constraints were used to design and to inform the high-throughput screening.

The new protease delivers the cleaning power of warm-water wash at temperatures below 20°C. The mutations introduced near the enzyme's active site serve to broaden its specificity toward more protein substrates and to cleave them faster. Other mutations serve to increase the affinity of structural calcium ligands, thus conferring stability in harsh detergent environments. The new enzyme demonstrated superior cleaning performance for a broad panel of protein stains, such as blood, grass, and dairy.

Current energy consumption for U.S. residential clothes washing is substantial, amounting to 54,000 GWh, annually. This represents CO_2 emissions of 40 million annual tons. About 80 percent of this energy is used to heat the wash water. Universal adoption of cold-water washing could save as much as 45,000 GWh each year, reducing CO_2 emissions by the equivalent of more than 6 million cars.

environmental impacts, use biological feedstocks sustainably, and operate according to high safety standards with respect to humans, animals, and the environment.

The governance framework should aim to achieve other goals as well, such as trustworthiness. For the industry to be trustworthy, its actors should adhere to high safety and environmental standards, and the public or its governmental representatives must have access to information that allows them to assess industry adherence to such standards.

For the governance framework to have legitimacy in the eyes of the public and the industry, it should be perceived as fair, transparent, effi-

Copyright © National Academy of Sciences. All rights reserved.

cient, and inclusive of diverse viewpoints. These values sometimes come into tension, and all governance systems must make trade-offs between or among them. For instance, efficiency may be hindered by the need for public participation (inclusion of diverse viewpoints) and fair processes. Governance values can also interact synergistically. For instance, public participation and citizen oversight are promoted by transparency, which permits the public to acquire information about industry activities, contemplated regulations, and other relevant issues on which individuals from outside the industry might want to provide feedback. For the governance framework to achieve a reasonable balance among the relevant goals and important values, the framework must be carefully designed. Furthermore, governance can be more or less adaptive, insofar as it explicitly builds on opportunities for learning and adjustment as technologies evolve along with the social, political, and environmental contexts. Adaptive governance poses information problems, however, because somebody must be collecting data on the functioning of the governance processes and on changes in the technology and in the surrounding context for policy makers to know how and when the governance should change.

Regulation is one component of governance—one mechanism by which governments, representing and balancing a broad array of interests, shape behavior in the industry by setting and enforcing standards. Regulation can be very prescriptive, so-called command-and-control style, or it can be more flexible, such as market-based frameworks (e.g., carbon trading), negotiated project-specific licensing, multiparty collaborative planning, and other alternative or second-generation approaches.

The current regulatory environment for approval and control of organisms used in bioprocessing is complex and still developing. Biologically produced chemicals can be regulated by the Environmental Protection Agency (EPA) under the Toxic Substances Control Act (TSCA) or the Federal Insecticide, Fungicide, and Rodenticide Act (FIFRA); the USDA's Animal and Plant Health Inspection Service (USDA-APHIS) under the Plant Protection Act; and the FDA under the Food, Drug, and Cosmetics Act. The applicable U.S. legal regime will depend on a product's intended use rather than on the method by which it was made. The Occupational Safety and Health Administration's (OSHA's) general regulations also address the health and safety of people who work in industries where biotechnology is used. OSHA does not have regulations specific to work with engineered organisms, but it does require that employers create a workplace free from serious, recognized hazards, and it lays out principles and precautions for working with hazardous chemicals. The overlap of legal regimes, and the uncertainty over how and whether regimes will apply to complex engineered organisms, can lead to uncertainty that may hinder

Copyright © National Academy of Sciences. All rights reserved.

Industrialization of Biology: A Roadmap to Accelerate the Advanced Manufacturing of Chemicals

technology investment, development of new products, or more efficient production processes (Box 2-3).

In 1986, the United States developed the Coordinated Framework for the Regulation of Biotechnology, a formal policy for coordinating the activities of the agencies that regulate biotechnology products and research. Where regulatory oversight involves more than one agency, the policy specifies that one agency will take the lead in consolidating and coordinating regulatory reviews. The Coordinated Framework provides a good basis for reconciling the overlapping jurisdiction of various agencies in situations where uncertainty might arise. However, despite the multiple statutory authorities under whom agencies can regulate industrial biology, the existing legal regimes may fail to adequately address some foreseeable risks. Neither the EPA nor USDA-APHIS regulates production processes, and both focus most of their biotechnology-specific regulation on the "premarket" phase of the product life cycle.[ii] It is, therefore, unclear whether these agencies have adequate authority or expertise to ensure that proper containment and disposal procedures will be used at commercial manufacturing facilities once a manufacturer engages in legal commercial production of a biologically produced chemical. It is also unclear whether either agency has authority to sanction a firm that creates a public health or environmental hazard by inappropriately disposing of host microorganisms, waste biomass, or co-products.

It is not clear whether any agency has adequate authority to oversee worker, environmental, or public safety when research to design and test industrial microorganisms is privately funded and conducted by commercial firms. Working with engineered microorganisms requires taking precautions commensurate with the level of risk. The design of manufacturing facilities should include appropriate containment features, and firms should include biosafety considerations in their standard practices. Firms that have long been involved in the production of chemicals by ordinary chemical synthesis will have a safety culture and safety engineering in their facilities, but their expectations and practices may have to evolve to accommodate biological processes. There currently are no unified federal standards for commercial production of chemicals through biological routes; there is nothing for industry that is comparable to the National Institutes of Health (NIH) Guidelines.

Similarly, the Centers for Disease Control and Prevention and NIH have created the Biosafety in Microbiological and Biomedical Labora-

[ii] It may be possible that some waste streams could fall under the Resource Conservation and Recovery Act (RCRA); however, the EPA implementation under this act provides two categories of regulated waste, "Listed Wastes" and "Characteristic Wastes," each with very specific lists of chemicals and their concentrations that may be regulated. See 40 CFR §261.

Copyright © National Academy of Sciences. All rights reserved.

BOX 2-3
Biosafety Design Considerations

The U.S. Department of Health and Human Services has put forward screening framework guidance for providers of double-stranded DNA. These guidelines have been implemented by the International Gene Synthesis Consortium's (IGSC's) harmonized screening protocol, which consists of both gene sequence and customer screening approaches to promote biosecurity. Biosafety and biosecurity gene sequence screening approaches can readily be incorporated into the integrated design toolchain. Although IGSC's protocol does not enable companies to identify and predict problematic biosafety and biosecurity properties that emerge when multiple components come together in an organism or bioprocess, it does when an individual component is of concern. A similar problem is the development of algorithms to assign or discern design attribution for legal or law enforcement purposes. Methods and tools that address emergent biosafety and biosecurity concerns, including tools that identify the metabolic dependencies and physical containment properties of designed organisms, should also be included in the integrated design toolchain. Furthermore, methods and tools for identifying hazardous chemical properties of feedstocks, intermediates, and products (from material safety data sheets or other information sources) will be very important for assessing the biosafety aspects of an overall bioprocess.

tories (BMBL) manual, which complements and expands on the NIH Guidelines. Now in its fifth edition, the BMBL sets forth principles for risk assessment and containment of hazardous biomaterials and has become the code of practice for biosafety. No agency, however, has specified the precautions to be taken in privately funded research and product development, or mandated adoption of the BMBL or similar principles.

"In modern democracies . . . the public plays a central role in determining how science is funded, used, and regulated."[74] The governance regime for industrializing biology should enhance opportunities for the public to engage with regulators and industry. By doing so, the regime can manifest the values of trustworthiness, transparency, and participation and can help ensure that science serves the greater public good. Such engagement enables members of the public to learn about technologies and oversight mechanisms, and to play a role in governance processes. There are numerous methods of public engagement, and the optimal approach will vary with the particular social context, technology, and applicable regulatory regime(s). This report does not recommend particular approaches for public engagement but does emphasize its importance.

Another set of social factors that will influence the pace and direction of industrial biology is the balance between open innovation and informa-

Copyright © National Academy of Sciences. All rights reserved.

Industrialization of Biology: A Roadmap to Accelerate the Advanced Manufacturing of Chemicals

tion sharing, on the one hand, and proprietary product development, on the other hand. In the context of synthetic biology, a great deal has been written about the virtues of open versus proprietary research as pathways for advancing the field. This report does not extend that previous work. Here, we merely note two important points. First, "open science" does not necessarily require that findings be put in the public domain, although that is one approach. Open science can also be promoted by licensing intellectual property to create a science commons, or otherwise to protect broad access to a discovery or product. Second, patents are often viewed as one component in the proprietary and private property–driven approach to innovation. To date, patents have played a significant role in attracting and protecting investments in the biotechnology and chemical industries; however, the patent law has recently changed in ways that constrain opportunities to patent some biotechnology inventions and processes. The committee did not attempt to predict the degree to which recent legal changes will affect patenting of this report's subject matter—complex engineered microorganisms and industrial production processes using those microorganisms. Patent protection will still be available for some important products and processes in industrial biology, but recent legal changes could affect the number and nature of patents, which may in turn influence the ways businesses organize and collaborate. To promote the industrialization of biology, academic and industry scientists in synthetic biology and related fields will need to determine an acceptable balance between open and proprietary approaches to innovation.

Related to balancing open and proprietary science is the increasingly prevalent practice of data sharing. In other contexts (e.g., biomedical science), sharing of precompetitive, detailed data has proven beneficial for a broad array of stakeholders, from academia to industry. Such sharing, which is being encouraged or mandated[75] by many public funders of science and across the U.S. government, requires that researchers make available types of data beyond the summarized, aggregated, highly analyzed data typically found in publications. Stakeholders in industrial biotechnology might also identify opportunities to advance the field through appropriate data sharing.

Copyright © National Academy of Sciences. All rights reserved.

3

Vision of the Future: What New Chemicals Could Be Made?

To date, most successful commercial products were carefully selected for their manufacturing via biological synthesis. As discussed in the preceding chapter, a large degree of chemical space is already known to be available for chemical manufacturing. The vision of the future put forth herein is one where biological synthesis and engineering and chemical synthesis and engineering are on par with one another for chemical manufacturing.

The recommendations and roadmap goals outlined throughout this report were all conceived in the context of this overarching vision and are designed with the understanding that, in order for the industrialization of biology to be realized, the use of biological and chemical routes must be thought of as equals. That is not to say that each would be used interchangeably, but that biological routes would be included the same way individual chemical reactions are considered when developing a synthetic route.

Determining whether both biological and chemical routes should be set on equal footing and understanding the potential diversity of chemical products that could be produced are critical to the industrialization of biology. The majority of this chapter is devoted to answering these questions.

WHAT CHEMICALS COULD BE MADE?

The industrialization of biology offers prospects not only for new commercial production and process methods but also for the opening

Copyright © National Academy of Sciences. All rights reserved.

of novel chemical space for the discovery of functional molecules (e.g., pharmaceuticals, materials, fuels). As discussed in Chapter 4, enzyme- or cell-based synthetic approaches can provide compounds ranging from drop-in replacements—made via processes with economic or environmental advantages over previous synthetic methods—to new structures with improved function or performance relative to their chemical precursors. Targets at either end of this spectrum are subject to very different factors with respect to technological and economic factors influencing their development (Figure 3-1). The Department of Energy's report, *Top Value Added Chemicals from Biomass,* [76] provides an excellent discussion of potential targets for the biological production of chemicals.

Both commodity and specialty chemicals can be approached using biological methods but should take advantage of the unique properties of living systems. For commodity chemicals, targets need to add economic value to the starting carbon source (e.g., glucose or cellulose) and can include preexisting high-volume chemicals, biologically sourced

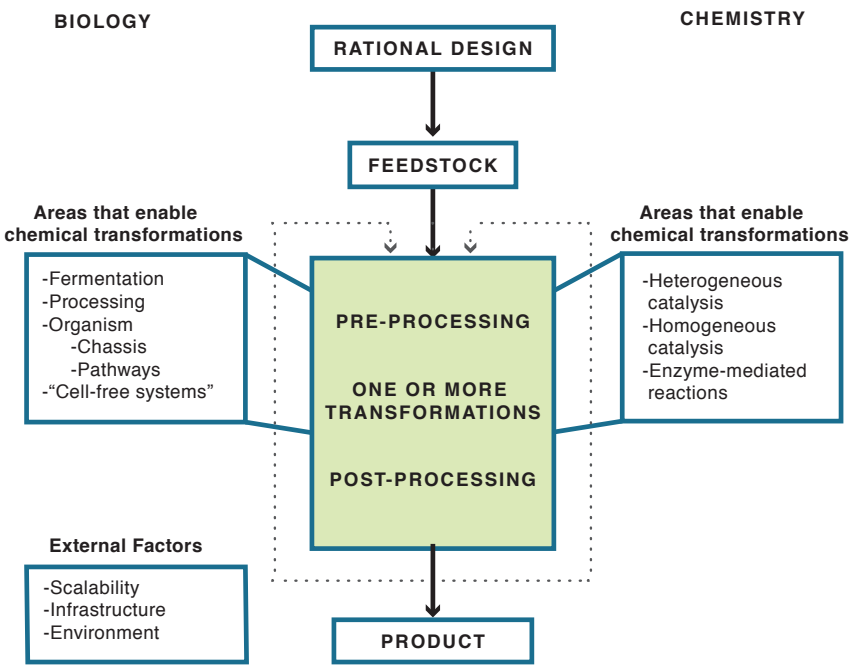


FIGURE 3-1 Chemical manufacturing flowchart, showing the report's conceptual schema of the chemical manufacturing process and highlighting the techniques of both biology and chemistry that enable chemical transformations.

Copyright © National Academy of Sciences. All rights reserved.

precursors that may be converted to the desired product through simple chemical transformations, or new structures. These types of targets can provide both economic and environmental benefits via the ability of cells to utilize biomass-derived carbon sources, grow in aqueous media, and carry out multistep transformation of substrate to product in a single reactor. Specialty or fine chemical targets yield more flexibility in approach and cost of manufacture based on their higher value. Indeed for many complex natural products, there may be no existing chemical method for their commercial manufacture. As such, a biological route can provide new access to the target or a semisynthetic intermediate. In addition to multistep cellular transformations, single enzyme-based transformations may also be important in this area because the utilization of enzymatic regio- and stereoselectivity can greatly streamline a chemical process.

The continued development of biotechnology related to chemical synthesis also enables new routes to discovery when combined with the more mature area of chemical synthesis, because it allows opportunities to mix orthogonal structural space. In this regard, living systems produce a wide range of compounds that often demonstrate relatively low structural overlap to those produced via synthetic methods (Figure 3-2). Much of this divergence in structure arises naturally from differences in building block availability and assembly. In general, synthetic compounds are ultimately derived from petrochemical sources with substitution patterns controlled by the selectivity of chemical reagents but can take advantage of a broader coverage of elemental composition, functional groups, and reaction space. In contrast, large classes of biological metabolites often share a biosynthetic logic in their assembly but can utilize the selectivity of enzymes to produce highly complex structures. As such, the development of methods for combining biological and synthetic chemistry could

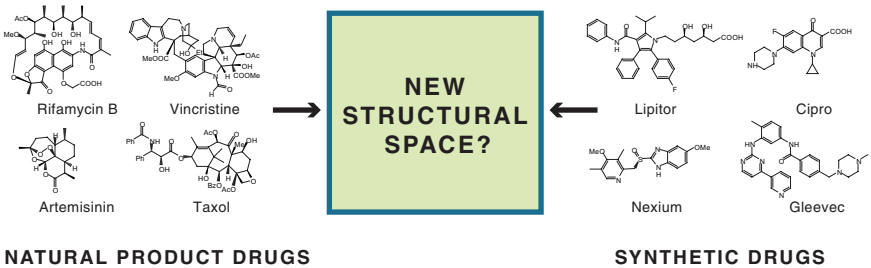


FIGURE 3-2 Low-structural-overlap compounds produced by living systems and synthetic methods. Some of the current targets of chemical manufacturing are identified.

Copyright © National Academy of Sciences. All rights reserved.

allow the expansion of the accessible structural space for screening new compounds for functional properties.

Natural Products

Natural products and their derivatives remain an important resource for the discovery of new bioactive compounds. They represent a significant portion of new chemical entities while also playing an important role in the identification of druggable targets and pathways for development of synthetic compounds.[77] Their success as pharmaceutical agents is likely derived from their natural evolution toward structures optimized to bind macromolecular biological targets, which requires a high structural complexity that can be oftentimes difficult to replicate in a synthetic compound. As such, it is estimated that it is several orders of magnitude more likely that a natural product will bind a cellular target compared to a synthetic compound. However, the use of natural products as lead compounds is quite challenging, because they are difficult to synthesize and structurally optimize for appropriate pharmacokinetic behavior. As such, natural products routinely remain underutilized in the drug discovery pipeline, and advances in biotechnology on many different fronts could greatly alter this landscape.

Genes to Products

It is widely accepted that natural products contain an enormous structural diversity. As previously discussed, this structural diversity typically accesses structural space outside of chemically synthesized compound libraries, yet poises natural products for macromolecular target binding. Thus, the inclusion of new natural product structures and their pharmacophores is important for expanding the available space for discovery. However, there are several roadblocks to achieving this goal: most natural products are produced at extremely low yield in their native host; the majority of genes encoding the production of natural products are silent, that is, displaying no detectable phenotype; and most environmental isolates are not culturable under laboratory conditions. Thus, new methods of moving from gene sequence to product are important and could potentially be provided both by the ability to synthesize and express large sets of genes in model hosts and by rapid approaches to domesticate environmental hosts.

Copyright © National Academy of Sciences. All rights reserved.

Natural Product Analogs

While the complexity of natural product structures serves as an advantage in their use as lead compounds, it rapidly becomes a disadvantage given that most lead compounds need to be optimized for proper potency, cross-reactivity, and pharmacokinetic behavior. Semisynthetic approaches, or the direct chemical modification of a natural product or biosynthetic intermediate, are limited in their ability to achieve a broad range of structural transformations of the natural product given their functional group density and lability to harsh chemical reaction conditions. Thus, enzymatic or biosynthetic modifications can provide a new route into structural diversification of natural products for tuning their performance as drugs. In this regard, the identification and characterization of tailoring enzymes that may oxidize, cross-link, or ligate on new groups to core structures are useful. In addition, methods to feed in different building blocks to the biosynthetic machinery can generate much-needed variations in the core structure. Advances in manipulating core structure and tailoring can further help to create diversity by introducing orthogonal chemical handles (e.g., halogens) or new linkage locations (e.g., amines) for downstream enzymatic or chemical reactions.

Tapping New Structural Diversity

Beyond the exploration of natural products classes with known genetic signatures, such as polyketides, nonribosomal peptides, and isoprenoids, there are many structural cores that have yet to be identified or genetically annotated. Among these are nitrogen-rich compounds of varied structure, including alkaloids, which are needed to augment our pool of compounds, as the more well-characterized classes of natural products tend to be oxygen rich (e.g., polyketides and isoprenoids). Modified peptides, both ribosomally and nonribosomally encoded, also represent interesting families for further characterization. Improvements in gene prediction, chromosome modification, host domestication, and small-molecule analysis can aid in this endeavor.

Advanced Molecules

For the development of advanced molecules, the relatively new interface between synthetic chemistry and biology needs to be further enlarged because synergy between these two areas can greatly accelerate the discovery process. For example, microbial fermentation can generate previously untapped monomers for polymer production, but the synthesis and characterization of the resulting materials is equally essential for identifying new properties or functions. Conversely, the analysis of

Copyright © National Academy of Sciences. All rights reserved.

Industrialization of Biology: A Roadmap to Accelerate the Advanced Manufacturing of Chemicals

synthetic bottlenecks in the production of complex targets could allow us to focus on engineering specific enzyme families with the highest potential to enable pharmaceutical research and production. Research directions in this area could include but are not limited to engineering enzymes or pathways for the biological production of complex building blocks, stereo- and regioselective transformation of synthetic building blocks, reactions involving key elements or functional groups, and catalysis of new C-C bond-making reactions. In addition, computational tools to combine biological and synthetic reaction space to analyze the efficiency of different hybrid preparation routes are also necessary to identify specific paths forward for further development.

Engineering the Production of Complex Building Blocks

Building blocks with a high density of stereocenters or functional groups can often be derived from biological sources. Some examples include isoprenoids, sugars, and other classes of metabolites, which are used as synthetic starting materials but also can affect the price and availability of the final product. One example was previously presented for the production of artemisinin from a microbially derived semisynthetic intermediate. In this case, both the intermediate and the target compounds are natural products and synthetic chemistry is used to scale up a biologically difficult reaction that ultimately opens access to a low-cost antimalarial drug.

A different type of advancement in this area can be illustrated in the commercial synthesis of oseltamivir (Tamiflu), a synthetic antiviral compound prescribed for avian flu that is made from shikimic acid. This biosynthetic intermediate is produced by microbes and plants but at such low levels that its availability controls that of Tamiflu. As a result, a strain of *E. coli* was engineered to highly overproduce shikimic acid and has greatly increased access to Tamiflu.[78] In contrast to the example of artemisinin, Tamiflu is not itself a natural product but simply takes advantage of the existing stereocenters in a biological metabolite to reduce the cost of the target compound. Without using the innate stereochemistry of shikimic acid, the synthesis of Tamiflu would likely require several steps resulting in higher prices and decreased availability.

Beyond traditional natural products, biological systems are also uniquely poised for the generation of other types of structures with challenges of regioselectivity and stereoselectivity. One example is represented by polysaccharides, which can be important modifiers of bioactive agents. Their chemical synthesis requires extensive and laborious protection and deprotection routines to achieve regioselective assembly but can potentially be put together instead from their unprotected parental sugars using glycosyl transferase enzymes.

Copyright © National Academy of Sciences. All rights reserved.

The ability to take these routes and use computational analysis or software to identify these points of overlap rather than relying on human insight could greatly accelerate similar projects. By extension, large-scale analysis of various synthetic routes could also help to identify classes of molecules or patterns of substitution that could be produced using biological systems as useful synthetic building blocks.

Engineering the Stereo- and Regioselective Transformation of Synthetic Building Blocks

Enzymes excel at selective transformations and have been used as reagents for individual transformations of synthetic intermediates when chemical reagents are difficult to optimize for a particular reaction. In many cases, the adoption of an enzymatic step could streamline the synthetic route, which may utilize a number of additional steps in order to avoid a particularly challenging problem in asymmetric catalysis. In this case, families of enzymes, such as ketoreductases, esterases, peptidases, and transaminases, have been well developed for these applications.[79] One important advance was achieved for sitagliptin (Januvia, Merck), where a partnership between enzyme engineering and chemical synthesis led to the insertion of a transaminase-catalyzed step, thereby reducing the step count in its preparation.[80] Currently, we are limited to a select group of enzyme families that are known to be naturally accommodating to wide ranges of substrates, which correspondingly limits the scope of transformations that are targeted for this approach. Thus, the identification and implementation of new target enzyme families and transformations could greatly accelerate advances in this area.

Catalysis with Key Functional Groups and for New C-C Bond-Making Chemistry

Compared to the chemical reaction scope, cells typically use a smaller set of functional groups and lower diversity of C-C bond-forming reactions, because enzymes can use substrate and product selectivity to form the correct bond amidst many different possibilities. In comparison, synthetic catalysts tend to use functional group orthogonality and/or protecting groups to achieve selective bond formation. Thus, an interesting area of development could incorporate enzymes to install rarer elements or synthetic functional groups for function or as synthetic handles for downstream chemical catalysis. In addition, new enzyme classes could also be evolved to catalyze C-C bond coupling reactions from synthetic building blocks. Some examples of useful functional groups could be fluorine to tune bioactivity and pharmacokinetic properties[81] as well as orthogonal synthetic handles such as other halides (X = Cl, Br, I), nitriles, boronic

Copyright © National Academy of Sciences. All rights reserved.

Industrialization of Biology: A Roadmap to Accelerate the Advanced Manufacturing of Chemicals

acids/esters, or alkynes for cross-coupling reactions such as those developed by Heck, Stille, Negishi, and Suzuki; Sonogashira; and Buchwald-Hartwig. One example where synthesis has inspired the development of new reaction chemistry involves the engineering of cytochrome P450s for the insertion of C or N rather than O to form cyclopropane or aziridine rings.[82] In addition, the exploration of biodiversity leads to the discovery of new families of enzymes that could be useful for synthetic applications, such as those catalyzing Pictet-Spengler[83] or Diels-Alder[84] reactions.

Polymers

Polymers are organic macromolecules made of repeating monomer units that are valued for their tunable functional and structural properties. Indeed, polymers are used for a broad range of applications, from use as plastics, rubbers, fibers, and paints to controlled drug release and electronic displays. They are derived from biological sources, such as natural rubber, silk, and cellulose, as well as synthetic origins, such as polyethylene, polystyrene, nylon, silicone, and polyvinyl chloride. Polymer properties are controlled by many variables, including monomer structure, bonding between monomers, tacticity, average molecular weight, polydispersity, and branching for homopolymers. Co-polymers made up of more than one monomer type expand this range even further to include attributes such as monomer arrangement (periodic, statistical, or random) or co-block characteristics. These structural features influence intra- and interchain microstructure that in turn control bulk material properties that are important for function, such as melting temperature, crystallinity, glass transition, tensile strength and elasticity, transport behavior, and electronic response.

The relationship between chemical and materials properties has been well explored but remains challenging to predict from a new monomer given the breadth of different polymers that can be accessed. At this time, many of the commodity polymers are constructed from building blocks that can be prepared from readily available petrochemical sources. However, living systems provide a vast array of bifunctional compounds that can be used as monomers, the majority of which have yet to be tapped for polymer synthesis. This section covers opportunities in metabolic engineering for existing and new monomers and polymers.

Existing Monomers

One approach is the direct replacement of existing monomers derived from petrochemical sources with the same structure made by microbial fermentation. A key advantage in this strategy is that a drop-in replacement already has a current market demand. However, two major chal-

Copyright © National Academy of Sciences. All rights reserved.

lenges are that it can be difficult to either ferment the monomer at a competitive price with the existing competitor given their low cost and the capital investment associated with building plants for a new process or to displace significant volume of the petrochemically derived monomer because of their high usage. An example of a microbially sourced monomer currently on the market is ethylene (or "bioethylene"). Ethylene represents one of the highest-volume monomers in use today (~140 million tons per year) because it is found in approximately half of all plastics as a homopolymer (e.g., high- and low-density polyethylene) and co-polymer (e.g., polystyrene, polyvinyl chloride, and polyethylene terephthalate).[85] Bioethylene is produced by microbial fermentation of sugar to ethanol followed by chemical dehydration and is produced at a large scale (~200 kilotons [kt] in 2013).[85] For comparison, if all the ethanol produced by microbial fermentation for transportation fuels were converted to bioethylene, then this volume could reach approximately 25 percent of the annual ethylene feedstock currently needed.[85] While it can be produced at a similar cost as petrochemical ethylene, the price depends greatly on the cost of sugar, which is currently a highly volatile feedstock. Other examples of monomers in the development pipeline at this time are butadiene (from dehydration of 1,4-butanediol, Genomatica),[86] acrylic acid (from dehydration of 3-hydroxypropionic acid, Cargill, OPXBIO; or lactate, Myriant),[87] and isoprene (Dupont and Goodyear). All three of these monomers are targeted toward large-volume markets. Other similar targets can be identified by examining the commodity chemicals market and could be prioritized by their biosynthetic complexity as well as the range of polymer products, because niche markets could potentially be easier to move toward biosourced monomers.

New Monomers

Another approach is the development of new monomers to produce novel polymers. Although the market for these new monomers is more difficult to characterize, they do not need to directly compete with an existing product made through a mature process. This approach also allows polymer chemists to explore greater chemical space to improve the material properties of polymers or to discover entirely new functions. In general, many commercial polymers have been developed from readily available petrochemical feedstocks and optimized for their intended application by controlling various parameters as discussed above. As such, compounds falling into the same chemical class as known monomer feedstocks, but with different substitution patterns, could be funneled into the same polymerization pipeline but impart altered properties to homo- and co-polymers.

Copyright © National Academy of Sciences. All rights reserved.

One example is Bio-PDO®: Before the advent of Bio-PDO, 1,3-propanediol (PDO) was considered a specialty monomer but still fell into a chemical class with known polymeric products generated from structurally similar but more readily available diols, such as ethylene glycol or propylene glycol. However, the new microbial process for its production enabled greater access to this monomer and led to the development of new polymers that have earned significant market share.

A second example that highlights the interplay between chemistry and biology is polylactic acid (or polylactide, PLA) made from the microbially sourced lactic acid monomer, developed by NatureWorks.[88] Similar polyesters, called polyhydroxyalkanoates (PHAs), are made by microbes for carbon storage from a variety of 3-hydroxyacids and are thus biodegradable.[89] As such, a significant research effort was made to develop plant- or microbe-based processes for its industrial production of PHA and PLA because a biosourced and biodegradable polyester could have interesting applications. The underlying biology of these systems involved in controlling important parameters including chain length and polydispersity is quite complex and remains insufficiently understood for rapid engineering. As a result, the polymer properties of bioengineered PHAs were difficult to tune compared to synthetic polyesters made from mature chemical processes. NatureWorks developed instead a process based on the chemical polymerization of lactate, which is also a 3-hydroxyacid even though it is not typically a physiological monomer for PHA biosynthesis. Using this approach, their overall process could rely on a robust fermentation process for the monomer, because many organisms are known to ferment glucose to lactate at near quantitative yield, and a well-characterized chemical polymerization to PLA, which allows for control over its material properties while maintaining the biodegradability of the polymer.

Taking this bioinspired approach, there already exist many classes of bifunctional small-molecule metabolites that fall into categories of known monomers or monomer precursors that can be processed within a few downstream chemical steps (e.g., dehydration, oxidation, and reduction). For example, different combinations of carboxylic acid, ester, ketone, aldehyde, amine, alcohol, olefin, and epoxide functional groups could be directly incorporated into polymers such as polyesters, polyamides, nylons, polyolefins, synthetic rubbers, polyethers, and others. Because small structural changes in monomer structure, such as stereochemistry, substitution patterns, or spacing between functional groups, can greatly affect polymer performance, these monomers could be explored for their behavior in homo- and co-polymers. The biosynthesis of some of these monomers could then be directly engineered from existing pathways or could also be greatly diversified by engineering pathways to accommodate greater structural diversity.

Copyright © National Academy of Sciences. All rights reserved.

Polymerases

An interesting area for the development of high-performance polymers could be templating or engineering the assembly of monomer units using polymerase enzymes to regulate important features that may be difficult to control using chemical catalysts, such as sequence, tacticity, block size, or branching. While the enzymatic selectivity filter for some of these properties may not yet be sufficiently understood for engineering purposes, the ability to precisely tune these properties could transform the scope of polymer behavior that can be achieved.

Protein polymers offer a key example of how precise control over sequence and chain length can impart key material properties. There are many examples of polypeptide-based materials, such as silks, wool, or collagen, which are genetically encoded and synthesized via the ribosome. Using the 20 canonical proteinogenic amino acids as well as others, an enormous amount of structural and functional space of the resulting polyamides can be examined. Currently, there already exists a large body of work on peptide-based materials and their self-assembly into materials with unusual properties.[90] In addition to side-chain diversity, it may also be possible to examine features such as tacticity by altering interchanging L- and D-stereochemistry around the alpha carbon or branching from side-chain functional groups by post-translational attachment of different structures. Another area of research is the use of the templating afforded by the ribosome to make different classes of polymers, such as polyesters.[91] A key challenge in this area for industrial-scale production is the development of robust methods for engineering the export of the target polypetides to allow for their scalable collection as individual polymers or as fibers.

In addition to genetically templated macromolecules, such as polypetides, polymerases can also catalyze the assembly of alkanes (fatty acid synthases), polyketide-based structures (polyketide synthases), mixed peptide and ketide structures (hybrid nonribosomal peptide and polyketide synthases), polyesters (PHA synthases), oligosaccharides (glycosyl transferases), and others. All of these structures can be produced using a broad range of monomers, which can either be selectively or nonselectively chosen by the particular enzyme. Developing a better understanding of how these systems control polymer structure and monomer selection could allow us to selectively generate either new monomers or polymers with high functionality.

Polymers for Templating the Formation of Inorganic Materials

In addition to the production of purely organic materials, biological systems can also use the self-assembly of these biopolymers to template

Copyright © National Academy of Sciences. All rights reserved.

the formation of inorganic and composite organic and inorganic materials made of calcium, silicon, iron, manganese, and copper. Some naturally occurring examples include bone, nacre, diatom frustules, and magnetite nanocrystals. In these cases, the nanostructure of these materials is highly controlled in terms of chemistry (e.g., composition and mineral structure) as well as structure (e.g., size and shape).[91b, 92] This approach has inspired the development of methods to evolve polymers, such as peptides, to template and control the shapes of different minerals. A major challenge in this area again is the consideration of cost in the scalable production of materials through this route, which could be correspondingly improved by the development of methods for extracellular delivery of the templating agents.

BUSINESS MODELS FOR FUTURE INDUSTRIAL BIOTECHNOLOGY

The term "vertically integrated development" is used to describe a future in which biomanufacturing process research and development is performed by vertically integrated corporations that develop the entire bioprocess from end to end: from feedstock sourcing to organism engineering to manufacturing and sales. In this future, successful industrial biotechnology companies are comparable to Intel: they encompass everything from design to manufacturing.

The term "horizontally stratified development" is used to describe a future in which there is a stratified industry for biomanufacturing process development in which different companies specialize in different steps along the supply or value chain. For example, one company may focus on feedstock sourcing, another on organism engineering, another on scale-up and manufacturing, and still another on marketing and sales. In this future, the industrial biotechnology industry is comparable to the PC industry of the 1990s in which different companies manufactured the hardware components, assembled the computers, wrote the operating system, and developed the software applications.[93]

The term "centralized production" is used to describe a future in which the biomanufacturing of chemicals occurs in a handful of very-large-capacity biorefineries that take advantage of economies of scale to eliminate inefficiency and produce chemicals with razor-thin cost margins and at volumes sufficient to meet world demand. In this future, chemical biomanufacturing looks similar to the oil and petrochemical industry in which there has been a persistent trend toward ever fewer and ever larger oil refineries over the past two decades.[94, iii]

[iii] In 1994, the United States had 179 operable crude oil refineries capable of distilling just over 1.5 million barrels per day. In 2014, the United States was down to only 142 operable refineries but had a distillation capacity of nearly 1.8 million barrels per day.

Copyright © National Academy of Sciences. All rights reserved.

The term "distributed production" is used to describe the local, small-scale manufacturing of chemicals. In this future, these specialized biorefineries might use geographically co-localized feedstocks and produce only enough product to meet local demand. In this future, chemical biomanufacturing looks similar to the home brewing or microbrewery industry of today.[95]

Examples of these definitions are presented for comparison in Box 3-1.

Although the envisioned future is presented here as discrete scenarios for simplicity, it should be noted that there exists a continuum of possibilities between these scenarios. For example, distributed production may result in biorefineries of sufficient size to supply a nation, a region, a city, a neighborhood, or just a single household. As a second example,

BOX 3-1

Vertically Integrated Development

Research and design for biomanufacturing is performed by corporations that develop the entire process end-to-end: from feedstock sourcing to organism engineering to manufacturing and sales.

Apple Inc. is a contemporary example of this, with design, operating system, sales, and service being provided by Apple Inc. itself.

Horizontally Stratified Development

Research and design for biomanufacturing is performed by different companies that each specialize in a different step along the production process.

The PC industry is a contemporary example of this, with design, components, assembly, operating systems, software, sales, and service being provided by specialized companies.

Centralized Production

Biomanufacturing occurs in a handful of very large capacity facilities that take advantage of economies of scale to eliminate inefficiency and produce chemicals with thin margins and at volumes sufficient to meet world demand

The petroleum industry is a contemporary example of centralized production.

Distributed Production

Biomanufacturing occurs in many local, small-scale facilities, potentially using geographically co-localized feedstocks and producing only enough product to meet local demand.

The home brewing or microbrewery industry is a contemporary example of distributed production.

Copyright © National Academy of Sciences. All rights reserved.

the degree of stratification in the horizontally stratified industry may vary. Organism engineering may constitute a layer within the supply chain, or it may be further stratified into design firms, DNA synthesis and assembly firms, and organism testing and validation firms. Finally, even in a future where horizontally stratified development is the norm, it may still result in centralization within particular strata—akin to Microsoft's dominance of operating systems on the PC in the 1990s.

Moreover, these discrete scenarios for the future are not mutually exclusive. It may be that some sectors of industrial biotechnology may lend themselves to centralized versus distributed production. For example, specialty ingredients—high-value chemicals that make up fast-moving consumer goods—will not require biorefineries at comparable scale to those needed for fuel production because the volumes needed to satisfy consumer demand are several orders of magnitude lower than for fuels. So the very nature of the specialty ingredients industry (hundreds of ingredients each at smaller volumes and higher price points) versus the commodity chemicals industry (dozens of chemicals at very large volumes with thin margins) may result in a hybrid chemical bioproduction model.

Furthermore, the degree of stratification or centralization of an industry can swing back and forth over time. As a particularly relevant case in point, DNA sequencing began as a highly distributed technology that was largely performed by individual researchers and labs. Then, driven in part by the Human Genome Project and the desire to push down the cost per base pair of sequencing DNA, there was a move to sequencing centers such as the Broad Institute of MIT and Harvard, the Sanger Centre, the Beijing Genome Institute, and the Department of Energy Joint Genome Institute. Today, while centralized sequencing centers continue, the falling costs of sequencing instruments are making genome sequencing at the individual laboratory scale possible once more.

Copyright © National Academy of Sciences. All rights reserved.

4

How Do We Get There?

OVERVIEW OF ISSUES

In order to realize a future of biological, chemical, and combined approaches being viewed as equally viable options for chemical manufacturing, a number of technical challenges must be overcome. As discussed previously, the use of biological systems for chemical manufacturing has already attained fairly widespread use in some specific sectors, but by comparison to traditional chemical manufacturing it is still a relatively small market. The promise shown by these previous successes, however, is significant.

One key technical consideration that has been less well integrated into planning and processing for bio-based methods than for traditional chemistry is the ability to model and fully understand the entire manufacturing process when considering the use of biological systems for chemical manufacturing. The characteristics of biological behavior make this a daunting task, but relatively recent advances in life sciences and chemical engineering make it attainable if the many factors that could cause progress to stagnate are avoided. In order to facilitate biomanufacturing for chemical production, a series of conclusions and roadmap goals are presented and discussed in this chapter. The discussion is organized into three broad categories: feedstocks, enabling transformations, and integrated design toolchain.

The feedstocks section discusses the promising array of feedstocks currently used in manufacturing as well as the array of opportunities that are possible with key technological advances. Starch and other simple

Copyright © National Academy of Sciences. All rights reserved.

Industrialization of Biology: A Roadmap to Accelerate the Advanced Manufacturing of Chemicals

sugars obtained from biomass are the most widely used feedstocks today, and the use of cellulosic biomass is expanding. There are still many challenges associated with using recalcitrant cellulosic material for manufacturing, but there are potential solutions to this issue as discussed herein. Although the discussion is focused on different forms of biomass, the discussion is not limited to biomass. There is active work in facilitating the use of syngas, methane, and carbon dioxide in manufacturing as well.

The enabling transformations section discusses the science, technology, and engineering knowledge and tools required to transform the feedstock material into a useful product or intermediate. One of the major engineering considerations is related to fermentation and processing that are required for biological production of chemicals. Fermentation can be facilitated in many ways, but it typically represents a large capital expense that must be overcome in order to begin production. To mitigate this capital expense, the ability to reliably and efficiently scale up processes is very important.

This section continues to discuss the research and development needed to facilitate chemical transformations. The majority of this section discusses synthetic biology and the use of chassis and pathways to develop microorganisms for use in chemical manufacturing. Although this work is ongoing in a number of sectors, the use of microorganisms for chemical manufacturing could be more widespread. This portion describes the priority research needs to enable chemical transformations using biological systems.

The final section discusses the overall needs in measurements science and technology in relation to the research and development needs discussed in this chapter.

FEEDSTOCKS

New Sources of Carbon

Carbon in the form of fermentable sugars is the primary raw material, and often the largest single input cost, for the biological production of chemicals. In the case of large-volume chemicals, sugar costs can represent the majority of the total product costs. In the extreme case of biofuels, sugar costs represent as much as 65 percent of the total product costs.[96] By contrast, for industrial enzyme and specialty chemical production, the overall cost of the carbon source is a small fraction of the total costs. The feedstock cost is so small for these products that changes to feedstock price are negligible.

For today's fermentation, the source of carbon is overwhelmingly dextrose derived from the starch in grain. In Brazil, abundant sucrose

Copyright © National Academy of Sciences. All rights reserved.

from sugarcane is used instead of dextrose. For the biological production of chemicals to reach its full potential, more abundant, more diverse, and less costly sources of carbon are needed.

Cellulosic material derived from agricultural residues, forestry by-products, and even dedicated energy crops are both abundant and diverse. Conversion of cellulose to fermentable sugars is the subject of active research and development. A number of challenges must be overcome if cellulose-based sugars are to become fully substitutable for starch-derived sugars, and with a cost advantage.

Today's agricultural economy has well-functioning markets for grain trading, and a well-established infrastructure for the production, transportation, and storage of the commodities. None of this exists for fermentable sugars derived from cellulosic feedstocks. As the first cellulose-based ethanol plants are coming on line, individual plants develop their own technology and (local) markets for originating cellulosic material. The cost of the cellulosic feedstock must be kept low—much less than $100/ton—in order for it to be a viable alternative to grain.

Cellulosic sugars are not the only alternative to starch-based carbon. Methane and methane derivatives are also potentially attractive feedstocks for bioprocessing. Abundant shale gas has dramatically increased the supply and reduced the price of methane.

Multiple Generations of Feedstocks

Grain-Derived Sugars

As mentioned above, the first-generation source of carbon for fermentation has been starch derived from grain. The U.S. ethanol industry has been built on grain feedstocks, and all current biological production of chemicals relies on grain. In the United States today, nearly 40 percent of the corn crop is consumed in nonfood or feed uses, primarily for the production of fuel ethanol.[97] Although this feedstock has served the industry well, there are limitations to the supply of grain and concerns about competition with the food and feed uses for grain. These concerns were anticipated in the Renewable Fuels Standard created by the 2007 Energy Independence and Security Act, which mandated dramatic expansion of the use of cellulose as a feedstock for fuel ethanol. The Renewable Fuels Standard mandated that, beyond 2010, most of the increases in fuel ethanol would derive from cellulosic sugar sources (Figure 4-1).

Ultimately, it is the land on which the grain is grown that is scarce. The supply of high-quality, arable land is finite, in the United States and globally. Yields per unit area will continue to increase, through improvements in agronomic practices, breeding of higher-yielding cultivars, and

Copyright © National Academy of Sciences. All rights reserved.

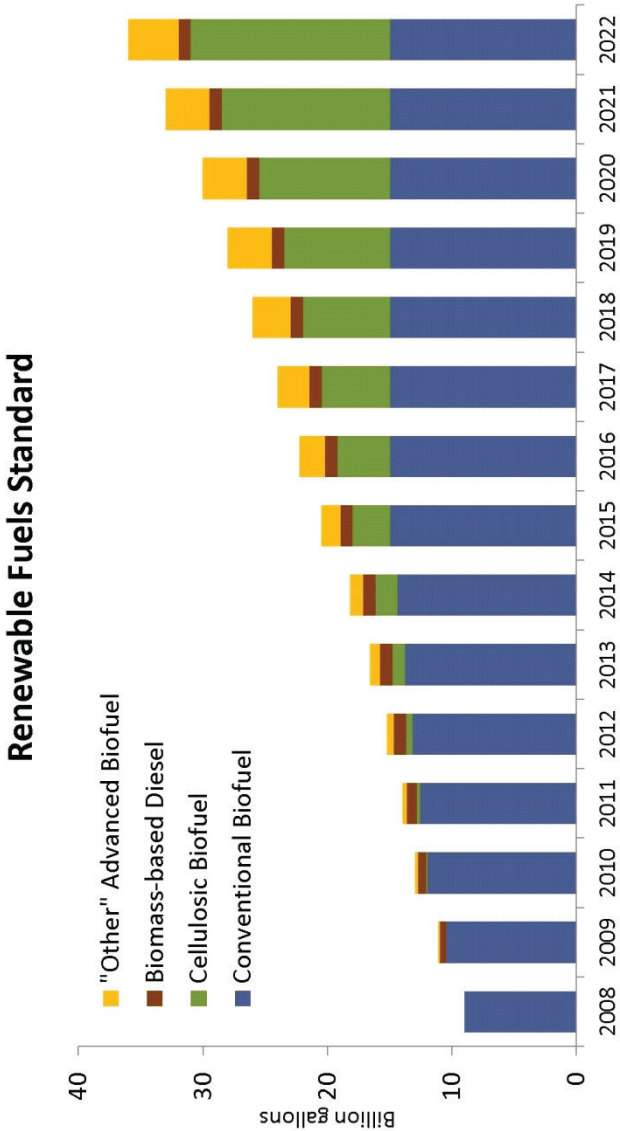


FIGURE 4-1 The quantity and sources of biofuels mandated by the 2007 Renewable Fuel Standard. SOURCE: EISA 2007 (RFS2).

Copyright © National Academy of Sciences. All rights reserved.

the application of agricultural biotechnology. Yields can be expected to improve by 1 to 2 percent per year in the developed world, and by somewhat higher rates in the developing world, where the yield baseline is lower. The projected rates of yield improvement will ensure an adequate supply of our food and feed needs. Alternative sources of carbon are needed to realize our full ambitions for the biological production of chemicals.

Lignocellulosic Biomass

Agricultural residues will be the first source of cellulose used in the biological production of chemicals. A current generation of cellulosic ethanol plants will rely on corn stover (stalks, leaves, and cobs) as the exclusive source of carbon. Other sources of cellulose are available from agriculture; wheat straw, rice straw, and sugarcane bagasse are all examples.

The use of cellulosic biomass as a source of sugars for fermentation requires a multistep process to digest the cellulose. The first step of the processing is size reduction to facilitate the flow of materials and to increase surface area for subsequent chemical steps. The second step is exposure to acid or base at elevated temperature, to break down the cellulosic structure and to unlock it from the lignin. The third and final step is saccharification, usually achieved using a cocktail of cellulases and hemicellulases, to hydrolyze the polysaccharides, yielding a mixture of simple sugars for fermentation.

The resulting sugar stream is a mixture of five-carbon and six-carbon sugars. Concentrations of sugars are much more dilute than that used in today's dextrose-based fermentation. The stream also contains recalcitrant polysaccharides, lignin, and other solids. For production of fuels or large-volume chemicals, economics require that the sugar stream be used without refining, separations, or concentration. Furthermore, economics dictates that the host fermentation organism be engineered to consume both five- and six-carbon sugars. Finally, the cost of cellulases and hemicellulases used for saccharification is a significant element. Enzyme efficiency must increase, and overall cost contribution from enzymes must decrease, as essential elements of cost reduction for sugars from cellulosic biomass.

Enzymes are eliminated entirely in an alternative process, based on supercritical CO_2 that is being developed yielding a cleaner, more concentrated sugar stream, with a somewhat higher associated cost.[98]

Lignin constitutes about 20 percent of corn stover mass. It is currently recovered and valued as a fuel. As the use of cellulosic feedstocks expands, strategies are needed to derive additional value from lignin so that it can be used as a co-product rather than a waste stream from fermentation.

Copyright © National Academy of Sciences. All rights reserved.

Beyond residues from agriculture, the forestry industry also produces residues that are a potential source of sugars for fermentation. The "hard" cellulose that constitutes wood is characterized by its higher hemicellulose and lignin contents (lignin approaching 40 percent by mass for some wood).

Forestry residues exist in large quantities, and they are often readily available at the saw mill for further processing. The disadvantages of woody biomass derive from its high lignin content. The lignin requires, and complicates, the intensive mechanical size-reduction operation. Because of these difficulties, woody biomass is considered recalcitrant in its release of sugars suitable for fermentation. New technology that would improve the release of sugars from wood could have significant economic value.

Dedicated energy crops will also play a significant future role as a source of carbon. Cropping patterns change slowly. It is unlikely that land used for today's crops will be converted to production of an energy crop. This said, annual crops such as sorghum have great potential to be a future source of cellulosic feedstock. Sorghum is well adapted to the more arid conditions of the U.S. western Great Plains. It is a versatile crop, for which the agronomic systems are well established. Sorghum has been bred to develop varietals with high yield of grain, or for cane sugar content, or to maximize biomass yield.

While the timeline for deregulation of biotech agricultural traits extends well over a decade, on that longer time horizon, additional technology can be brought to bear to produce sugars from cellulose. Use of advanced breeding techniques and transgenic traits can lead to cultivars designed for biomass disassembly into its constituent sugars. Modifications to the level of lignin and the nature of the hemicellulose content will lead to less recalcitrant biomass, yielding more fermentable sugars per ton, and further reducing the cost of usable cellulosic sugars.

Perennial grasses may also be adapted to cultivation on marginal land not currently used for row crops. As such, they have the potential to augment biomass supply without competing for today's agricultural land. Switchgrass, for example, is native to the United States and can yield large quantities of biomass per unit area. Grasses also offer greater flexibility in the timing of harvest, when compared to row crop residues. The principal disadvantage of perennial grasses is the 2 to 3 years needed to establish the crop. This constitutes a significant economic penalty at startup.

Fast-growing trees also hold potential as sources of fermentable sugars, in much the same way that they provide feedstocks for pulping processes. This source faces the twin challenges of the long time needed to establish the crop, and the challenge associated with the high lignin content of "hard" cellulose sources.

Copyright © National Academy of Sciences. All rights reserved.

C1 Feedstocks

Cheap, abundant natural gas (whose composition is essentially methane with trace amounts of other hydrocarbons) from unconventional sources is revolutionizing the U.S. energy and feedstocks landscape. Natural gas is replacing the products of naphtha crackers as the preferred feedstock for many chemical products. In addition to unconventional gas, there are also biological sources of methane from landfill gas or the biological digestion of biomass. Methane and its derivatives such as methanol, syngas, or formate all have potential as carbon sources for fermentation.

Despite the potentially attractive costs of C1 feedstocks, considerable technical challenges exist. Two-phase gas-liquid fermentation reactors are complex and costly. Both methane and hydrogen are sparingly soluble in aqueous media. Gas-liquid mass transfer is a significant impediment to high volumetric productivity in the fermenter. However, at least three demonstration-scale syngas-to-ethanol facilities are operating today. Additional process engineering and host organism research are needed to expand the economic viability of C1 feedstocks for the biological production of chemicals. Additional advances in C1s are coming from ICI, INEOS Bio, LanzaTech, and Newlight Technologies.

Key Conclusions

Conclusion: Improvements in availability of economically feasible and environmentally sustainable feedstocks are necessary to accelerate the production of fuels and high-volume chemicals via bioprocessing.

Conclusion: Improvements in the availability, reliability, and sustainability of biofeedstocks including

- cellulosic feedstocks from plants, including plants engineered for disassembly with special attention to low-cost saccharification;
- full use of lignin co-product from feedstocks;
- utilization of dilute sugar streams;
- ability to convert complex feedstocks into clean, fungible, usable intermediates via biological pathways;
- dramatic lowering of environmental impact;
- utilization of methane, methane derivatives, carbon dioxide, and formate as feedstocks; and
- use of noncarbon feedstocks (e.g., metals, silicon)

would increase the range of economically viable products, provide more predictive levels and quality of feedstock, and lower barriers to entry into the biological production of chemicals.

Copyright © National Academy of Sciences. All rights reserved.

Conclusion: Improving the basic understanding of C1-based fermentation, including both host organism and fermentation processes, is enabling in light of the increased availability of natural gas in the United States.

Roadmap Goals

- Within 4 years, for biological processing, achieve widespread use of novel sources of carbon, such as fermentable sugars derived from soft cellulose at a full cost less than $0.50 per kilogram of substrate.
- Within 7 years, for biological processing, achieve widespread use of novel sources of carbon, such as fermentable sugars derived from soft and hard cellulose at a full cost less than $0.40 per kilogram.
- Within 10 years, for biological processing, achieve widespread use of diverse sources of carbon, such as lignin, syngas, methane, methanol, formate, and CO_2, in addition to fermentable sugars derived from soft and hard cellulose at a full cost less than $0.30 per kilogram.

ENABLING TRANSFORMATIONS

Fermentation and Processing

Economic challenges have slowed the industrialization of biology. To accelerate the use of industrial biology for the production of chemicals, overall economics must be improved.

The product targets for industrial biotechnology must be selected using economics as a primary factor. It is difficult for a bioprocess to compete directly with large-volume chemicals produced from common petrochemical feedstocks in fully depreciated assets. High-valued specialties that take advantage of the high specificity of biology are advantaged. In the case of molecules that cannot be practically produced using conventional chemistry, an economically feasible bioprocess has no competition. For chemicals with a value of less than $20/kg, the market size must justify production of more than 1 kt/year.

Based on committee background and interactions with industry experts, for commodity chemicals, having a value of $2-5/kg, the potential market must be as large as 50 kt/year. For such products, both feedstock costs and capital costs are critical considerations. Hence, both product costs and capital costs must be reduced for industrial biology to compete effectively with conventional petrochemical processing. Moreover,

Copyright © National Academy of Sciences. All rights reserved.

it is recognized that bioprocesses should be viewed as complementary to thermochemical processes, rather than competing with them. In the future, many chemicals will be produced by a combination of biological and conventional chemical synthetic steps.

The host organism is generally viewed as the most important determinant of the economics of a biological production process. The biocatalyst determines three important economic parameters: the production rate, product titer, and yield from feedstocks. These factors greatly influence both the product costs and the plant capital expenses required. High-productivity, high-efficiency bioprocessing is needed to accelerate the industrialization of biology to produce chemical products. A step-change improvement in space-time yields for bioprocessing is essential to achieving needed reductions in product and capital costs. A typical fermentation reactor will produce 3-5 g/L-hr of product. This is at least one order of magnitude lower than that achieved in a typical chemical reactor. Such improvement can only come from more productive host organisms, combined with improvements in process engineering.

A bioprocessing facility for chemical production consists of a series of operations. Fermentation assets represent the largest capital expense in bioprocessing, but there are several essential operations. Feedstock pretreatment may be needed if the feedstock is anything other than a clean sucrose or glucose stream. Pretreatment is discussed in the preceding section on feedstock. Feedstock pretreatment operations may be integrated with fermentation or performed remotely. Pretreatment is followed by fermentation. Fermentation normally includes the use of seed fermenters to grow the biocatalyst cell population before its introduction into the production-scale fermenters. Following fermentation, separation is needed to remove the product from the cells and fermentation broth. This is accomplished through a variety of filtration or centrifugation steps. Finally, concentration and purification of the product is achieved using ultrafiltration, extraction, evaporation, distillation, ion exchange, and other processes. It is important to note that separation steps can be some of the most expensive steps in the manufacturing process and should be considered carefully.

Fermentation

Fermentation assets represent the largest capital expense in bioprocessing. Chemical production is normally done in an aerobic fermenter, equipped with cooling coils to maintain temperature and with agitation for both mixing and to facilitate gas-liquid mass transfer of oxygen and heat transfer for cooling. A typical aerobic fermentation plant for production of a specialty chemical typically costs \$200,000/m^3 and produces

Copyright © National Academy of Sciences. All rights reserved.

0.1-1 g/L-hr. A large-volume chemical produced via aerobic fermentation costs typically \$50,000-100,000/m^3 and produces 1-5 g/L-hr. In contrast, an anaerobic corn ethanol plant, operating at a vastly larger scale, costs typically \$7,500/m^3 (including dry-mill saccharification) and produces 3-5 g/L-hr.

Fermentation has been conducted in batch mode for a long time. Batch fermentation gave way to "fed-batch" reactors in which the carbon source and co-factors needed to grow the biocatalyst, maintain its metabolism, and deliver the product, which was done on a continuous basis.

Improvements have been made in fermenter performance through better agitation, heat transfer surfaces, and better gas-liquid contacting. Better heat and mass transfer have led to larger fermenters that can operate efficiently. Space-time yield remains low because of constraints of the microorganism and temperature and shear limitations.

Historically, host organisms have been selected and engineered to optimize productivity in terms of production rate, fermentation titer, and product yield (per unit feedstock). Additional characteristics of host organisms are needed that are developed in tandem with the overall process development. For example, the need to maintain a sterile fermenter environment contributes significantly to energy costs in the form of steam needed for sterilization. Organisms capable of operating in a less sterile environment, or having tolerance to allow for pH-based versus steam sterilization, would reduce product costs. Host organisms that exhibit greater temperature or shear resistance, or that require less oxygen, would contribute to improved space-time yield. Hosts exhibiting better strain stability can be adapted to continuous fermentations and longer, more productive batch fermentation.

Little attention has been paid to the continuous removal of the product. In typical batch fermentations, the end point is determined by the loss of productivity of the production host, which, in turn, is caused by the deleterious effects of accumulated metabolites, including the targeted product. Continuous removal of metabolites can reduce the costs associated with growing the host cells—both the costs of the carbon substrate, and the less productive fermenter hours, during the cell growth phase of the batch.

The chemical process industry evolved from batch reactors to continuous processes. The reasons for this were improved uniformity (elimination of batch-to-batch variation) and enhanced process control. The two are related, but the ultimate driver has been economics. It is hard to imagine the petrochemical process industry, operating at its enormous scale, without highly efficient continuous processes. Industrial biotechnology, true to its origins in brewing, has clung to batch and fed-batch fermentation processes. The development of continuous fermentation is important

Copyright © National Academy of Sciences. All rights reserved.

to improving the economics of industrial biology. This must be done in tandem with the development of host organisms built to this purpose.

The ability to build predictive models at the level of individual metabolic pathways, at the level of whole-cell metabolism, and at the level of the overall fermenter operation is a significant need. Available modeling tools for fermenters are helpful in constructing mass and energy balances, and flowsheeting of fermentation processes. Dynamic modeling tools that predict the effects of perturbation at the cell or fermenter level are a gap. These tools would be useful for development of batch, fed-batch, and especially continuous fermentation.

Scaling

Improvements in the host organism are essential to high-productivity, high-efficiency fermentation processing. While the host may be the most important determinant of the economics of a bioprocess, improvements in the engineering of the bioprocess are also a significant factor, with clear impact on both capital costs and operating costs. It must be recognized that the development of the host organism and of the bioprocess must be done in concert.

Process scale-up represents a key challenge and a potential hurdle to production of chemicals and fuels. The challenge of translating the host organism performance across scales, starting from microtiter, to small-scale fermenters, and eventually to production-scale fermentation is a significant one. Getting this right can assist rapid progress of the field. As the promise of synthetic biology starts to deliver, and the design-build-test-learn cycle (described below) begins to churn, high-throughput screens are needed to select the variants to be used in higher-scale testing. These decisions can be helped by assay protocols that can mimic at the microscale the performance of the strain during large-scale fermentation. For a specialty product, fermentation may occur at the 1,000-L scale, whereas a large-volume chemical could be produced in a fermenter of >100,000 L. Bioethanol is typically produced in fermenters of 1 million liters, or larger. The goal is to scale from the microtiter to the production scale, as quickly as possible, with the fewest number of intermediate scales of testing and rework. This challenge requires an interdisciplinary effort that includes chemical engineering, cell physiology, automation, statistics, and modeling.

Enzyme-Mediated Reactions

The use of enzymes in the production of biochemicals, or organic fine chemicals, has been practiced commercially for many decades. Early

Copyright © National Academy of Sciences. All rights reserved.

embodiments made use of naturally occurring enzymes, isolated from living organisms. As recombinant technologies developed, from the 1970s more efficient enzymes were developed that improved the process economics of enzyme-mediated reactions and broadened the base of applications. Enzyme catalysts are produced via fermentation, via the process described above. Enzymatic catalysis is typically used to effect reactions such as hydrolysis, aminolysis, amidation, or resolution of racemic mixtures. Typical commercial uses include a broad range of alcohols, amines, amino acids, and organic acids.

Enzyme-mediated reactions can be carried out at high yield. The stereo- and regioselectivity of enzymes heightens their utility. Increasingly, these reactions can be conducted in organic solvents, further broadening the use of enzymes. While enzyme-mediated reactions are often performed via homogenous catalysis, the development of stable, engineered enzymes has increased the ability to immobilize enzymes on a variety of substrates.[99]

Cell-Free Processing

The potential to conduct complex, multistep biocatalysis outside the cell offers tremendous promise. Cell-free processing is just this: the activation of complex biological processes without the use of living cells.[100] In practice, cell extracts have been used for many years to conduct simple reactions, along the lines of the enzyme-mediated reactions described in the section above. Cell-free processing utilizes the biochemicals of the cell, without the disadvantages of the cell's metabolism. The biocatalyst organism is grown via fermentation. The cells are then lysed, destroying the cells but allowing the biochemistry of the enzymes and co-factors to persist. The advantages of cell-free processing include the ability to add, or to remove, catalysts and/or reagents, and reduced effects of toxicity, because cell viability is not a concern. Energy and mass transfer may be enhanced by the absence of cell walls. The reaction medium is homogeneous, facilitating measurements of concentrations without concern about gradients across the cell wall. Cell-free bioprocessing is not without its challenges. Metabolic networks that are essential to the desired synthesis must be maintained. Co-factors must be recycled, to make the processes economical. To date, production rates remain modest. Complex, multistep syntheses have not been achieved. Operations at a scale suitable for large-volume chemicals are still to be demonstrated, but, given the great potential and numerous advantages of this technology, its development is likely to continue.[101]

Copyright © National Academy of Sciences. All rights reserved.

Additional Bioprocessing Operations

A number of unit operations are required downstream of the fermenter, enzyme-catalyzed reactor, or cell-free bioreactor. As in any chemical production, product separation and purification are necessary steps. These steps add operating and yield costs and represent a significant capital cost for the facility. Thermal separation processes are both energy and water intensive. Greater efficiency is needed to reduce the capital and operating costs of thermal separations. Use of alternative separations technologies such as extraction and membranes should be expanded. Lower-cost, cleanable membranes can reduce the costs of microfiltration and ultrafiltration. Separation processes adapted for continuous removal of product and other metabolites from batch fermenters is an additional need.

Fermentation processing requires the use of water. Water is used both as the fermentation process medium and as steam and cooling water in product recovery. The amount of water required per gallon of fuel ethanol has decreased from 5.8 gallons in 1998 to about 3 gallons today. Further improvements are needed in water reuse, with a goal of achieving near-zero net water usage.

While bioprocesses are often considered environmentally benign—"greener" than chemical plant operations—they do generate solid and liquid wastes. Dramatic increases in the use of bioprocessing will require disposal of larger quantities of these streams. Alternatives to current disposal methods will be required. Waste streams must be recognized for the additional value they can present. Co-product value will need to be derived from waste streams to improve the environmental footprint and to improve the economics of bioprocesses.

Key Conclusions

Conclusion: Aerobic, fed-batch, monoculture fermentation has been the dominant process for bioproduction of chemicals for many decades. Successful improvement efforts have focused on more productive host organisms. Little research has been conducted to improve the productivity of the fermentation process, by means of enhanced mass and heat transfer, continuous product removal, and more extensive use of co-cultures, co-products, and co-substrates.

Conclusion: The development of predictive computational tools based on small-scale experimental models that realistically predict performance at scale would accelerate the development of new products and processes for the production of chemicals via industrial biotechnology.

Copyright © National Academy of Sciences. All rights reserved.

Conclusion: Unlike many traditional chemical processes, industrial biotechnology generates large aqueous process streams that require efficient mechanisms for product isolation and for efficient water reuse.

Roadmap Goals

- Within 3 years, achieve an operating process for an economically viable bioreactor that overcomes mass-transfer and separations limitations associated with gaseous feedstocks and/or gaseous products.
- Within 5 years, develop data-based modeling tools and scale-up technologies that enable reliable scale-up of any bioproduction process from 10 L to 10,000 L in less than 6 weeks.
- Within 7 years, consistently and reliably achieve fermenter productivity of 10 g/L-hr at steady state in a continuous fermenter, or following the growth phase in a batch fermentation.

- Within 5 years, for all bio-based aqueous processes, achieve 80 percent reuse of process water.
- Within 7 years, for all bio-based aqueous processes, achieve 90 percent reuse of process water.
- Within 10 years, for all bio-based aqueous processes, achieve 95 percent reuse of process water.

ORGANISM

The core of an expanded industry emerging from the accelerated biological production of chemicals will consist of specialized organisms capable of producing a given compound at titers, productivities, and yields sufficient for economical production. These microbes will almost certainly be highly engineered, featuring many genetic modifications, including but not limited to insertion of genes encoding new enzymatic activities, deletion of genes encoding competing and undesired activities, and modification of genes to alter regulatory and feedstock, intermediate, and product tolerance processes. Hence, the core of this industry will consist not only of the microbes themselves but also of advanced methods for the facile production of these engineered organisms. The advances necessary to generate these next-generation production strains fall into several categories: first, the development of modeling and design tools capable of the predictive tailoring of pathways, genomes, and capabilities of industrial microorganisms, from discovery to large-scale fermentation; second, the underlying science and technology for genome manipula-

Copyright © National Academy of Sciences. All rights reserved.

tion, including in organisms that are not part of the current pantheon of established production strains or that may yet be discovered in the wild; third, informative measurement techniques to assess the performance of engineered organisms and pathways; and finally, approaches to learn from previous efforts so as to repeat successes and avoid past failures.

The development of an engineered organism for the production of chemicals begins with a technical specification for the desired bioprocess, with particular emphasis on those aspects of the specification that influence the selection of host organism and metabolic pathway. The initial specification may include one or more of the following: (1) the chemical(s) to be produced, (2) the target price point of the finished chemical (e.g., dollars per kilogram), (3) the target volume of the chemical (e.g., metric tons per year), and (4) the target feedstock (e.g., glucose). These are most relevant to the host and pathway, as they establish the primary set of parameters and objectives around which strain engineering will commence. As proof of concept is established for biological production and the model needs expand to consider the full integration of process design and development, additional specifications may include (5) the quality specification of the finished product (e.g., purity); (6) the target titer, productivity, and yield (as determined from a technoeconomic model of the full bioprocess); (7) additional bioprocess considerations (e.g., batch versus fed batch versus continuous fermentation, use of co-solvent, aeration level) that may influence the design of the organism; and (8) designs that expedite scale-up and ongoing quality control measurements.

Engineering organisms for the production of chemicals thus requires modeling across many different levels of resolution, spanning from (re)design of host metabolism to support the carbon, energy, and cofactor needs of chemical production to design of the genetic sequences that encode the cellular machinery needed for manufacturing the chemical to the desired specification. Each of these levels presents its own set of technical challenges, needs, and opportunities. Further, biomanufacturing of any chemical compound will certainly require extensive strain engineering if the molecule is heterologous; however, even for hosts in which the molecule is a naturally occurring metabolite it is highly likely that additional modifications will be necessary to achieve a commercially viable process.

If the target molecule is not a known biological metabolite but its synthesis is believed to be accessible through biology, then a novel pathway must be designed to produce the product of interest from either an existing metabolic intermediate or a readily supplied carbon source. Once a pathway has been specified, the next step is to select the enzymes needed to catalyze each biosynthetic step. The mining, design, and evolution of discrete steps will lead to a functioning pathway, but typically

Copyright © National Academy of Sciences. All rights reserved.

with low yield. After a functioning pathway has been established, the development process must continue in order to produce the ultimate, engineered microbe generating the desired product at specified rates, titers, and yields. As multilayered as the yield may be, it can ultimately be specified in terms of mass and energy balances, which will have concomitant impacts on the organism as a whole that must be taken into account.

Introduction: The Design-Build-Test-Learn Loop

An essential element of engineering biology is the application of the time-honored, iterative scheme of design-build-test-learn (DBTL) that is a hallmark of all engineering disciplines. Metabolic engineering first applied engineering principles toward strain construction for production of small molecules. Synthetic biology has endeavored to expand and greatly enhance the DBTL loop throughout all aspects of the engineering of biological systems. For a given desired bioprocess, this DBTL loop spans from the selection and tailoring of a suitable host and metabolic pathway, the enzymes that will constitute the pathway, the genetic systems that will express the enzymes, and the implementation plans for how to build and test what has been selected and tailored (design); to employing DNA synthesis, assembly, transformation, and genomic modification tools to generate the designed strain variants (build); to culturing these variants to assess the performance of the built strains, for example through approaches such as transcriptomics, proteomics, metabolomics, and one or another means of metabolic flux analysis[102] (test); and finally to evaluating the resulting test data to determine whether the design was successfully realized and whether the initial design model(s) or build and test processes require further improvement (learn). Each aspect of this cycle will be considered in turn as it applies to the overall foundational science and driving conclusions that support the acceleration of biomanufacturing.

Fully Integrated Design Toolchain

Across each of the levels of resolution described at the outset, we note a common gap between the scientific design tools available today and the engineering design tools needed to achieve the envisioned future presented in this report. To date, most tools used in organism design are what are colloquially referred to as "pull" tools. Pull tools are tools that enable the user to ask and answer a specific question regarding a proposed design. For example, mFold allows a user to submit a nucleic acid sequence for secondary structure prediction. PROSITE allows a user to submit a protein sequence for known protein domain motifs.[103] COBRA

Copyright © National Academy of Sciences. All rights reserved.

allows users to use a genome-scale model to predict cellular metabolism under different conditions among other functionalities.[104] Although each of these tools can be useful in the overall organism design process, they all require that the biological engineer formulate a specific question regarding a proposed design, identify and apply the tool that can answer that question, and then interpret the validity of the results. This approach limits the detection of flaws in a proposed design to those issues that the biological engineer opts to study—the engineer must "pull" information from each available tool. To realize the grand challenges presented in this report, it will increasingly be necessary to develop and deploy "push" tools that can preemptively provide useful information regarding potential flaws in proposed designs. For example, a comprehensive push tool for designed genetic sequences might scan the input nucleic acid sequence for gene expression regulatory motifs (promoters, transcription factor binding sites, ribosome binding sites/Kozak sequences, codon usage, translational pause sites, terminators, and RNase sites), structural motifs (DNA, RNA, and protein secondary and tertiary structure), as well as functional motifs at the protein level (known protein domains, signal sequences, and proteolytic cleavage sites) and "push" the summarized results of this analysis to the biological engineer. More sophisticated push tools may even be able to prioritize the results based on both estimated confidence in each prediction as well as the likelihood that each result might adversely impact organism performance. Push tools free the biological engineer from needing to query each design against a library of tools and instead rely on software to point out all potential issues in a proposed design. It should be clarified that push tools extend beyond merely more autonomous and integrated software systems. Push tools have the potential to notify the biological engineer, asynchronously with the engineer pulling information from an integrated system, of new concerns or opportunities as additional information or tool improvements emerge. For example, if a desired biosynthetic route is currently inaccessible because no known enzyme exists to perform a key step in the pathway, then a push tool could notify the engineer when such an enzyme is identified. Or, if new information indicates that a metabolic intermediate of a previously designed pathway poses a significant human health risk, a push tool notification could arrest the deployment of the designed biological system implementing that potentially hazardous pathway.

The realization of a fully integrated design toolchain will require the establishment of standardized software tool application programming interfaces (APIs) so that the tools can effectively send "push" notifications to each other, apply data-exchange standards that specify how the content of the notifications should be structured, and make use of standardized data repositories (relating in particular to bioprocesses, bioreactors,

Copyright © National Academy of Sciences. All rights reserved.

Industrialization of Biology: A Roadmap to Accelerate the Advanced Manufacturing of Chemicals

organisms, pathways, enzymes, expression systems, and build and test methodologies) that design tools can pull information from. Standardization is a well-established concept and practice in synthetic biology, dating back to at least the development of the BioBrick DNA assembly.[105] More recent efforts have sought to move standardization beyond physical DNA assembly to data-exchange and visual design representation standards, including the Synthetic Biology Open Language (SBOL) and its visual notation (SBOL Visual).[42] Complementary efforts have sought to leverage and adapt other established standards, such as Digital Imaging and Communications in Medicine (DICOM),[106] into the service of synthetic biology. Repositories of information concerning organisms, DNA sequences, and expression systems have begun to emerge, including the iGEM Registry of Standard Biological Parts,[107] the ICE repository platform,[108] the Virtual Parts Repository,[109] the DNASU plasmid repository,[110] and AddGene.[111] The first three of these specific repositories have established APIs for design tools to access their contents, and efforts are under way to develop a standardized API across these repositories to enable a united "Web of Registries." While these efforts demonstrate that some progress has been and is being made toward the establishment of the standardized APIs, data-exchange standards, and standardized data repositories that will be required to enable a fully integrated design toolchain, it is clear that much work remains (in particular around establishing repositories of experimental measurement and characterization data). It should also be noted that there is a delicate balance between the organic and prescriptive development of standards, namely that, although standards are essential to realizing the fully integrated design toolchain and new incentives (whether resource or social) are required for their development, there is a risk that prematurely institutionalizing a standard could create cumbersome legacy disincentives to make improvements that might adversely impact innovation and rates of progress. It is likely that if an integrated design toolchain is developed and becomes widely used, data standards for feeding this toolchain will follow naturally.

Key Conclusions

Conclusion: The development and use of a robust integrated design toolchain across all scales of the process—individual cells, cells inside reactor, and the fermentation reactor itself—is an important step in bringing biomanufacturing onto the same level as traditional chemical manufacturing.

Conclusion: The development of predictive modeling tools within and for integration across all scales of the process—individual cells,

Copyright © National Academy of Sciences. All rights reserved.

cells inside reactor, and the fermentation reactor itself—would accelerate the development of new products and processes for the production of chemicals via industrial biotechnology.

Roadmap Goals

- Within 4 years, develop and demonstrate an integrated design toolchain for the design of a biomanufacturing process at and below the level of an individual organism (i.e., everything inside the cell).
- Within 7 years, develop and demonstrate an integrated design toolchain for designing a biomanufacturing process at and below the level of an individual biological reactor (i.e., everything inside the reactor).
- Within 8 years, develop and demonstrate an integrated design toolchain for designing an entire biomanufacturing process (i.e., everything from concept to product).

Design

Pathway Design

The first step in the design process is to select an appropriate metabolic pathway for biosynthesis. In this case, even the knowledge of an elucidated pathway for target synthesis does not necessarily render the choice of metabolic pathway obvious. For example, the isoprenoid/terpenoid family of compounds can be produced using the mevalonate or nonmevalonate (DXP) pathway, or a hybrid incorporating elements of both.[112] Similarly, succinic acid can be generated from either the oxidative or nonoxidative branch of the tricarboxylic acid cycle, or a hybrid of both.[113]

For more novel conversion steps, in which the enzymatic chemistry is validated but transformation of the specific substrate of interest has not been experimentally confirmed, new tools are needed to increase the predictability of proposed designs. Such tools would ideally provide a rank order for pathway designs based on predicted experimental feasibility. Factors to consider may include the chemical distance between known and target substrates,[114] diversity of enzymatic sequences encoding the activity of interest, knowledge and understanding of reaction mechanism of target enzyme activities (to aid rational design of enzymes; see below), and extent of functional validation of substrate diversity and range.

Copyright © National Academy of Sciences. All rights reserved.

Enzyme Design

For known enzymatic reactions, the design tool should include available experimental data to rapidly identify variants likely to possess the highest activity. It should be noted that in biology, context matters and thus typical experimental data, such as measurements of activity in idealized *in vitro* conditions, may not translate into high activity in the cellular host. Nonetheless, the integration of detailed biochemical information, where available, can aid the selection process. When enzymes are not identified that meet the target specifications, alternatives must be found. One option is the search for alternatives based on homology to known variants, for example, using BLAST alignments.[115] This search method does not require isolated or functionally validated sequences but relies solely on similarity to suggest additional options. The advantage of this approach is that it facilitates access to the treasure trove of genomic and metagenomic data to access new variants; however, the disadvantage lies in the uncertainty associated with sequences that have not been functionally validated. Even as improvements in build throughput emerge, it is still desirable to avoid unnecessary synthesis of enzyme-encoding DNA sequences that fail to be functionally useful. Accelerated industrialization demands increased predictability to link protein sequence to enzymatic function. Design tools that improve the accuracy of functional prediction—and, ultimately, the ability to predict not just whether an enzyme will be active but how active it will be—can greatly accelerate the initial steps of establishing proof of concept for biosynthesis. The integration of pathway design and enzyme specification tools, resulting in exquisite computational tools that can reliably present a feasible *de novo* pathway toward a target compound, would herald a revolution in industrial biology as these tools would immediately and dramatically expand the scope of chemical compounds that would be candidates for biomanufacturing.

As indicated above, supporting and building on existing enzyme databases will accelerate efforts in enzyme design. As these databases are built out, data fields should be modified to include knowledge that is particularly relevant to pathway design, such as known side reactions, substrate specificities, allosteric controls, evolvability (based on phylogenetic or experimental knowledge), and potential functional analogues.

Systems Biology Design

Implanting pathways and enzymes in a chassis for screening or production usually requires a (re)design of host metabolism and/or physiology to achieve the desired performance standard. Metabolic design objectives typically include reengineering of competing by-products and minimization of biomass formation. Both of these contribute to maximiz-

Copyright © National Academy of Sciences. All rights reserved.

ing product yield. However, biosynthetic pathways often involve redox reactions such that electron flow has to be considered in combination with carbon flow. Additionally, specific transformations may require coupled reactions or the generation of activated substrates to provide the energy needed to catalyze thermodynamically unfavorable reactions. The fully integrated design toolchain should be able to satisfy these layered objectives, accounting for endogenous metabolism, heterologous product formation, and redox and energy balances to predict the optimal combination of genetic manipulations. To this end, it would be desirable to also have registries containing the characteristics of hundreds of host organisms and their phenotypes under a wide range of conditions, such as different temperatures, pressures, salinities, and carbon sources. Such a registry would be a public good in the same manner as PubMed and could be accessed to accelerate both corporate efforts and to provide fodder for the fuller development of systems biology tool sets for organismal and pathway design.

It will be especially important not to neglect systems-level effects on overall cellular physiology. A commonly viewed obstacle toward bio-based small-molecule production is toxicity, in which the product greatly or completely reduces cell viability and, in doing so, affects the production capacity of the host. These effects are often not easily classified in mass and energy balance equations and often manifest in both physical and biological ways. For example, a product may be inhibitory to enzymes in the pathway or to other endogenous reactions essential for cell performance. In this case, identification and incorporation of feedback-resistant enzymes may alleviate the most harmful effects. While certainly not trivial to implement, this form of toxicity has a clearly assigned biological cause and can be addressed as such. On the other hand, if the product associates physically with the cell membrane, disrupting membrane integrity and causing leakage of cytoplasmic contents, then this mode of toxicity must be understood on a more fundamental level to rationally propose a solution, perhaps through engineering the composition of the cell wall to withstand higher concentrations of the toxic production. In either case—or in combinations thereof—design tools are needed that can propose both a mechanism of toxicity and a means to address it, given knowledge of the system. It should be noted that adaptation and evolution could certainly be used to obtain strain variants with more tolerant phenotypes and, in this case, the design tool chain should be able to incorporate findings from these experiments to learn and thus implement that knowledge in future design scenarios.

Copyright © National Academy of Sciences. All rights reserved.

Bioprocess Design

While the design toolchain as described above is focused on the cellular organism, a fully integrated design process must operate across scales to incorporate bioprocess considerations. The strain that performs to specification will operate reliably as designed. These performance specifications can then be translated into well-established parameters, including, for example, observable product yield on substrate, product yield on biomass, and specific productivity, that have been successfully used for decades to model and design bioprocesses. As cellular behavior is more complex, for example, exhibiting dynamic behaviors through the incorporation of feedback control mechanisms, these behaviors can be modeled at the bioreactor scale to predict overall process performance, ultimately generating the predictions in titer, yield, and productivity that are necessary to evaluate the commercial viability of a process. Overall, models should predict cell behavior in culture over a wide variety of culture volumes and under a wide variety of bioprocess conditions. The systems biology–based registries imagined above would assist in building tools to eventually predict scale-up and scale-out.

Build

The construction of new organisms for industrial biology applications can be further broken down into the identification, characterization, and modification of "chassis" for production, and the construction of appropriate pathways in these chassis for the production of a given compound.

The modification of chassis and the construction of new pathways will be greatly enabled by the ongoing revolution in DNA synthesis. To the extent that we remain on an exponential trajectory for the acquisition of longer and cheaper pieces of DNA, much larger and many more constructs can be generated and tested. Synthesis technologies will make the DBTL paradigm particularly powerful. That said, there is clearly a growing need for biofoundries that can scale the production of subgenomic assemblies or pathways. While public funding may lead to the establishment of more centers for synthesis, it is also possible that synthesis and assembly technologies can be developed to the point where DNA designs could be synthesized and assembled by bench-scale equipment (a "DNA printer") in essentially every research lab.

Pathways

Pathways are typically composed of a series of enzymatic transformations, integrated with central metabolism via sensors and regulatory interactions. In order to develop pathways capable of generating virtually

Copyright © National Academy of Sciences. All rights reserved.

any small organic product of interest, it will be first and foremost necessary to enable the acquisition of enzymes that can carry out virtually any transformation. Such enzymes can likely come from three sources: first, mining phylogeny for novel enzymes; second, elaborating the catalytic activities and biophysical properties of known enzymes, by either design or selection; and finally, the generation of enzymes with wholly new properties not previously found in nature.

Bioinformatic approaches to mining new enzymes are already in vogue,[116] and the integrated design toolchain described above is likely to continue to both fill informatics databases with alternatives and better target enzymes to new purposes and pathways. Although mining and characterization have yielded numerous parts that have proven to be useful for microbial engineering, in many cases the specific roles of parts or their performance in new contexts must be further optimized. Two methods have shown promise in the generation of parts for virtually any genetic circuit: computational design and directed evolution. Such methods are equally useful for proteins. The computational design of proteins has advanced to the point that it is now possible to generate novel protein folds and to frequently improve the functionalities of extant proteins, including their stabilities and interfaces with both small molecules and biopolymers. There have been several enabling improvements in protein design tools, most notably the widespread use of the Rosetta suite. In concert with the improvements in DNA synthesis that have been noted elsewhere, this has meant that it is frequently possible to redesign a given protein scaffold for novel structure, synthesize tens to hundreds of predicted variants, and quickly assay for those that have the required capabilities. Roadblocks that remain to future progress primarily have to do with improvements in physics-based approaches and algorithms that will better specify the energetics of interactions, especially with small molecules. As these barriers are overcome, it should be possible to redesign enzyme active sites to accommodate a wide range of substrates and co-factors, and thereby to more completely enable the development of virtually any transformative pathway. A reach goal would be the ability to design enzymes *de novo* for chemical reactions that currently have no biocatalytic equivalent.

Similarly, the directed evolution methods described for organisms also apply to enzymes, and there are a variety of techniques for altering enzyme properties. Directed evolution complements molecular design in that it can sieve through large numbers of molecules for those few with the required capabilities. However, directed evolution is frequently capable of sieving very large libraries of millions to billions of variants, thus partially obviating the need for design. On the other hand, the sequence spaces that are accessible by even small proteins are so large

Copyright © National Academy of Sciences. All rights reserved.

that design tools have proven to be extremely valuable for delimiting the libraries that will be constructed for a given directed evolution experiment.

The key issue that restrains more widespread use of directed evolution as a means of optimizing parts is that novel selections or screens must be developed for each new molecular functionality. If an enzyme with new substrate specificities is desired, then either the enzyme's functionality must be linked to cell growth or a high-throughput assay specific for that enzyme must be devised. In order to overcome these problems researchers have begun to develop more generalized schemes for directed evolution, such as phage-assisted continuous evolution[117] and compartmentalized partnered replication, that attempt to generally connect the phenotype of a part with function in a system, thus enabling more modular selections. In this regard, improvements in rational design may enable smaller libraries of sequence space to produce desired activities with limited screening throughput.

Going beyond nature and beyond the capabilities of directed evolution is still mostly notional. It is possible that wholly new enzymes can be designed or selected that incorporate a variety of novel elements to carry out complex bioinorganic transformations. Similarly, the 20 amino acids available for enzyme chemistry can be greatly augmented by nonstandard amino acids that are better able to perform specific chemistries or that can "harden" proteins to the requirements of bioprocessing streams operating at high temperatures or under acidic conditions, intracellularly or in isolation.

This space is well populated (although not saturated) by industry. Between improvements in computational design and directed evolution, the prospect exists for taking a relatively small list of parts and endlessly morphing their function to suit the needs of industry. This in turn suggests that there will likely be productive niches within the corporate ecosystem devoted to parts improvement. Companies such as Codexis regularly develop novel enzymes for customers carrying out large-scale bioprocesses.[118] It is not unreasonable to expect that if "conceptual barriers" between design and synthesis remain in place and are propagated, part of a future system will specify the characteristics of a part, rather than the part itself, and if those characteristics are not satisfied by something already in a database, then the specifications will be delivered to a parts foundry as a standing order.

Key Conclusion

Conclusion: Improvements in the ability to rapidly design enzymes with respect to catalytic activity and specific activity and engineer

Copyright © National Academy of Sciences. All rights reserved.

their biophysical and catalytic properties would significantly reduce the costs associated with biomanufacturing and scale-up.

Roadmap Goals

- Within 7 years, have the ability to insert 1 megabase of wholly designed, synthetic DNA into the genome of an organism at an error rate of less than 1 in 100,000 base pairs, at a cost of $100, in 1 week.
- Within 7 years, have the ability to design *de novo* enzymes with new catalytic activities with a high turnover rate.

Chassis

In the bioprocessing considered here, cells are the unit of engineering. Although enzymes or pathways can be embedded in cells, the cellular metabolism and physiology that supports chemical transformations are often critical aspects of bioprocess engineering and scale-up. While a great deal of basic metabolic engineering can take place in *E. coli* and other model organisms, these cellular "chassis" may not always be suitable for production.

The diversity of metabolic and physiological requirements for the production of different compounds necessitates a range of cellular chassis for metabolic engineering. For example, microorganisms with a naturally high tolerance for long-chain alcohols may be more suitable as hosts for new biofuel production, while strains with very low pH tolerance are advantageous for production of organic acids by minimizing downstream separation costs. The reason that *E. coli*, *S. cerevisiae*, and other model organisms are so highly used is the extensive repertoire of genetic tools available for these hosts. As a result, the correlation between genomic, proteomic, metabolic, and other information is relatively complete and is already laid down into systems biology models that are increasingly being quantified (as apparent from the Design Toolchain described above). Therefore, it is critical that additional foundational research be carried out on the systems biology and physiology of organisms that are better suited to bioprocess engineering and production.[119] Beyond capturing the genome sequences of laboratory strains, sequencing greater numbers of microorganisms that are actually involved in production should prove useful. Ancillary proteomic and metabolic analyses, and follow-on quantitative and predictive models for these systems as a whole, will provide fodder for grafting new enzymes and pathways to these chassis and therefore for producing a new cornucopia of compounds at the industrial scale.

Copyright © National Academy of Sciences. All rights reserved.

As we garner better understanding of industrially relevant chassis, new tools for the manipulation of organismal genomes will become increasingly important. This is especially true because of the limitations on transformation and because the breadth of different chassis under consideration will require more generic mechanisms for undertaking site-specific genome modifications. In this regard, the ongoing innovations with CRISPR-derived systems promise to revolutionize the modification of many organisms, including those relevant for chemical production, either via targeted genomic editing or via regulation of individual pathways by catalytically inactive, programmable ribonucleoproteins such as dCas9.[120] There are other systems for programmable site-specific modification, including Targetrons,[121] TALENS,[122] and zinc-finger endonucleases;[123] modifications of all of these systems often allow the site-specific insertion or mutation of genes, as well as their deletion. Overall, continued advances in these areas promise to widen the reach of methods like MAGE,[124] in which there is iterative optimization of function across the entire organismal genome.

In contrast to these methods, many synthetic biologists have focused on developing orthogonal systems that can operate beside or on top of extant genomes. Such orthogonal systems may come to represent very large, programmable subsystems with their own replication, transcription, and translation capabilities, as well as internally programmed regulatory and metabolic pathways. In essence, episomes carrying these features would be subgenomes that would both direct their own function and redirect their host's genome toward a desired functionality, such as the production of a particular metabolite or compound. To promote the development of this new generation of programmable, self-sufficient episomes may require a renaissance in plasmid and epsiome biology. Indeed, this may be an area where synthetic biology can provide modules that go well beyond regulation or metabolism. Into the future, it should be possible to take a toolbox of standardized and orthogonal origins, polymerases, promoters, ribosomes, and encoded amino acid biosynthetic and charging capacities and create made-to-order episomes for any of a variety of industrially relevant bacteria. The addition of CRISPR or other elements would allow these subgenomic control systems to finely control host expression.

Following site-specific genome engineering or the introduction of subgenomic control systems on episomes, the stabilization of an engineered chassis would be paramount. Most organisms have evolved not to produce a metabolite or compound in great yields, but instead to grow and survive. Redirecting metabolic flux for human purposes is usually an evolutionary dead end. Thus, either the rate of mutation and genetic change must be greatly reduced, or the engineered organisms or episomes

Copyright © National Academy of Sciences. All rights reserved.

must be evolutionarily robust—able to retain function even in the presence of multiple mutations. For example, proteins may be engineered to tolerate multiple amino acid substitutions and would thereby exist on a large neutral fitness landscape that would greatly delay loss of function. When such proteins are expressed in a slow-evolving chassis that contains antimutator polymerases or enzymes that can remove nucleotide modifications even prior to incorporation, it may be possible to slow evolution to the point where it is no longer a consideration over the industrial lifetime of a biosynthesized product.

Paradoxically, before a chassis is fixed into an evolutionarily stable trajectory, directed evolution methods applicable to whole organisms will be of increasing importance. As systems biology approaches provide increasingly excellent "roadmaps" for metabolic and regulatory engineering in a wide variety of organisms, it should be possible to delimit what pathways, loci, or regulatory networks should be the focus of directed evolution. In the past, strain improvement via random chemical mutagenesis was one of the primary tools for generating a production strain. Now random or semirandom approaches to modifying organismal genomes, coupled with well-designed selections or the high-throughput screens described below, will allow organisms to be driven into more productive states. In particular, the sequence-directed approaches to manipulating organismal genomes described above will likely prove more useful not only for model-based manipulation but also for directed evolution. These include methods such as recombineering libraries (as embodied in MAGE) and Cas9/dCas9 libraries. Again, an issue with many of these approaches is that they are targeted largely to *E. coli* as a platform, and their use in nonstandard laboratory strains, especially those that may be of greatest importance for production, is limited. This will require the adaptation of these tools to new organisms, potentially via the development of broadly useful episomes for horizontal transfer, as described above. In this paradigm, the tools and libraries for site-specific or random modification might initially be created in a tractable chassis, such as *E. coli*, and then moved by horizontal transfer to a new host to execute.

Genetic designs are currently limited to approximately a dozen genes, whereas genomes consist of thousands and many of the potential products of biology will require large numbers of regulated genes. As such, as the desired products become more complex, so too will the need to push our design capacity to this scale. This will require pathway design involving dozens of genes that collectively build the desired product. This will have to be integrated into the broader cellular metabolism and cellular functions, for example, those involved in nutrient and feedstock acquisition (e.g., cellulases, nitrogen fixation), secretion and import of precursors, and stress response. These functions require precise timing as to the con-

Copyright © National Academy of Sciences. All rights reserved.

ditions or order in which they are expressed as part of building a product or coordinating responses. This will require the ability to build synthetic regulation of a sophistication of the natural regulatory networks in cells. All of these genes are going to tax the host's resources, which will require a better understanding of how to allocate cellular machinery. Collectively, these designs will require combining hundreds of DNA parts and being able to predict how they work in concert. All of these considerations will have to be integrated into future computer-aided design packages that facilitate the management of large genetic engineering projects. In essence, the domestication of an organism as a suitable chassis in industrial biotechnology, as *E. coli* is today.

Key Conclusions

Conclusion: Continued development of fundamental science and enabling technologies is required for the rapid and efficient development of organismal chassis and pathways.

Conclusion: Expanding the palette of domesticated microbial and cell-free platforms for biomanufacturing is critical to expanding the repertoire of feedstocks and chemicals accessible via bio-based manufacturing.

Conclusion: The design, creation, and cultivation of robust strains that remain genetically stable and retain performance stability over time in the presence of diverse feedstocks and products will reduce the costs involved in the use and scaling of biological production.

Roadmap Goals

- Within 2 years, achieve domestication (including >1 percent transformation competency, genetic and genomic modification tools) across five phenotypically diverse microbial types other than established models (such as *E. coli* and *S. cerevisiae*).
- Within 5 years, achieve domestication across an additional 10 or more industrially relevant recalcitrant microbial types and the ability to domesticate any other microbial type within 3 months.
- Within 7 years, develop the ability to achieve domestication in any new microbial type within 6 weeks.
- Within 7 years, have a suite of domesticated organisms (including cell-free systems) that can utilize diverse feedstocks and generate a range of products with high yield and productivity under various process conditions while maintaining process robustness.

Copyright © National Academy of Sciences. All rights reserved.

Test and Measurement

Although the ability to design and evolve parts and circuits is of fundamental importance for the improved practice of synthetic biology, developing improved methods for measuring the results of experiments will perhaps have an even greater impact. Design and evolution can provide basal circuitry that frequently requires additional optimization. Improvements in design tools can reduce the number of circuits that need to be tested and can improve the overall quality of those circuits, while facile directed evolution methods allow an ever larger number of variants to be screened and selected for improved function. But in neither case will the tools developed cover all challenges; they will likely continue to run well behind the sheer size of the sequence spaces being explored.

By enabling the underlying data needed to create and continuously improve the design methods envisioned above, measurement technologies will play a strong role in the subsequent emergence, practice, and advancement of engineering biology. The comprehensive measurement of DNA, RNA, proteins, metabolites, their chemical and structural variants, and their interactions enabled the advances in molecular cell biology knowledge and methods that have brought us to our current state of capabilities and understanding. The new knowledge and technological advances in turn motivate new questions and unmet needs for measurement that must be addressed to enable the efficient and effective future pursuit of engineering biology. New innovations in measurement would help accelerate the DBTL cycle, improve predictive design, broaden the scope of directed evolution, support manufacturing development and process control, propagate standards, improve regulatory decisions, and ensure safe practices.

Many advances in measurement technology are driven by medical applications. These same advances, with modification and extension, can also be useful for engineering organisms. Creatively extending such measurement technologies to the needs of engineering biology in an application-specific way will be valuable.

A preeminent example of a revolution in measurement methods primarily motivated by biomedical research that simultaneously enables leaps in engineering biology is nucleic acid sequencing. Now well integrated into multiple parts of the DBTL cycle, current practice would be inconceivable without it. Advanced methods for manufacturing DNA constructs, characterization of the structure and stability of transformed genomes, quantification of the impact of genomic alterations on expressed transcripts, clarification of the behavior of regulatory elements, and identification of the genomic alterations accompanying phenotypes of interest on the basis of nucleic acid sequencing are prevalent. There remain, however, opportunities to usefully further extend high-throughput sequencing

Copyright © National Academy of Sciences. All rights reserved.

Industrialization of Biology: A Roadmap to Accelerate the Advanced Manufacturing of Chemicals

by lowering its error rate to keep up with the very low and increasingly lower error rates of DNA synthesis; by improving its ability to delineate large-scale structural rearrangements as complex, large-scale, and precise genome engineering becomes more common; and by improving its sensitivity to single cells without compromising throughput to better discern the appearance and influence of heterogeneity among populations of cells. Improvements in error rate, read length, and sensitivity are also sought by biomedical researchers and clinicians concerned with complex diseases such as cancer. However, the requirements for throughput, data quality, data analysis methods, sample preparation, and integration of the results with complementary methods are quite distinct, leading to a divergence in the needed advances and their best implementation. These differences have already resulted in segmentation of platforms themselves across these very different fields. For example, some leading-edge single-molecule sequencing platforms have so far found more utility within studies of the microbial world than of mammalian systems.

The extraordinary utility of next-generation sequencing (NGS) makes it an ideal technology for many different types of measurements, beyond just sequencing genomes, constructs, and RNA expression levels. To the extent that protein and other analytes can be transduced into nucleic acids it may be possible to deconvolute extremely complex mixtures using NGS. For example, an antibody library tagged with unique DNA tags could be used to coordinately identify the presence and amounts of proteins on a cell surface or in a lysate. Transduction schemes for small molecules based on ligand-dependent nucleic acid conformational changes can also be envisioned. Protein modification states and epigenetic tags could be followed using similar implementations. The downside to such measurements is that they are not in real time and resolution may be lost through the transduction process.

Molecular sensing, molecular recognition, and cell signaling comprise a diverse set of fundamental biological processes. Commensurately, there is a diverse set of design options for engineering responses to environmental or internal cellular conditions. This flexibility is further increased by the success of taking a modular approach to the sensing process, making it easier to vary what is sensed and what happens as a consequence. Integration with designed cellular circuitry creates the potential for many options for control, memory, logical operations, and multiplexing. Building context-dependent sensors into microorganisms is one potential path to obtain subcellular measurements despite their small size. Overall, these phenomena can be used to help debug a living system under development, to provide feedback to living cells, or as subsystems within an *ex vivo* measurement solution for research, production, diagnostics, or environmental monitoring. In some cases, biosensor systems can be run

Copyright © National Academy of Sciences. All rights reserved.

in vitro and thus be exported to cell-free systems in solution or on solid supports. Conversely, advanced cell-free systems can be used to debug the biosensor before it is deployed *in vivo*. If the technical challenges can be overcome, then the number of situations in which biosensors can be expected to enable rapid, low-cost, high-throughput testing of individually engineered cells or, if desired, entire populations of engineered cells seems certain to increase.

Beyond sequencing, many additional measurements can be made to assess the performance of a given circuit. A field of measurement that is particularly essential to both engineering biology and the elucidation of human biology is metabolomics. The chemical industry's interest in new biology-based routes for producing products is very much focused on metabolite production. Perhaps there is no better indicator than the nomenclature "metabolic engineering" for the traditional development of new organisms for better industrial bioprocessing. There is, as a consequence, a long history of developments that adapt metabolite measurements to the interests and needs of engineering biology. For example, there is already a rich diversity of laboratory and data analysis methods designed to identify and follow the pathways of metabolite production and modification. Because of the viability of progressing toward the creation of quantitative models for the associated chemical reactions in microorganisms, there is an especially close relationship between modeling and measurement in this field. They are advancing together and synergistically. Still, the most universal metabolite measurement solutions are too slow to meet the potential of the information gained while the highest-throughput methods require specialized optimization on a case-by-case basis. As is also true of proteomics measurements, faster, generalizable analysis of metabolites would greatly accelerate learning and the associated models that can encapsulate the results in ways that illuminate preferred steps throughout the engineering cycle. Here too, sensitivity to the level of single cells without compromising throughput will be of value. While there are important aspects of human biology that would also advance greatly with the advent of higher-throughput metabolomics, again there is a divergence of needs, especially because of the difference in scope, prior information, and sophistication of models between studies of industrially relevant microorganisms and human biology.

At the same time broadly useful measurement platforms are extended for use by biological engineers, synthetic biology is enabling and introducing new measurement paradigms particularly well suited for the needs of engineering biology. Circuits can be easily linked to readily observed reporters, such as green fluorescent protein, and high-throughput devices and methods that have already been developed, such as plate readers or fluorescence-activated cell sorting, can be used to parse performance.

Copyright © National Academy of Sciences. All rights reserved.

Into the future, the development of additional reporters and analytical methods that can scale to even greater numbers may be important. *In vivo* measurements, genetic manipulations, intervention in regulatory processes, interference with molecular intermediates, such as expressed RNA transcripts, and chemical signal-induced alterations have been long-standing tools for generating and validating molecular biological hypotheses. Engineering biology brings a perspective of altering organisms for utilitarian purposes including the development of biosensor measurement devices at both the molecular and the organism levels for readout, feedback, and control.

In parallel, as circuits become increasingly complex there will be a need to increase the number of different parameters that can be measured in parallel, such as the expression of multiple genes or the production of multiple metabolites. This would argue for the development of higher-throughput methods that are capable of analyzing whole organisms or chemical mixtures, such as mass spectrometry or nuclear magnetic resonance. There are foundational technologies in miniaturization, micro- and nanofluidics, photonics, nucleic acid synthesis chemistry, and data analysis that could feed into and enable these desired advances in measurement solutions. Standards for metrics and materials will also facilitate convergence and deployment of the best methods along with ensuring reproducibility and transferability across laboratories, manufacturing sites, and institutions.

A second consideration in the development of analytical methodology for the assessment of synthetic circuitry is ensuring that the measurements being performed accurately reflect the performance of a given organism in an industrial setting. It does little good to optimize a circuit in the laboratory only to find that it does not work in a vat. In contrast, it is very difficult to carry out high-throughput experiments that scale to even small fermenters. Thus, it becomes important to rationally understand how the readouts of organismal metabolism scale from benchtop experiments to test bed to production. This in turn requires greater integration between systems modeling tools for gene expression and metabolism and analytical methods. The results of baseline experiments with a given chassis or circuit under a variety of conditions need to be compared with similar results in batch, or under different fermentation conditions, in order to develop feedback loops that will allow prediction of how augmentation of the chassis or circuit will perturb both the initial readings and the final performance.

Copyright © National Academy of Sciences. All rights reserved.

Key Conclusions

Conclusion: The ability to rapidly, routinely, and reproducibly measure pathway function and cellular physiology will drive the development of novel enzymes and pathways, which are needed to increase the array of efficient and low-cost chemical transformations available for use in biomanufacturing.

Conclusion: The fall in cost and increase in throughput of measurement technologies should track that of strain engineering technologies and vice versa.

Roadmap Goals

- Within 4 years, develop the ability to routinely and reproducibly measure nucleic acids, proteins, and metabolites targeted to characterize 50 or more high-priority, selected model parameters for 2,000 strains and measure 1,000 or more parameters for 200 strains within 1 week at a cost no higher than the full cost of designing and building those strains.
- Within 10 years, have the ability to routinely and reproducibly measure 50 or more high-priority, selectable model parameters *in vivo* at the same cost and speed as above.

Copyright © National Academy of Sciences. All rights reserved.

Copyright © National Academy of Sciences. All rights reserved.

5

What Is Success and How to Get There: Recommendations

This report has described the structure of the current chemical manufacturing process and explored the promise of increased application of biological processes to chemical production. Lowered costs, increases in production speed, flexibility of manufacturing plants, and increased production capacity are among the many potential benefits that the increased industrialization of biology may bring to producers and consumers of chemical products. As outlined in Chapter 2, the production of chemicals through biological processes may help to reduce toxic by-products, to reduce greenhouse gas emissions, and to lower fossil fuel consumption in chemical production. The advanced manufacturing of chemicals through biology can help address global challenges related to energy, climate change, sustainable and more productive agriculture, and environmental sustainability.

Realizing the significant benefits of the continued and more efficient industrialization of biology requires the sustained effort of multiple stakeholders. This chapter offers several recommendations to specific stakeholders designed to facilitate the achievement of the technical milestones detailed in Chapter 4.

Additionally, recognizing the significant role that societal factors will play in the continued industrialization of biology, this chapter puts forth recommendations focused on the impact of economics, education and workplace issues, and governance in facilitating the industrialization of biology. This chapter addresses these societal factors and offers sev-

Copyright © National Academy of Sciences. All rights reserved.

Industrialization of Biology: A Roadmap to Accelerate the Advanced Manufacturing of Chemicals

eral recommendations to foster the achievement of critical societal goals related to the industrialization of biology.

Ensuring the rapid industrialization of biology will require, one, the selection of advantageous chemicals, materials, and fuel targets, based on technical and economic criteria as well as social benefits as embodied in governance criteria; two, the development of broader and deeper scientific understanding in support of the industrialization of biology; and three, engagement with the public at large who are impacted by the acceleration of this industry. The recommendations presented in this chapter are designed to address each of these three factors, with the ultimate goal of putting biological synthesis and engineering on par with chemical synthesis and engineering for chemical manufacturing.

HOW DO WE GET THERE?

Realization of the promise of the industrialization of biology for chemical manufacturing can only be achieved through a sustained effort among multiple stakeholders. The challenge is even more daunting in an era of fiscal austerity, of technological complexity, and of regulatory uncertainty. To meet the Statement of Task, the Committee has constructed a roadmap based on a current view of technology, markets, and societal considerations. Any roadmap is accurate only at a point in time. In a fast-evolving field, a roadmap can only remain useful if it is updated at some frequency. As a result, the Committee believes it is essential to create a mechanism that provides for an ongoing road-mapping process.

The UK recently established the Synthetic Biology Leadership Council (SBLC) to maintain the momentum of the UK Synthetic Biology Roadmap. The UK SBLC has representatives from multiple stakeholder groups, including government, academia, and industry. In 2012, Research Councils UK convened a coordination group to oversee the creation of the UK Synthetic Biology Roadmap. Subsequently, the UK government instituted the SBLC as a steering structure governance body to assess progress and update recommendations and shape priorities for future implementation of the roadmap in the UK. The SBLC provides a visible point for strategic coordination between the funding agencies, the research community, industry, government sponsors, and other stakeholders, including societal and ethical representatives.

Within the United States, Synberc is a multi-university research center established in 2006 with a grant from the National Science Foundation (NSF) to help lay the foundation for synthetic biology. Eighteen institutions are currently involved. Synberc has also added nearly 50 industrial partners. Its mission does not include roadmapping, but it does focus on the foundational science and technology for synthetic biology, as well as

Copyright © National Academy of Sciences. All rights reserved.

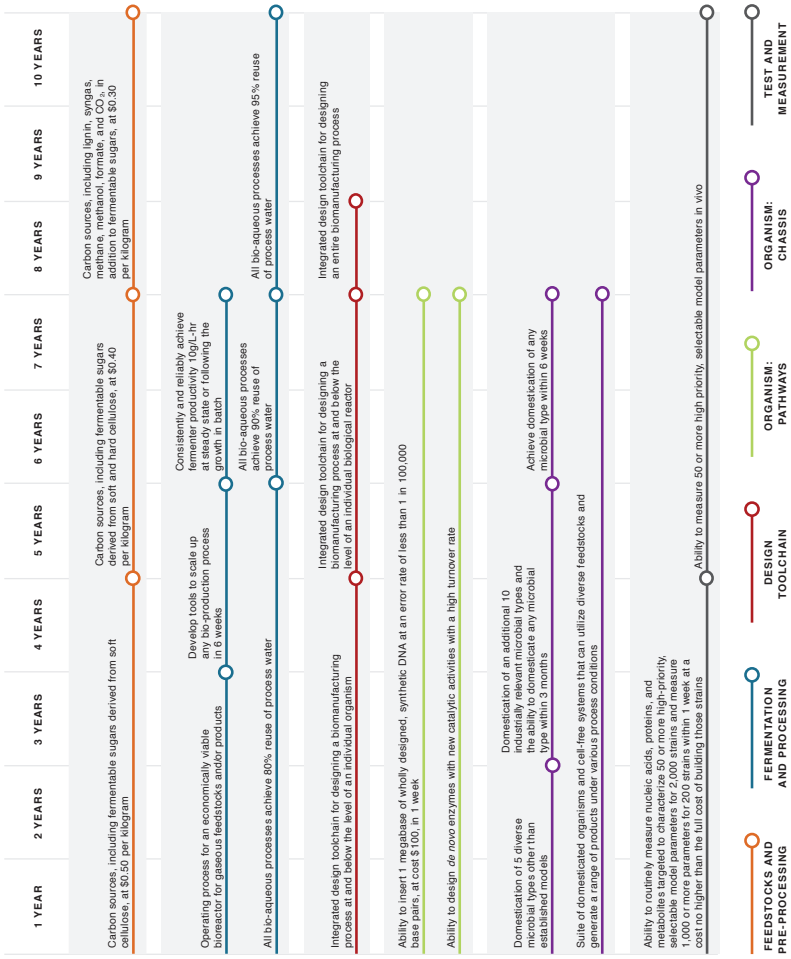


FIGURE 5-1 Technical roadmap to enable the industrialization of biology.
NOTE: A larger version of this roadmap can be found as a foldout at the end of this book.

Copyright © National Academy of Sciences. All rights reserved.

Industrialization of Biology: A Roadmap to Accelerate the Advanced Manufacturing of Chemicals

capability-building and public engagement. Synberc is a potential model for taking on the ongoing roadmapping work

One successful example of roadmapping in another technology area is Sematech. Dating back nearly 30 years, Sematech was founded as a consortium between the U.S. government and the American semiconductor industry, with some initial funding from DARPA (Defense Advanced Research Projects Agency). Among its important functions was the maintenance of the technology roadmap for semiconductors. Since its founding, Sematech has evolved to a global industry consortium, fully funded by its members.

The Committee recommends that the relevant government agencies consider establishment of an ongoing roadmapping mechanism to provide direction to technology development, translation, and commercialization at scale. This roadmapping effort would bring together participants from public and private research, and participants with all skill sets needed for the industrialization of biology. In addition to maintaining the roadmap, this effort could assist in sharing the knowledge, tools, and data needed to accelerate progress. It is recognized that a number of well-functioning processes and organizations are already meeting needs in industrialization of biology, and the suggested roadmapping effort would not usurp the existing mechanisms, but would help to coordinate these activities with other elements. Moreover the roadmapping effort could help to address a set of difficult, core technical challenges that must be overcome. It might help to develop, share, and diffuse common interoperable standards, languages, and measurements. Roadmapping would also assist in prioritizing efforts for creating new enabling tools or data.

The lessons learned from roadmaps and consortia in comparable domains demonstrate that well-designed and well-executed strategic processes can accelerate time frames, help prioritize objectives, and make the industrialization more transparent, responsible, and accessible.

The Committee recognizes that any decisions on a roadmapping process would be within the purview of the interested federal agencies. Based on the UK experience, it would be possible to have a functioning process within 2 years. Within 5 years, such an effort could contribute materially to our national capability to develop and scale up bioprocesses for the manufacture of chemicals. Within 10 years, one can foresee the broad diffusion and acceptance of bioprocessing as a core foundation of the chemical economy, of advanced manufacturing, and of American competitiveness in the bioeconomy.

TECHNICAL NEEDS AND ROADMAP

Chapter 4 laid out critical technical milestones and roadmap goals (see Figure 5-1) for feedstocks, chemical transformations, organism and

Copyright © National Academy of Sciences. All rights reserved.

pathway design, and measurement techniques. Achieving these milestones will take predictable and consistent investment to develop the scientific knowledge and technical tools.

Conclusion: Biomanufacturing of chemicals is already a significant element of the national economy and is poised for rapid growth during the next decade. Both the scale and scope of biomanufacturing of chemicals will expand and will involve both high-value and high-volume chemicals. Progress in the areas identified in this report will play a major role in achieving the challenge of increasing the contribution of biotechnology to the national economy.

Recommendation: In order to transform the pace of industrial biotechnology by enabling commercial entities to develop new biomanufacturing processes, the National Science Foundation (NSF), Department of Energy, National Institutes of Health, National Institute of Standards and Technology, Department of Defense, and other relevant agencies should support the scientific research and foundational technologies required to advance and integrate the areas of feedstocks, organismal chassis and pathway development, fermentation, and processing as outlined in the roadmap goals.

Supporting foundational research in these areas is critical to the growing commercial viability of biological processes in chemical manufacturing. Specifically, it is recommended that these agencies support research focused on the following:

- Improving the availability of economic and environmentally sustainable feedstocks;
- Increasing the availability, reliability, and sustainability of biofeedstocks, in order to increase the range of economically viable products, provide more predictive levels and quality of feedstock, and lower barriers to entry into the biological production of chemicals;
- Improving the basic understanding of C1-based fermentation, in light of the increased availability of natural gas in the United States;
- Improving the productivity of the fermentation process, by means of enhanced mass and heat transfer, continuous product removal, more extensive use of co-cultures, co-products, and co-substrates, where continued development of fundamental science and enabling technologies is required for the rapid and efficient development of organismal chassis and pathways;

Copyright © National Academy of Sciences. All rights reserved.

- Expanding the palette of domesticated microbial and cell-free platforms for biomanufacturing;
- Cultivating robust strains that remain genetically stable and retain performance stability over time in the presence of diverse feedstocks and products;
- Developing the ability to rapidly develop enzymes with respect to catalytic activity and specific activity; and
- Rapidly, routinely, and reproducibly measuring pathway function and cellular physiology.

This list is not intended to be exhaustive but rather to highlight those areas that are most directly related to the technical roadmap goals.

NONTECHNICAL INSIGHTS AND SOCIETAL CONCERNS

Economic

Meeting the technical and scientific challenges involved in the industrialization of biology is necessary to realize the potential benefits, but ensuring that those benefits accrue rapidly and with maximum positive impact requires accurate and detailed information about the role of bio-based production in the economy. The ability to predict economic trends, to assess economic impact, and to more completely understand the role of bio-based products in the economy will enable better decision making for all stakeholders involved in the industrialization of biology.

Recommendation: The U.S. government should perform a regular quantitative measure of the contribution of bio-based production processes to the U.S. economy to develop a capacity for forecasting and assessing economic impact.

Improved quantitative measures of the impact of bio-based production will be valuable to a range of stakeholders, but these measures will directly affect both policy makers and business leaders: policy makers will be better able to set budget estimates and projections, and business leaders will be better able to assess market size and direction. By measuring this area of economic activity, those involved will be able to make more informed decisions, potentially leading to significantly increased efficiency.

Education and Workforce

The industrialization of biology will create new structures of work, place new skills in demand, and necessitate the development of new

Copyright © National Academy of Sciences. All rights reserved.

expertise in the biological and chemical sciences, engineering, and computing in the workforce. Changing workforce demands will require changes in education and training.

Recommendation: Industrial biotechnology firms individually, and especially through industry groups, should strengthen their partnerships with all levels of academia, from community colleges, undergraduate institutions, and graduate institutions, to communicate changing needs and practices in industry in order to inform and influence academic instruction.

Without communication and partnership between academia and industry, skills that are emphasized in academia may not be useful or valued in industry. Developing balanced training portfolios for technicians, subject-matter experts, and biological designers is important, but is only possible through the active engagement of both industry and academic institutions. The aforementioned collaboration framework can likely be one mode by which these connections are facilitated.

Affording students the opportunity to experience industrial lab settings carries significant benefits to both students and future employers. The ability to plan for large-scale production and skill in developing significant scientific results into tangible, useful products are critical capabilities that the present and future chemical manufacturing demands. Ensuring that academia is providing students with the ability to function in both academic and industrial settings requires the active participation of both industry and academia.

Biology is already playing a large role in chemical manufacturing in the United States. Chemical manufacturers utilizing bio-based processes can help to develop the workforce necessary for the future structure of chemical manufacturing. By encouraging the training of a skilled workforce prepared to work in this emerging field, students and trainees should have the opportunity to explore the field early in their academic careers.

Recommendation: Federal agencies, academia, and industry should devise and support innovative approaches toward expanding the exposure of student trainees to design-build-test-learn paradigms in a high-throughput fashion and at industrial scale.

The needs and tools of industry are rapidly changing. Chemical production at very large scales and with extensive automation is frequently very distinct from academic experience. Partnerships between universities and industry will allow students and trainees to be exposed to the

Copyright © National Academy of Sciences. All rights reserved.

concerns, techniques, and needs of industry, which will help to create a workforce better prepared to think and function in this new economic environment.

Governance

The impact of the industrialization of biology on society will be mediated by a governance framework. Ensuring that this framework balances important social values is critical. In order to do so, a governance framework should involve a variety of policy approaches, including education, self-governance through standard setting, accreditation, government regulation, public engagement and public scrutiny, and tort liability, among other methods.

Safety, sustainability, security, and resilience are critical goals for any governance framework. These values sometimes cause tension and any governance framework will have to balance these competing demands. In order to do so, a governance framework must have legitimacy in the eyes of the public and the industry. To be successful, a governance framework should be perceived as fair, transparent, efficient, and inclusive of diverse viewpoints.

Recommendation: The administration should ensure that the Environmental Protection Agency (EPA), Commerce Department, U.S. Department of Agriculture (USDA), Food and Drug Administration (FDA), Occupational Safety and Health Administration, National Institute of Standards and Technology (NIST), and other relevant agencies work together to broadly assess, and regularly reassess, the adequacy of existing governance, including but not limited to regulation, and to identify places where industry, academia, and the public can contribute to or participate in governance.

Recommendation: Science funding agencies and science policy offices should ensure outreach efforts that facilitate responsible innovation by enabling the extension of existing relevant regulatory practices, concordance across countries, and increased public engagement.

Coordination across government bodies, combined with a commitment to transparency and public contribution and participation, will enable a governance framework that is at once navigable, perceived as legitimate, and achieves the societal goals critical to the public welfare.

Copyright © National Academy of Sciences. All rights reserved.

Moreover, the governance framework established should be capable of gathering and utilizing information about the risks posed by new techniques and products.

Recommendation: Government agencies, including EPA, USDA, FDA, and NIST should establish programs both for the development of fact-based standards and metrology for risk assessment in industrial biotechnology and for the use of these fact-based assessments in evaluating and updating the governance regime.

CONCLUDING REMARKS

The industrialization of biology offers the prospect of addressing global as well as American national interests. The recommendations put forward are designed to facilitate the achievement of the roadmap goals and, ultimately, the challenge posed by the committee: to double the percentage of gross domestic product that comes from the bioeconomy by putting biological synthesis and engineering on par with chemical synthesis and engineering for chemical manufacturing. It is important to note the urgency of these recommendations: scientific, technological, environmental, and economic trends are converging *now* that are creating positive conditions for the rapid industrialization of biology. Advanced chemical manufacturing through the industrialization of biology will require new tools, new knowledge, and new financial mechanisms. It promises new investment opportunities, new platforms for designing biological systems for next-generation American manufacturing, and opportunities to enhance competitiveness and create well-paying jobs.

Copyright © National Academy of Sciences. All rights reserved.

Copyright © National Academy of Sciences. All rights reserved.

References

1. *National Bioeconomy Blueprint*; The White House: Washington, DC, 2012.
2. (a) Carlson, R. Synthesis. The U.S. Bioeconomy in 2012 Reached $350 Billion in Revenues, or About 2.5% of GDP. http://www.synthesis.cc/2014/01/the-us-bioeconomy-in-2012.html (accessed July 18, 2014); (b) Solomon, D. Industrial Views on Synthetic Biology. Presented at Tooling the U.S. Bioeconomy: Synthetic Biology Conference, Washington, DC, November 5, 2013. ACS Science & the Congress Project, 2013.
3. Cha, A. E. Companies Rush to Build 'Biofactories' for Medicines, Flavorings and Fuels. *The Washington Post*, October 24, 2013. http://www.washingtonpost.com/national/health-science/companies-rush-to-build-biofactories-for-medicines-flavorings-and-fuels/2013/10/24/f439dc3a-3032-11e3-8906-3daa2bcde110_story.html (accessed December 2, 2014)
4. Golden, J. S.; Handfield, R. B. *Why Biobased? Opportunities in the Emerging Bioeconomy*; U. S. Department of Agriculture: Washington, DC, 2014.
5. (a) Kosuri, S.; Church, G. M. Large-scale de novo DNA synthesis: technologies and applications. *Nat. Methods* **2014**, *11(5)*, 499-507; (b) Carlson, R. The Pace and Proliferation of Biological Technologies. *Biosecurity and Bioterrorism* **2003**, *1(3)*, 203-14.
6. National Human Genome Research Institute. The Human Genome Project Completion: Frequently Asked Questions. http://www.genome.gov/11006943 (accessed February 2, 2015).
7. (a) 1000 Genomes Project Consortium. An Integrated Map of Genetic Variation from 1,092 Human Genomes. *Nature* **2012**, *491(7422)*, 56-65; (b) Clark, L. Illumina Announces Landmark $1,000 Human Genome Sequencing. http://www.wired.co.uk/news/archive/2014-01/15/1000-dollar-genome (accessed December 30, 2014).
8. (a) Benson, D. A.; Karsch-Mizrachi, I.; Lipman, D. J.; Ostell, J.; Wheeler, D. L. GenBank. *Nucleic Acids Res.* **2008**, *36(Database Issue)*, D25-30; (b) Benson, D. A.; Karsch-Mizrachi, I.; Lipman, D. J.; Ostell, J.; Sayers, E. W. GenBank. *Nucleic Acids Res.* **2009**, *37(Database Issue)*, D26-31.

Copyright © National Academy of Sciences. All rights reserved.

9. Wang, H. H.; Isaacs, F. J.; Carr, P. A.; Sun, Z. Z.; Xu, G.; Forest, C. R.; Church, G. M. Programming cells by multiplex genome engineering and accelerated evolution. *Nature* **2009**, *460(7257)*, 894-8.
10. Haurwitz, R. E.; Jinek, M.; Wiedenheft, B.; Zhou, K.; Doudna, J. A. Sequence- and Structure-Specific RNA Processing by a CRISPR Endonuclease. *Science* **2010**, *329(5997)*, 1355-8.
11. (a) Ladisch, M. The Role of Bioprocess Engineering in Biotechnology. *The Bridge* **2004**, *34(3)*, 26-32; (b) Mosier, N. S.; Ladisch, M. R. Biotechnology. In *Modern Biotechnology: Connecting Innovations in Microbiology and Biochemistry to Engineering Fundamentals*; John Wiley & Sons: Hoboken, NJ, 2011; pp 1-25.
12. OECD (Organisation for Economic Co-operation and Development). *The Application of Biotechnology to Industrial Sustainability*; OECD Publishing: France, 2001.
13. National Research Council. *A New Biology for the 21st Century;* The National Academies Press: Washington, DC, 2009.
14. Obama, B. Remarks by the President on the Economy in Osawatomie, Kansas. http://www.whitehouse.gov/the-press-office/2011/12/06/remarks-president-economy-osawatomie-kansas (accessed December 20, 2014).
15. Merriam-Webster. Biotechnology in *Merriam-Webster*. http://www.merriam-webster.com/dictionary/biotechnology (accessed February 3, 2015).
16. Merriam-Webster. Genetic Engineering in *Merriam Webster*. http://www.merriam-webster.com/dictionary/genetic%20engineering (accessed February 3, 2015).
17. UK Synthetic Biology Roadmap Coordination Group. A Synthetic Biology Roadmap for the UK; Technology Strategy Board: Swindon, Wiltshire, 2012.
18. Mutalik, V. K.; Guimaraes, J. C.; Cambray, G.; Lam, C.; Christoffersen, M. J.; Mai, Q. A.; Tran, A. B.; Paull, M.; Keasling, J. D.; Arkin, A. P.; Endy, D. Precise and Reliable Gene Expression via Standard Transcription and Translation Initiation Elements. *Nat. Methods* **2013**, *10(4)*, 354-60.
19. The European Commission. *The European Bioeconomy in 2030: Delivering Sustainable Growth by Addressing the Grand Societal Challenges*, 2012. http://www.epsoweb.org/file/560 (accessed January 12, 2015).
20. de Jong, E.; Higson, A.; Walsh, P.; Wellisch, M. *Bio-based Chemicals Value Added Products from Biorefineries* [online]; IEA Bioenergy: Wageningen, The Netherlands, 2012. http://www.ieabioenergy.com/wp-content/uploads/2013/10/Task-42-Biobased-Chemicals-value-added-products-from-biorefineries.pdf (accessed December 12, 2014).
21. (a) OECD. *Industrial Biotechnology and Climate Change: Opportunities and Challenges* [online]; OECD Publishing: 2011. http://www.oecd.org/sti/biotech/49024032.pdf (accessed December 11, 2014); (b) OECD. *Emerging Policy Issues in Synthetic Biology* [online]; OECD Publishing, 2014. http://dx.doi.org/10.1787/9789264208421-en (accessed December 11, 2014).
22. OECD. *The Bioeconomy to 2030: Designing a Policy Agenda*; OECD Publishing, 2009.
23. Palsson, B. Cell Factory Design. Presented at Workshop on the Industrialization of Biology, May 28, 2014.
24. BCC Research. *Synthetic Biology: Global Markets*; BCC Research: Wellesley, MA, 2014.
25. Milken Institute. *Unleashing the Power of the Bio-Economy*; Milken Institute: Santa Monica, CA, 2013.
26. McKinsey Global Institute. *Disruptive Technologies: Advances that Will Transform Life, Business, and the Global Economy*; McKinsey & Company: Washington, DC, 2013.
27. Kelley, N. J.; Whelan, D. J.; Kerr, E.; Apel, A.; Beliveau, R.; Scanlon, R. Engineering Biology to Address Global Problems: Synthetic Biology Markets, Needs, and Applications. *Ind. Biotechnol.* **2014**, *10(3)*, 140-9.
28. MIT (Massachusetts Institute of Technology). *The Third Revolution: The Convergence of the Life Sciences, Physical Sciences, and Engineering* [online]; MIT Washington Office:

Copyright © National Academy of Sciences. All rights reserved.

Washington, DC, 2011. http://dc.mit.edu/sites/dc.mit.edu/files/MIT%20White%20Paper%20on%20Convergence.pdf (accessed December 4, 2014).

29. AAAS (American Academy of Arts and Sciences). ARISE II: Unleashing America's Research & Innovation Enterprise [online]; AAAS: Washington, DC, 2013. https://www.amacad.org/multimedia/pdfs/publications/researchpapersmonographs/arise2.pdf (accessed October 10, 2014).
30. Kopchik, K. Bucknell Forum: Designer Neri Oxman to Speak Tonight. The Bucknellian, 2010.
31. OECD. Emerging Policy Issues in Synthetic Biology [online]; OECD Publishing, 2014. http://dx.doi.org/10.1787/9789264208421-en (accessed December 11, 2014).
32. Serger, S. S.; Breidne, M. China's Fifteen-Year Plan for Science and Technology: An Assessment. *Asia Pol'y* **2007**, *4(1)*, 135-64.
33. (a) Kodumal, S. J.; Patel, K. G.; Reid, R.; Menzella, H. G.; Welch, M.; Santi, D. V. Total synthesis of long DNA sequences: Synthesis of a contiguous 32-kb polyketide synthase gene cluster. *Proc. Natl. Acad. Sci. U. S. A.* **2004**, *101(44)*, 15573-8; (b) Bayer, T. S.; Widmaier, D. M.; Temme, K.; Mirsky, E. A.; Santi, D. V.; Voigt, C. A. Synthesis of Methyl Halides from Biomass Using Engineered Microbes. *J. Am. Chem. Soc.* **2009**, *131(18)*, 6508-15.
34. (a) Gibson, D. G.; Glass, J. I.; Lartigue, C.; Noskov, V. N.; Chuang, R.-Y.; Algire, M. A.; Benders, G. A.; Montague, M. G.; Ma, L.; Moodie, M. M.; Merryman, C.; Vashee, S.; Krishnakumar, R.; Assad-Garcia, N.; Andrews-Pfannkoch, C.; Denisova, E. A.; Young, L.; Qi, Z.-Q.; Segall-Shapiro, T. H.; Calvey, C. H.; Parmar, P. P.; Hutchison, C. A.; Smith, H. O.; Venter, J. C. Creation of a Bacterial Cell Controlled by a Chemically Synthesized Genome. *Science* **2010**, *329(5987)*, 52-56; (b) Annaluru, N.; Muller, H.; Mitchell, L. A.; Ramalingam, S.; Stracquadanio, G.; Richardson, S. M.; Dymond, J. S.; Kuang, Z.; Scheifele, L. Z.; Cooper, E. M.; Cai, Y.; Zeller, K.; Agmon, N.; Han, J. S.; Hadjithomas, M.; Tullman, J.; Caravelli, K.; Cirelli, K.; Guo, Z.; London, V.; Yeluru, A.; Murugan, S.; Kandavelou, K.; Agier, N.; Fischer, G.; Yang, K.; Martin, J. A.; Bilgel, M.; Bohutski, P.; Boulier, K. M.; Capaldo, B. J.; Chang, J.; Charoen, K.; Choi, W. J.; Deng, P.; DiCarlo, J. E.; Doong, J.; Dunn, J.; Feinberg, J. I.; Fernandez, C.; Floria, C. E.; Gladowski, D.; Hadidi, P.; Ishizuka, I.; Jabbari, J.; Lau, C. Y.; Lee, P. A.; Li, S.; Lin, D.; Linder, M. E.; Ling, J.; Liu, J.; Liu, J.; London, M.; Ma, H.; Mao, J.; McDade, J. E.; McMillan, A.; Moore, A. M.; Oh, W. C.; Ouyang, Y.; Patel, R.; Paul, M.; Paulsen, L. C.; Qiu, J.; Rhee, A.; Rubashkin, M. G.; Soh, I. Y.; Sotuyo, N. E.; Srinivas, V.; Suarez, A.; Wong, A.; Wong, R.; Xie, W. R.; Xu, Y.; Yu, A. T.; Koszul, R.; Bader, J. S.; Boeke, J. D.; Chandrasegaran, S. Total Synthesis of a Functional Designer Eukaryotic Chromosome. *Science* **2014**, *344(6179)*, 55-8.
35. Nielsen, A. A. K.; Segall-Shapiro, T. H.; Voigt, C. A. Advances in Genetic Circuit Design: Novel Biochemistries, Deep Part Mining, and Precision Gene Expression. *Curr. Opin. Chem. Biol.* **2013**, *17(6)*, 878-92.
36. Hsu, P. D.; Scott, D. A.; Weinstein, J. A.; Ran, F. A.; Konermann, S.; Agarwala, V.; Li, Y.; Fine, E. J.; Wu, X.; Shalem, O.; Cradick, T. J.; Marraffini, L. A.; Bao, G.; Zhang, F. DNA targeting specificity of RNA-guided Cas9 nucleases. *Nat. Biotechnol.* **2013**, *31(9)*, 827-32.
37. (a) Becker, S. A.; Feist, A. M.; Mo, M. L.; Hannum, G.; Palsson, B. O.; Herrgard, M. J., Quantitative Prediction of Cellular Metabolism with Constraint-Based Models: the COBRA Toolbox. *Nat. Protoc.* **2007**, *2(3)*, 727-738; (b) Burgard, A. P.; Pharkya, P.; Maranas, C. D. Optknock: A Bilevel Programming Framework for Identifying Gene Knockout Strategies for Microbial Strain Optimization. *Biotechnology & Bioengineering* **2003**, *84(6)*, 647-57.
38. Smanski, M. J.; Bhatia, S.; Zhao, D.; Park, Y.; B A Woodruff, L.; Giannoukos, G.; Ciulla, D.; Busby, M.; Calderon, J.; Nicol, R.; Gordon, D. B.; Densmore, D.; Voigt, C. A. Functional Optimization of Gene Clusters by Combinatorial Design and Assembly. *Nat. Biotechnol.* **2014**, *32(12)*, 1241-9.

Copyright © National Academy of Sciences. All rights reserved.

39. Brophy, J. A. N.; Voigt, C. A. Principles of Genetic Circuit Design. *Nat. Methods* **2014**, *11(5)*, 508-520.
40. Cardinale, S.; Arkin, A. P., Contextualizing Context for Synthetic Biology – Identifying Causes of Failure of Synthetic Biological Systems. *Biotechnol. J.* **2012**, *7(7)*, 856-66.
41. (a) Mutalik, V. K.; Guimaraes, J. C.; Cambray, G.; Lam, C.; Christoffersen, M. J.; Mai, Q. A.; Tran, A. B.; Paull, M.; Keasling, J. D.; Arkin, A. P.; Endy, D. Precise and reliable gene expression via standard transcription and translation initiation elements. *Nat. Methods* **2013**, *10(4)*, 354-60; (b) Baker, D.; Church, G.; Collins, J.; Endy, D.; Jacobson, J.; Keasling, J.; Modrich, P.; Smolke, C.; Weiss, R. Engineering Life: Building a FAB for Biology. *Sci. Am.* **2006**, *294(6)*, 44-51.
42. Galdzicki, M.; Clancy, K. P.; Oberortner, E.; Pocock, M.; Quinn, J. Y.; Rodriguez, C. A.; Roehner, N.; Wilson, M. L.; Adam, L.; Anderson, J. C.; Bartley, B. A.; Beal, J.; Chandran, D.; Chen, J.; Densmore, D.; Endy, D.; Grunberg, R.; Hallinan, J.; Hillson, N. J.; Johnson, J. D.; Kuchinsky, A.; Lux, M.; Misirli, G.; Peccoud, J.; Plahar, H. A.; Sirin, E.; Stan, G.-B.; Villalobos, A.; Wipat, A.; Gennari, J. H.; Myers, C. J.; Sauro, H. M. The Synthetic Biology Open Language (SBOL) Provides a Community Standard for Communicating Designs in Synthetic Biology. *Nat. Biotechnol.* **2014**, *32(6)*, 545-50.
43. (a) Donia, M. S.; Cimermancic, P.; Schulze, C. J.; Wieland Brown, L. C.; Martin, J.; Mitreva, M.; Clardy, J.; Linington, R. G.; Fischbach, M. A. A Systematic Analysis of Biosynthetic Gene Clusters in the Human Microbiome Reveals a Common Family of Antibiotics. *Cell* **2014**, *158(6)*, 1402-14; (b) Scharschmidt, T. C.; Fischbach, M. A. What Lives On Our Skin: Ecology, Genomics and Therapeutic Opportunities Of the Skin Microbiome. *Drug Discovery Today: Dis. Mech.* **2013**, 10(3-4), e83–9.
44. (a) Hatzimanikatis, V.; Li, C.; Ionita, J. A.; Henry, C. S.; Jankowski, M. D.; Broadbelt, L. J. Exploring the Diversity of Complex Metabolic Networks. *Bioinformatics* **2005**, *21(8)*, 1603-1609; (b) Li, C.; Henry, C. S.; Jankowski, M. D.; Ionita, J. A.; Hatzimanikatis, V.; Broadbelt, L. J. Computational Discovery of Biochemical Routes to Specialty Chemicals. *Chem. Eng. Sci.* **2004**, *59(22-23)*, 5051-60.
45. Srivastava, S.; Kotker, J.; Hamilton, S.; Ruan, P.; Tsui, J.; Anderson, J. C.; Bodik, R.; Seshia, S. A. In Pathway Synthesis Using the Act Ontology in Proceedings of the 4th International Workshop on Bio-Design Automation (IWBDA): San Francisco, CA, 2012.
46. Lu, T. K.; Khalil, A. S.; Collins, J. J. Next-generation synthetic gene networks. *Nat. Biotechnol.* **2009**, *27(12)*, 1139-50.
47. Zhang, F.; Carothers, J. M.; Keasling, J. D. Design of a Dynamic Sensor-Regulator System for Production of Chemicals and Fuels Derived from Fatty Acids. *Nat. Biotechnol.* **2012**, *30(4)*, 354-9.
48. Chen, A. Y.; Deng, Z.; Billings, A. N.; Seker, U. O. S.; Lu, Michelle Y.; Citorik, R. J.; Zakeri, B.; Lu, T. K. Synthesis and patterning of tunable multiscale materials with engineered cells. *Nat. Mater.* **2014**, *13(5)*, 515-23.
49. Bennett, J. W. The Time Line Adrenalin and cherry trees. *Mod. Drug Discovery* **2001**, *4*, 47-8.
50. Shuler, M. L.; Kargi, F. *Bioprocess Engineering: Basic Concepts*. Prentice Hall: Upper Saddle River, New Jersey, 2002.
51. Cohen, S. N.; Chang, A. C.; Boyer, H. W.; Helling, R. B. Construction of biologically functional bacterial plasmids in vitro. *Proc. Natl. Acad. Sci. U. S. A.* **1973**, *70(11)*, 3240-4.
52. Bailey, J. E. Toward a Science of Metabolic Engineering. *Science* **1991**, *252(5013)*, 1668-75.
53. Stephanopoulos, G.; Vallino, J. Network rigidity and metabolic engineering in metabolite overproduction. *Science* **1991**, *252(5013)*, 1675-81.
54. Bornscheuer, U. T.; Huisman, G. W.; Kazlauskas, R. J.; Lutz, S.; Moore, J. C.; Robins, K. Engineering the third wave of biocatalysis. *Nature* **2012**, *485(7397)*, 185-94.
55. Savile, C. K.; Janey, J. M.; Mundorff, E. C.; Moore, J. C.; Tam, S.; Jarvis, W. R.; Colbeck, J. C.; Krebber, A.; Fleitz, F. J.; Brands, J.; Devine, P. N.; Huisman, G. W.; Hughes,

Copyright © National Academy of Sciences. All rights reserved.

G. J. Biocatalytic Asymmetric Synthesis of Chiral Amines from Ketones Applied to Sitagliptin Manufacture. *Science* **2010**, *329(5989)*, 305-9.

56. WHO (World Health Organization). *World Malaria Report 2005*. UNICEF: Geneva, 2005.
57. (a) Korenromp, E. L.; Williams, B. G.; Gouws, E.; Dye, C.; Snow, R. W. Measurement of trends in childhood malaria mortality in Africa: an assessment of progress toward targets based on verbal autopsy. *Lancet Infect. Dis.* **2003**, *3(6)*, 349-358; (b) Marsh, K. Malaria disaster in Africa. The *Lancet* **1998**, *352(9132)*, 924.
58. Enserink, M. Source of New Hope Against Malaria is in Short Supply. *Science* **2005**, *307(5706)*¸ 33.
59. Schmid, G.; Hofheinz, W. Total Synthesis of Qinghaosu. *J. Am. Chem. Soc.* **1983**, *105(3)*, 624-5.
60. (a) Haynes, R. K.; Vonwiller, S. C. Cyclic peroxyacetal lactone, lactol and ether compounds. U.S. Patent 5,420,299, May 30, 1995; (b) Roth, R. J.; Acton, N., A simple conversion of artemisinic acid into artemisinin. *J. Nat. Prod.* **1989**, *52(5)*, 1183-5.
61. Paddon, C. J.; Westfall, P. J.; Pitera, D. J.; Benjamin, K.; Fisher, K.; McPhee, D.; Leavell, M. D.; Tai, A.; Main, A.; Eng, D.; Polichuk, D. R.; Teoh, K. H.; Reed, D. W.; Treynor, T.; Lenihan, J.; Fleck, M.; Bajad, S.; Dang, G.; Dengrove, D.; Diola, D.; Dorin, G.; Ellens, K. W.; Fickes, S.; Galazzo, J.; Gaucher, S. P.; Geistlinger, T.; Henry, R.; Hepp, M.; Horning, T.; Iqbal, T.; Jiang, H.; Kizer, L.; Lieu, B.; Melis, D.; Moss, N.; Regentin, R.; Secrest, S.; Tsuruta, H.; Vazquez, R.; Westblade, L. F.; Xu, L.; Yu, M.; Zhang, Y.; Zhao, L.; Lievense, J.; Covello, P. S.; Keasling, J. D.; Reiling, K. K.; Renninger, N. S.; Newman, J. D. High-level semi-synthetic production of the potent antimalarial artemisinin. *Nature* **2013**, *496(7446)*, 528-32.
62. WHO Prequalificaion of Medicines Programme. *Acceptance Of Non-Plant-Derived-Artemisinin Offers Potential To Increase Access To Malaria Treatment* [online]. 2013. http://apps.who.int/prequal/info_press/documents/PQ_non-plant_derived_artemisinin_1.pdf (accessed December 12, 2014).
63. Marris, C. SciDeveNet. Synthetic biology's malaria promises could backfire [Online], 2013. http://www.scidev.net/global/biotechnology/opinion/synthetic-biology-s-malaria-promises-could-backfire.html (accessed January 5, 2015).
64. Rude, M. A.; Schirmer, A. New Microbial Fuels: A Biotech Perspective. *Curr. Opin. Microbiol.* **2009**, *12(3)*, 274-81.
65. (a) Buelter, T.; Meinhold, P.; Feldman, R.; Hawkins, A.; Bastian, S.; Arnold, F. H.; Urano, J. Engineered microorganisms capable of producing target compounds under anaerobic conditions. U.S. Pat. Appl. 0058532 A1, 2012; (b) Donaldson, G. K.; Eliot, A.; Flint, D.; Maggio-Hall, A.; Nagarajan, V. Fermentative production of four carbon alcohols. U.S. Pat. Appl. 0313206 A1, 2007.
66. Hong, K. K.; Nielsen, J. Metabolic engineering of Saccharomyces cerevisiae: a key cell factory platform for future biorefineries. *Cell. Mol. Life Sci.* **2012**, *69(16)*, 2671-90.
67. Donaldson, G. K.; Eliot, A.; Flint, D.; Maggio-Hall, A.; Nagarajan, V. Fermentative production of four carbon alcohols. U.S. U.S. Pat. Appl. 0092957 A1, 2007.
68. Festel, G.; Boles, E.; Weber, C.; Brat, D. Fermentative production of isobutanol with yeast. U.S. Patent 8,530,226 B2, September 10, 2013.
69. Knothe, G. Biodiesel and renewable diesel: A comparison. *Prog. Energy Combust. Sci.* **2010**, *36*, 364-73.
70. Trimbur, D.; Im, C.-S.; Dillon, H.; Day, A.; Franklin, S.; Coragliotti, A. Production of oil in microorganisms. U.S. Patent 8,889,401, November 18, 2014.
71. Burk, M. Personal Comments. Presented at Workshop on the Industrialization of Biology, May 28, 2014.
72. (a) Yim, H.; Haselbeck, R.; Niu, W.; Pujol-Baxley, C.; Burgard, A.; Boldt, J.; Khandurina, J.; Trawick, J. D.; Osterhout, R. E.; Stephen, R.; Estadilla, J.; Teisan, S.; Schreyer, H. B.; Andrae, S.; Yang, T. H.; Lee, S. Y.; Burk, M. J.; Van Dien, S. Metabolic engineering of

Copyright © National Academy of Sciences. All rights reserved.

Industrialization of Biology: A Roadmap to Accelerate the Advanced Manufacturing of Chemicals

Escherichia coli for direct production of 1,4-butanediol. *Nat. Chem. Biol.* **2011**, *7(7)*, 445-52; (b) Burk, M. J. Sustainable production of industrial chemicals from sugars. *Int. Sugar J.* **2010**, *112(1333)*, 30.

73. BCC Research. Global Markets for Enzymes in Industrial Applications; BCC Research: Wellesley, MA, 2014.
74. Scheufele, D. A. Communicating science in social settings. *Proc. Natl. Acad. Sci. U. S. A.* **2013**, *110(Supplement 3)*, 14040-7.
75. NIH (National Institutes of Health). Final NIH Genomic Data Sharing Policy. *Fed Regist*. **2014**, 79(167), 51345–54.
76. Werpy, T.; Petersen, G. Top Value Added Chemicals from Biomass: Volume I—Results of Screening for Potential Candidates from Sugars and Synthesis Gas; U.S. Department of Energy: Oak Ridge, TN, 2004.
77. Newman, D. J.; Cragg, G. M.; Snader, K. M. Natural products as sources of new drugs over the period 1981-2002. *J. Nat. Prod.* **2003**, *66(7)*, 1022-37.
78. (a) Draths, K. M.; Knop, D. R.; Frost, J. W. Shikimic acid and quinic acid: Replacing isolation from plant aources with recombinant microbial biocatalysis. *J. Am. Chem. Soc.* **1999**, *121(7)*, 1603-4; (b) Krämer, M.; Bongaerts, J.; Bovenberg, R.; Kremer, S.; Müller, U.; Orf, S.; Wubbolts, M.; Raeven, L. Metabolic engineering for microbial production of shikimic acid. *Metab. Eng.* **2003**, *5(4)*, 277-83.
79. (a) Pollard, D. J.; Woodley, J. M. Biocatalysis for pharmaceutical intermediates: the future is now. *Trends in Biotechnol.* **2007**, *25(2)*, 66-73; (b) Clouthier, C. M.; Pelletier, J. N. Expanding the organic toolbox: A guide to integrating biocatalysis in synthesis. *Chem. Soc. Rev.* **2012**, *41(4)*, 1585-605.
80. Savile, C. K.; Janey, J. M.; Mundorff, E. C.; Moore, J. C.; Tam, S.; Jarvis, W. R.; Colbeck, J. C.; Krebber, A.; Fleitz, F. J.; Brands, J.; Devine, P. N.; Huisman, G. W.; Hughes, G. J. Biocatalytic asymmetric synthesis of chiral amines from ketones applied to sitagliptin manufacture. *Science* **2010**, *329(5989)*, 305-9.
81. (a) Müller, K.; Faeh, C.; Diederich, F. Fluorine in pharmaceuticals: Looking beyond intuition. *Science* **2007**, *317(5846)*, 1881-6; (b) Purser, S.; Moore, P. R.; Swallow, S.; Gouverneur, V. Fluorine in medicinal chemistry. *Chem. Soc. Rev.* **2008**, *37(2)*, 320-30; (c) Eustaquio, A. S.; O'Hagan, D.; Moore, B. S. Engineering fluorometabolite production: Fluorinase expression in Salinispora tropica yields fluorosalinosporamide. *J. Nat. Prod.* **2010**, *73(3)*, 378-82; (d) Runguphan, W.; Qu, X.; O'Connor, S. E. Integrating carbon-halogen bond formation into medicinal plant metabolism. *Nature* **2010**, *468(7322)*, 461-4; (e) Walker, M. C.; Thuronyi, B. W.; Charkoudian, L. K.; Lowry, B.; Khosla, C.; Chang, M. C., Expanding the Fluorine Chemistry of Living Systems Using Engineered Polyketide Synthase Pathways. *Science* **2013**, *341(6150)*, 1089-94.
82. (a) Coelho, P. S.; Brustad, E. M.; Kannan, A.; Arnold, F. H. Olefin cyclopropanation via carbene transfer catalyzed by engineered cytochrome P450 enzymes. *Science* **2013**, *339(6117)*, 307-10; (b) McIntosh, J. A.; Coelho, P. S.; Farwell, C. C.; Wang, Z. J.; Lewis, J. C.; Brown, T. R.; Arnold, F. H. Enantioselective intramolecular C-H amination catalyzed by engineered cytochrome P450 enzymes in vitro and in vivo. *Angew. Chem., Int. Ed.* **2013**, *52(35)*, 9309-12.
83. Treimer, J. F.; Zenk, M. H. Purification and Properties of Strictosidine Synthase, the Key Enzyme in Indole Alkaloid Formation. *Eur. J. Biochem.* **1979**, *101(1)*, 225-33.
84. Kim, H. J.; Ruszczycky, M. W.; Choi, S. H.; Liu, Y. N.; Liu, H. W. Enzyme-catalysed [4+2] cycloaddition is a key step in the biosynthesis of spinosyn A. *Nature* **2011**, *473(7345)*, 109-12.
85. IREA (International Renewable Energy Agency). Production of bio-ethylene (Technology Brief I13); International Renewable Energy Agency and Energy Technology Systems Analysis Programme; IREA: Abu Dhabi, UAE, 2013.

Copyright © National Academy of Sciences. All rights reserved.

86. Yim, H.; Haselbeck, R.; Niu, W.; Pujol-Baxley, C.; Burgard, A.; Boldt, J.; Khandurina, J.; Trawick, J. D.; Osterhout, R. E.; Stephen, R.; Estadilla, J.; Teisan, S.; Schreyer, H. B.; Andrae, S.; Yang, T. H.; Lee, S. Y.; Burk, M. J.; Van Dien, S. Metabolic Engineering of Escherichia coli for Direct Production of 1,4-Butanediol. *Nat. Chem. Biol.* **2011**, *7(7)*, 445-52.
87. Tullo, A. H. Hunting for Biobased Acrylic Acid. *Chem. Eng. News* **2013**, *91(46)*, 18-9.
88. Madhavan Nampoothiri, K.; Nair, N. R.; John, R. P. An overview of the recent developments in polylactide (PLA) research. *Bioresour. Technol.* **2010**, *101(22)*, 8493-501.
89. Anderson, A. J.; Dawes, E. A. Occurrence, Metabolism, Metabolic Role, and Industrial Uses of Bacterial Polyhydroxyalkanoates. *Microbiol. Rev.* **1990**, *54(4)*, 450-72.
90. Zhang, S., Fabrication of Novel Biomaterials through Molecular Self-Assembly. *Nat. Biotechnol.* **2003**, *21(10)*, 1171-8.
91. (a) Fahnestock, S.; Rich, A. Ribosome-catalyzed polyester formation. *Science* **1971**, *173(3994)*, 340-3; (b) Mao, C.; Solis, D. J.; Reiss, B. D.; Kottmann, S. T.; Sweeney, R. Y.; Hayhurst, A.; Georgiou, G.; Iverson, B.; Belcher, A. M. Virus-based toolkit for the directed synthesis of magnetic and semiconducting nanowires. *Science* **2004**, *303(5655)*, 213-7; (c) Ohta, A.; Murakami, H.; Higashimura, E.; Suga, H. Synthesis of polyester by means of genetic code reprogramming. *Chemistry & Biology* **2007**, *14(12)*, 1315-22.
92. (a) Addadi, L.; Weiner, S. Interactions between acidic proteins and crystals: Stereochemical requirements in biomineralization. *Proc. Natl. Acad. Sci. U. S. A.* **1985**, *82(12)*, 4110-4; (b) Mann, S.; Archibald, D. D.; Didymus, J. M.; Douglas, T.; Heywood, B. R.; Meldrum, F. C.; Reeves, N. J. Crystallization at Inorganic-organic Interfaces: Biominerals and Biomimetic Synthesis. *Science* **1993**, *261(5126)*, 1286-92; (c) Belcher, A. M.; Wu, X. H.; Christensen, R. J.; Hansma, P. K.; Stucky, G. D.; Morse, D. E. Control of crystal phase switching and orientation by soluble mollusc-shell proteins. *Nature* **1996**, *381(6577)*, 56-8; (d) Banfield, J. F.; Welch, S. A.; Zhang, H.; Ebert, T. T.; Penn, R. L. Aggregation-Based Crystal Growth and Microstructure Development in Natural Iron Oxyhydroxide Biomineralization Products. *Science* **2000**, *289(5480)*, 751-4; (e) Sundar, V. C.; Yablon, A. D.; Grazul, J. L.; Ilan, M.; Aizenberg, J. Fibre-optical features of a glass sponge. *Nature* **2003**, *424(6951)*, 899-900.
93. Christensen, C. M.; Raynor, M. E. *The Innovator's Solution: Creating and Sustaining Successful Growth*. Harvard Business School Press: Boston, MA, 2003.
94. U.S. Energy Information Administration. U.S. Number of Operable Refineries as of January 1. http://www.eia.gov/dnav/pet/hist/LeafHandler.ashx?n=PET&s=8_NA_8O0_NUS_C&f=A (accessed July 25, 2014).
95. Carlson, R.; Wehbring, R. *Microbrewing the Bioeconomy: Innovation and Changing Scale in Industrial Production*. http://www.biodesic.com/library/Microbrewing_the_Bioeconomy.pdf (accessed January 5, 2015).
96. Kojima, M.; Johnson, T. Potential for biofuels for transport in developing countries. *ESMAP Knowledge Exchange Series* **2005**, *4*, 1-4.
97. Agricultural Marketing Resource Center. A National Information Resource for Value-Added Agriculture: Corn. http://www.agmrc.org/commodities__products/grains__oilseeds/corn_grain/ (accessed December 30, 2015).
98. (a) Fang, Z. Converting Lignocellulosic Biomass to Low Cost Fermentable Sugars. In *Pretreatment Techniques for Biofuels and Biorefineries;* Springer: Berlin, 2013; pp 133-150; (b) Beckman, E. J. Supercritical and near-critical CO 2 in green chemical synthesis and processing. *J. Supercrit. Fluids* **2004**, *28(2)*, 121-91.
99. Singh, R. K.; Tiwari, M. K.; Singh, R.; Lee, J.-K. From Protein Engineering to Immobilization: Promising Strategies for the Upgrade of Industrial Enzymes. *Int. J. Mol. Sci.* **2013**, *14(1)*, 1232-77.
100. He, M. Cell-free protein synthesis: applications in proteomics and biotechnology. *New Biotechnol.* **2008**, *25(2-3)*, 126-32.

Copyright © National Academy of Sciences. All rights reserved.

101. Rollin, J. A.; Tam, T. K.; Zhang, Y. H. P. New biotechnology paradigm: cell-free biosystems for biomanufacturing. *Green Chem.* **2013**, *15(7)*, 1708-19.
102. Vallino, J. J.; Stephanopoulos, G. Metabolic Flux Distributions in Corynebacterium Glutamicum During Growth and Lysine Overproduction. *Biotechnology and Bioengineering* **2000**, *67(6)*, 872-85.
103. (a) Bairoch, A. PROSITE: A Dictionary of Sites and Patterns in Proteins. *Nucleic Acids Res.* **1991**, *19(Supplemental)*, 2241-5; (b) Hulo, N.; Bairoch, A.; Bulliard, V.; Cerutti, L.; De Castro, E.; Langendijk-Genevaux, P. S.; Pagni, M.; Sigrist, C. J. A. The PROSITE database. *Nucleic Acids Res.* **2006**, *34(Supplemental)*, D227-30.
104. Xiong, Z.; Laird, P. W. COBRA: A Sensitive and Quantitative DNA Methylation Assay. *Nucleic Acids Res.* **1997**, *25(12)*, 2532-4.
105. Hillson, N. DNA Assembly Method Standardization for Synthetic Biomolecular Circuits and Systems. In *Design and Analysis of Biomolecular Circuits*, Koeppl, H.; Setti, G.; di Bernardo, M.; Densmore, D., Eds. Springer: New York, 2011; pp 295-314.
106. Onken, M.; Eichelberg, M.; Riesmeier, J.; Jensch, P. Digital Imaging and Communications in Medicine. In *Biomedical Image Processing*, Deserno, T. M., Ed. Springer: Berlin, 2011; pp 427-54.
107. (a) Canton, B.; Labno, A.; Endy, D. Refinement and standardization of synthetic biological parts and devices. *Nat. Biotechnol.* **2008**, *26(7)*, 787-93; (b) Brown, J. The iGEM competition: building with biology. *Synthetic Biology, IET* **2007**, *1(1.2)*, 3-6.
108. Ham, T. S.; Dmytriv, Z.; Plahar, H.; Chen, J.; Hillson, N. J.; Keasling, J. D. Design, implementation and practice of JBEI-ICE: an open source biological part registry platform and tools. *Nucleic Acids Res.* **2012**, 40(18), e141.
109. Cooling, M. T.; Rouilly, V.; Misirli, G.; Lawson, J.; Yu, T.; Hallinan, J.; Wipat, A. Standard virtual biological parts: a repository of modular modeling components for synthetic biology. *Bioinformatics* **2010**, *26(7)*, 925-31.
110. Seiler, C. Y.; Park, J. G.; Sharma, A.; Hunter, P.; Surapaneni, P.; Sedillo, C.; Field, J.; Algar, R.; Price, A.; Steel, J.; Throop, A.; Fiacco, M.; LaBaer, J. DNASU plasmid and PSI:Biology-Materials repositories: resources to accelerate biological research. *Nucleic Acids Res.* **2013**, *42(Database Issue)*, D1253-60.
111. Herscovitch, M.; Perkins, E.; Baltus, A.; Fan, M. Addgene provides an open forum for plasmid sharing. *Nat. Biotechnol.* **2012**, *30(4)*, 316-7.
112. (a) Eisenreich, W.; Bacher, A.; Arigoni, D.; Rohdich, F., Biosynthesis of isoprenoids via the non-mevalonate pathway. *Cell. Mol. Life Sci.* **2004**, *61(12)*, 1401-26; (b) Rohmer, M. The discovery of a mevalonate-independent pathway for isoprenoid biosynthesis in bacteria, algae and higher plants. *Natural Products Reports* **1999**, *16(5)*, 565-74.
113. Raab, A. M.; Lang, C. Oxidative versus reductive succinic acid production in the yeast Saccharomyces cerevisiae. *Bioengineered Bugs* **2011**, *2(2)*, 120-3.
114. Rahman, S. A.; Cuesta, S. M.; Furnham, N.; Holliday, G. L.; Thornton, J. M. EC-BLAST: a tool to automatically search and compare enzyme reactions. *Nat. Methods* **2014**, *11(2)*, 171-4.
115. Altschul, S. F.; Madden, T. L.; Schäffer, A. A.; Zhang, J.; Zhang, Z.; Miller, W.; Lipman, D. J. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.* **1997**, *25(17)*, 3389-402.
116. De Ferrari, L.; Mitchell, J. B. From sequence to enzyme mechanism using multi-label machine learning. *BMC Bioinformatics* **2014**, *15*, 150.
117. Esvelt, K. M.; Carlson, J. C.; Liu, D. R. A system for the continuous directed evolution of biomolecules. *Nature* **2011**, *472(7344)*, 499-503.
118. (a) Fox, R. J.; Davis, S. C.; Mundorff, E. C.; Newman, L. M.; Gavrilovic, V.; Ma, S. K.; Chung, L. M.; Ching, C.; Tam, S.; Muley, S.; Grate, J.; Gruber, J.; Whitman, J. C.; Sheldon, R. A.; Huisman, G. W. Improving catalytic function by ProSAR-driven enzyme evolu-

Copyright © National Academy of Sciences. All rights reserved.

tion. *Nature Biotechnol.* **2007**, *25(3)*, 338-44; (b) Luetz, S.; Giver, L.; Lalonde, J. Engineered enzymes for chemical production. *Biotechnology and Bioengineering* **2008**, *101(4)*, 647-53.

119. Adrio, J.-L.; Demain, A. L. Recombinant organisms for production of industrial products. *Bioengineered Bugs* **2010**, *1(2)*, 116-131.
120. (a) Niewoehner, O.; Jinek, M.; Doudna, J. A., Evolution of CRISPR RNA recognition and processing by Cas6 endonucleases. *Nucleic Acids Res.* **2014**, *42(2)*, 1341-53; (b) Gao, X.; Tsang, J. C. H.; Gaba, F.; Wu, D.; Lu, L.; Liu, P. Comparison of TALE designer transcription factors and the CRISPR/dCas9 in regulation of gene expression by targeting enhancers. *Nucleic Acids Res.* **2014**, *42(20)*, e155.
121. Heap, J. T.; Pennington, O. J.; Cartman, S. T.; Carter, G. P.; Minton, N. P. The ClosTron: a universal gene knock-out system for the genus Clostridium. *J. Microbiol. Methods* **2007**, *70(3)*, 452-64.
122. Joung, J. K.; Sander, J. D. TALENs: a widely applicable technology for targeted genome editing. *Nat. Rev. Mol. Cell Biol.* **2013**, *14(1)*, 49-55.
123. (a) Carroll, D. Genome Engineering With Zinc-Finger Nucleases. *Genetics* **2011**, *188(4)*, 773-782; (b) Guo, J.; Gaj, T.; Barbas, C. F. Directed evolution of an enhanced and highly efficient FokI cleavage domain for Zinc Finger Nucleases. *J. Mol. Biol.* **2010**, *400(1)*, 96-107; (c) Cathomen, T.; Keith Joung, J. Zinc-finger Nucleases: The Next Generation Emerges. *Mol. Ther.* **2008**, *16(7)*, 1200-7.
124. Wang, H. H.; Isaacs, F. J.; Carr, P. A.; Sun, Z. Z.; Xu, G.; Forest, C. R.; Church, G. M. Programming Cells by Multiplex Genome Engineering and Accelerated Evolution. *Nature* **2009**, *460(7257)*, 894-8.
125. National Renewable Energy Laboratory. *Biomass Research*. http://www.nrel.gov/biomass/biorefinery.html (accessed January 13, 2015).
126. (a) Schellenberger, J.; Que, R.; Fleming, R. M. T.; Thiele, I.; Orth, J. D.; Feist, A. M.; Zielinski, D. C.; Bordbar, A.; Lewis, N. E.; Rahmanian, S.; Kang, J.; Hyduke, D. R.; Palsson, B. O. Quantitative prediction of cellular metabolism with constraint-based models: the COBRA Toolbox v2.0. *Nat. Protoc.* **2011**, *6(9)*, 1290-307; (b) Ebrahim, A.; Lerman, J. A.; Palsson, B. O.; Hyduke, D. R. COBRApy: constraints-based reconstruction and analysis for python. *BMC Systems Biology* **2013**, *7(1)*, 74.
127. Galzie, Z. What is protein engineering? *Biochem. Educ.* **1991**, *19(2)*, 74-75.

Copyright © National Academy of Sciences. All rights reserved.

Copyright © National Academy of Sciences. All rights reserved.

Appendix A

Glossary

ABE Process—Acetone–butanol–ethanol (ABE) fermentation process. Acetone and butanol are produced from glucose using strains of *Clostridia*.

Act Ontology—A formal ontology that describes the molecular function of any entity participating in a biochemical reaction. Act Ontology provides a formal description of the species' chemical behavior according to a controlled vocabulary to support querying, synthesis, and verification.[45]

APHIS—USDA Animal and Plant Inspection Service.

BDO—1,4-butanediol.

Bifunctional—A molecule or compound that has properties of two different types of functional groups.

Biocatalysis—The use of natural catalysts to perform chemical transformations on organic compounds.

Bioeconomy—The portion of the economy that is derived from biological processes and manufacturing.

Bioinformatics—The science of collecting and analyzing complex biological data.

Copyright © National Academy of Sciences. All rights reserved.

BIOFAB—International Open Facility Advancing Biotechnology.

Biomanufacturing—The production of biology-based chemicals and products.

Biorefinery—A biorefinery is a facility that integrates biomass conversion processes and equipment to produce fuels, power, and chemicals from biomass.[125]

BMBL—*Biosafety in Microbiological and Medical Laboratories.* A publication of the Centers for Disease Control and Prevention's Office of Safety, Health, and Environment and the National Institutes of Health.

BNICE—Biochemical Network Integrated Computational Explorer. A framework for identification and thermodynamic assessment of all possible pathways for the degradation or production of a given compound.

Biotechnology—The use of living cells, bacteria, etc., to make useful products.[15]

COBRA—COnstraint-Based Reconstruction and Analysis. It is a leading software package for genome-scale analysis of metabolism.[126]

CRISPR/Cas9—Clustered Regularly Interspaced Short Palindromic Repeats and CRISPR-associated protein 9.

DNA—Deoxyribonucleic Acid.

Domesticated Microorganism—The domestication of an organism to be a suitable chassis in industrial biotechnology.

Enzyme—Typically proteins, enzymes are macromolecular biological catalysts.

EPA—U.S. Environmental Protection Agency.

FDA—U.S. Food and Drug Administration.

Feedstock—The starting material used in the manufacturing process. This may be a form of biomass, a crude or refined petroleum hydrocarbon product, or a material that has already been chemically modified in some way.

Copyright © National Academy of Sciences. All rights reserved.

Fermentation—A metabolic process that converts sugar into a product.

Continuous—A fermentation process in which nutrients are added to the bioreactor continuously and product is continuously removed from the bioreactor.

Fed Batch—A fermentation processes in which nutrients are added to the bioreactor during cultivation with the product remaining in the bioreactor.

FIFRA—Federal Insecticide, Fungicide, and Rodenticide Act. The act provides federal government control of pesticide distribution, sale, and use.

Financial Instruments—Forms of research and innovation financing asset classes and financing mechanisms.

Genetic Engineering—The process of manually adding new DNA to an organism, typically with the goal of expressing one or more traits not already found in that organism.

Horizontally Stratified Development—A stratified industry for process development in which different companies specialize in different steps along the supply or value chain.

MAGE—Multiplex Automated Genomic Engineering. MAGE simultaneously targets many locations on the chromosome for modification in a single cell or across a population of cells.[124]

Metabolic Engineering—Optimizing genetic and regulatory processes within cells to increase the production of a desired product.

Metrology—The science of measurement.

Monomer—A molecule that may bind chemically to other molecules to form a polymer.

NIH—U.S. National Institutes of Health.

OSHA—U.S. Occupational Health and Safety Administration.

PDO—1,3-propanediol.

Copyright © National Academy of Sciences. All rights reserved.

PHA—Polyhydroxyalkanoates. Polyesters produced by bacterial fermentation of sugar or lipids.

Polymer—A large molecule, or macromolecule, composed of many repeated subunits.

PLA—Polylactic acid. A biodegradable polyester derived from biological feedstocks.

Polymerase—An enzyme that synthesizes long chains or polymers of nucleic acids.

Protein Engineering—The introduction of practical improvements into proteins.[127]

Rational Design—A design strategy that takes into consideration the capabilities available in science and engineering, as well as possible chemical transformations that will lead to a product of choice.

RNA—Ribonucleic Acid.

Synthetic Biology—A field that applies engineering principles to reduce genetics into DNA "parts" and understand how they can be combined to build desired functions in living cells.

Systems Biology—The study of systems of biological units.

TAG—triacyl glyceride.

Transformation—The conversion of a substrate to a product.

TSCA—Toxic Substances Control Act. Provides the federal government with authority to require reporting, record-keeping and testing requirements, and restrictions relating to chemical substances and/or mixtures.

USDA—U.S. Department of Agriculture.

Vertically Integrated Development—Process research and development is performed by vertically integrated corporations that develop the entire processes from end to end.

WHO—World Health Organization.

Copyright © National Academy of Sciences. All rights reserved.

Appendix B

The Current Regulatory Framework

ENGINEERED MICROORGANISMS AND THE CHEMICALS THEY PRODUCE

The Environmental Protection Agency (EPA) regulates entities that produce and commercialize new chemicals under the Toxic Substances Control Act of 1976 (TSCA). Under this Act, the EPA also regulates the commercial research and development, manufacturing, importing, and processing of "inter-generic microorganisms" intended for commercialization. An inter-generic microorganism is one "formed by the deliberate combination of genetic material originally isolated from organisms of different taxonomic genera." The EPA has reviewed applications for commercialization of approximately 75 microorganisms since 1998, and it construes the regulatory definition to include organisms created through synthetic biology. However, commentators have noted that microorganisms made using synthetic DNA sequences not found in any existing organism might not fall within the EPA's definition. In the next 5 to 10 years, however, it is unlikely that the DNA sequences used to engineer organisms for industrial chemical production will use completely *de novo* DNA sequences, and thus the regulatory definition is probably sufficient for this stage of industrial biology.

At least 90 days prior to manufacturing, importing, or processing a new, inter-generic microorganism for commercial purposes, the responsible firm must submit a complete Microbial Commercial Activity Notice (MCAN), or an exemption request, to the EPA for review. The MCAN must contain test data the manufacturer possesses or controls, and information

Copyright © National Academy of Sciences. All rights reserved.

from the scientific literature, about the microorganism's effects on humans (including worker exposure), animals, plants, and other microorganisms. In addition, the MCAN must include information about the identity of the organism, genetic manipulations used to construct it, its properties and phenotype, the traits that have been selected for or modified, the by-products it produces, and its proposed uses and environmental releases.

The EPA's review of an MCAN should determine whether the microorganism, under the proposed conditions of use, is reasonably safe for humans and the environment. After the 90-day review period has expired, a firm can manufacture, import, or process the microorganism if the EPA has not taken regulatory action to prevent or constrain such activities. If the agency determines that a chemical or microorganism presents an unreasonable risk of injury to health or the environment, then it can prohibit or limit the manufacture, distribution, or processing of the chemical or microorganism by writing a new regulation.

Review of an exemption request will determine whether the firm's commercial use of the new, inter-generic microorganism is exempt from TSCA requirements. Exemptions are granted for use of microorganisms, genetic manipulations, and production processes with which the agency has long experience and for which there is a record of safety.

If the new chemical will be produced by a microorganism intended for release into the environment, such as engineered algae grown in an open-pond system, then the manufacturer would have to conduct field tests prior to submitting the MCAN. Field tests of microorganisms present more uncertainty about health and environmental impacts than field tests of chemicals, because microorganisms can replicate and might proliferate beyond the immediate test site or might transfer genes to related organisms in the wild. One cannot as easily control the quantity of a microorganism released, or counteract problems should they occur.

To field test a microorganism intended for commercial use, the manufacturer must submit a TSCA Experimental Release Application (TERA) to the EPA at least 60 days before any field test could commence, unless the proposed testing is eligible for an exemption. The TERA must be approved, with or without conditions, before testing begins. Whoever conducts the research (manufacturer or a contractor) must comply with all terms and conditions in the TERA. The EPA only approves a TERA if it determines the proposed research "does not present an unreasonable risk of injury to health or the environment." The EPA can also revoke or modify a TERA if it receives new information concerning research risks. Data from field tests conducted under the TERA must be included in the MCAN if the tests are completed during the MCAN review period.

There is some question regarding whether the EPA can adequately review TERA applications for release of organisms created through syn-

Copyright © National Academy of Sciences. All rights reserved.

thetic biology approaches within the 60-day time frame. Synthetic biologists cannot always predict the effects of complex combinations of synthesized DNA on the organisms into which they are engineered, and the agency lacks models for assessing the health and environmental risks of such organisms to determine whether they are likely safe enough to release in a field test. As scientists and the agency become more experienced, some combinations of DNA segments in some organisms will become more predictable and regulatory risk-prediction models will be developed.

A second concern about precommercial testing of microorganisms has to do with whether the EPA has developed adequate insights into and guidance concerning containment for microbial organisms tested "inside a structure" (not intentionally released into the environment). Manufacturers conducting tests inside a structure need not report to the EPA if they meet certain criteria, although they must keep specified records. The EPA does not require that such interior testing meet the National Institutes of Health's (NIH's) Guidelines for Research Involving Recombinant or Synthetic Nucleic Acid Molecules, although it gives exemptions from EPA oversight to research conducted with federal funding from another agency and that is required to meet the NIH Guidelines. Researchers and administrators from several agencies have extensive experience applying the NIH Guidelines, and industry might be well served to voluntarily adopt these guidelines for testing of industrial microorganisms in contained spaces.

Another concern about the EPA's authority under the TSCA and its implementing regulations is that the agency's ability to require safety testing of engineered organisms (or new chemicals) requires substantial administrative process and might not be used effectively. The submitter of the MCAN must only provide information within her possession or control; the firm generally need not conduct research to generate new information. If the EPA believes information in the MCAN is inadequate for judging safety, then it has some authority under sections 4 (the "test rule") and 5(e) of TSCA to require that a manufacturer conduct and report safety research (such as generating and reporting toxicity data). Usually, firms and the EPA negotiate an agreement regarding the data to be generated, and the firm voluntarily signs a Consent Order. However, if the manufacturer does not voluntarily agree, then the agency must write a new regulation to require testing. Such a procedure can take several years and can be challenged in court.

A third concern is whether the EPA has sufficient postmarket authority to regulate industrial microorganisms. The EPA can regulate intergeneric microorganisms that produce chemicals even after the MCAN 90-day review period has expired and the product has entered the market;

Copyright © National Academy of Sciences. All rights reserved.

however, the agency must have evidence of harm to health or the environment before it may take action. The TSCA and its implementing regulations require manufacturers and distributors to keep records of specified negative health or environmental effects, but the regulations do not require that such information be reported to the EPA unless the agency explicitly requested reports, and the regulations do not require that firms conduct studies to identify problems. In the event of an accidental environmental release of an inter-generic microorganism used to manufacture chemicals, the firm responsible would have to keep records of health and environmental effects of the "spill," but TSCA does not require that a firm limit or mitigate those effects.

The EPA could have jurisdiction over some types of accidental microbial release under other laws (beyond TSCA). For instance, if algae used for industrial production of a chemical, such as a plastic, were to escape containment and get into natural waterways or lakes, the EPA might have jurisdiction under the Clean Water Act. However, in many instances people affected by an accidental release of an industrial microorganism might have to bring a tort action against the manufacturer, distributor, or other relevant party to obtain environmental cleanup and monitoring.

In the event of an accidental or uncontrolled harmful release of industrial microorganisms, the EPA could act under section 6 of TSCA to impose new containment conditions to prevent similar incidents in the future. However, to take such an action the agency must make a finding, based on risk-benefit calculations, that the microorganism as manufactured and distributed prior to the release posed an unreasonable risk, and that the regulatory action proposed was the least burdensome regulation that could provide adequate protection. The courts have ensured that the evidentiary burden for such findings is high, and the EPA had only issued "section 6 rules" for five chemicals as of 2005.

EPA officials believe they have adequate authority under TSCA to protect the environment, public health, and worker safety by imposing conditions on the manufacture and use of engineered microorganisms. The agency can prohibit or limit the use of new chemicals or engineered microorganisms that pose unreasonable risks to health or the environment, it can enjoin the manufacturing or processing of a commercial microorganism, and it can impose criminal or civil penalties on manufacturers who do not comply with TSCA. The EPA has been creative in using its premarket authority under TSCA to protect health and the environment and has developed cooperative relationships with chemical producers. It has used Consent Orders to impose conditions on particular manufacturers of microorganisms, including conditions for containment and worker safety, and it has used Significant New Use Rules (SNURs) to impose industry-wide standards. However, as with other regulations, a

Copyright © National Academy of Sciences. All rights reserved.

SNUR can take years to promulgate. In 2012, the EPA proposed a SNUR for the fungus *Trichoderma reesei*, inter-generic versions of which are used to make enzymes for ethanol production. Agency officials were concerned that the microbe could produce toxic peptides under some growth conditions and that it might not be properly contained. As of this report's writing, the *Trichoderma* SNUR has not been finalized.

A final concern about the EPA's regulation of industrial microbes is that the growth of industrial biology could result in a flood of MCANs and TERAs, which might overwhelm the agency. The agency has reviewed approximately 75 engineered microbes since 1998, a very limited number when compared, for instance, to the thousands of decisions the U.S. Department of Agriculture (USDA) has made over that same time period regarding the field testing and release of engineered crops. If the EPA lacks sufficient staff or resources, then the quality of its reviews could suffer or it could become a bottleneck in the pathway to market. At some point, the EPA might have to reallocate personnel and resources toward review of biological production of new chemicals, or it may need additional resources to carry out such reviews.

CROPS AS BIOREACTORS

If crops are used as bioreactors to produce industrial chemicals, then the USDA Animal and Plant Health Inspection Service (APHIS) is usually the lead agency with jurisdiction, which it could share with the EPA or the Food and Drug Administration (FDA), depending on the chemical's intended uses. Under the Plant Protection Act, APHIS regulates the importation, interstate movement, and environmental release of "plant pests," including genetically engineered organisms that might pose a risk to plant health. To date, most transgenic plants intended for commercial use have been modified using vectors or genes from plant pests, and thus APHIS has had jurisdiction over nearly all of them.

If a new, genetically engineered organism (including a plant) is known or suspected to cause damage or disease to a plant or plant product, then it cannot be introduced into the environment without a field trial conducted under APHIS's authorization. To conduct such a trial, the plant's developer must file either a notification or permit application with APHIS. Prior to approving the field release of regulated material, APHIS will conduct a review of the notification or application to ensure that under the proposed conditions of use—handling, confinement, and disposal—the risks to plant health and the environment have been appropriately minimized. Permits are generally more restrictive than notifications and are used for types of plants that pose heightened risk to plant health or the environment, or for plants with novel modifications whose risks are

Copyright © National Academy of Sciences. All rights reserved.

uncertain and with which the agency has less regulatory experience. Notification is used for plants that pose low risk and contain modifications with which the agency has had familiarity. APHIS typically uses permits for plants that produce biopharmaceuticals and industrial chemicals. Sites where a plant-produced biopharmaceutical or industrial chemical is being field tested will be inspected several times during the trial.

Once a new, genetically modified plant has been field tested the developer can file a petition for deregulation, which must include information about the plant's biology (including genetic modifications) and the field test results.

In evaluating such petitions, APHIS considers numerous factors, including "the expression of gene products, new enzymes, or changes to plant metabolism; weediness and impact on sexually compatible plants; agricultural or cultivation practices; effects on non-target organisms; and the potential for gene-transfer to other types of organisms." In addition to assessing the petition, the agency will prepare an environmental assessment or an environmental impact statement and seek public comment on the plant's risks. If APHIS reaches the conclusion that the new plant does not pose a plant-pest risk, it will approve the petition. Once the organism has been deregulated, it can be introduced into fields, and into commerce, without APHIS oversight. As of August 2013, APHIS had overseen the deregulation of 95 genetically engineered crops.

Some new, engineered organisms can be deregulated using a "request for extension of non-regulated status." This process was established in 1997 and assumes that, from a safety standpoint, many regulated organisms will have only negligible differences from previously deregulated ones. In the extension request, the petitioner compares the regulated organism to an antecedent, deregulated organism to show that the molecular manipulation used to make the new organism raises no serious, new risks to plants or the environment. Extensions are used for interventions such as synonymous nucleotide changes (ones that do not change the amino acid sequence of the encoded protein) and may be less often available to plants engineered using synthetic biology approaches.

When both APHIS and the FDA have jurisdiction over a genetically engineered plant, for instance, if the commercial aim is production of a plant-made pharmaceutical, then APHIS takes the lead in regulating premarket field tests of the engineered plant and the FDA would later subject the plant-made chemical to the premarket approval process typically used for drugs. And, APHIS could share jurisdiction with the EPA if, for example, the engineered organism was a known bacterial plant pest.

APHIS only has jurisdiction over genetically modified plants or other organisms under the Plant Protection Act if there is reason to believe that the engineered organism would harm plants. APHIS does not have broad

Copyright © National Academy of Sciences. All rights reserved.

jurisdiction to consider possible environmental or health hazards. Thus, there is a category of engineered plants whose field tests or commercial use would not fall under APHIS's jurisdiction, and that the EPA also does not have clear authority or experience to regulate. If such plants were producing industrial chemicals not intended for a FDA-regulated purpose, then their release into the environment and their commercial distribution would be completely unregulated. Switchgrass engineered for optimal use as a feedstock in biofuel production is an example of a genetically modified plant that did not fall within APHIS's regulatory authority and is otherwise unregulated.

Copyright © National Academy of Sciences. All rights reserved.

Copyright © National Academy of Sciences. All rights reserved.

Appendix C

Committee Member and Staff Biographies

CHAIR

Thomas M. Connelly, Jr. is the Executive Director and CEO of the American Chemical Society. Dr. Connelly retired from DuPont in December 2014, where he was Executive Vice President, Chief Innovation Officer, and a member of the company's Office of the Chief Executive. At DuPont, he was responsible for Science & Technology and the geographic regions outside the United States, as well as Integrated Operations which includes Operations, Sourcing & Logistics and Engineering. At DuPont, Dr. Connelly led businesses and R&D organizations, while based in the US, Europe and Asia. Dr. Connelly graduated with highest honors from Princeton University with degrees in Chemical Engineering and Economics. As a Winston Churchill Scholar, he received his doctorate in chemical engineering from the University of Cambridge. He is a Director of Grasim Industries, an Indian listed company. He has served in advisory roles to the U.S. government and the Republic of Singapore.

MEMBERS

Michelle Chang is associate professor of chemistry at the University of California (UC), Berkeley. Her research applies the approaches of mechanistic biochemistry, molecular and cell biology, metabolic engineering, and synthetic biology to address problems in energy and human health. Among her projects are the design and creation of new biosynthetic pathways in microbial hosts for *in vivo* production of biofuels from abundant

Copyright © National Academy of Sciences. All rights reserved.

crop feedstocks and pharmaceuticals from natural products or natural product scaffolds. She earned a Ph.D. from the Massachusetts Institute of Technology in 2004, did postdoctoral work at UC Berkeley, and joined its faculty in 2007. She received the Arnold and Mabel Beckman Young Investigator Award in 2008 and the Agilent Early Career Award in 2010.

Lionel Clarke is Co-Chair of the U.K. Synthetic Biology Leadership Council. He is also Team Leader, Biodomain Open Innovation for Shell Projects and Technology at the Shell Technology Centre, Thornton, United Kingdom. In this role he is responsible for planning and delivery of Shell strategic research and technology programs across the biodomain, deploying internal and external resources to deliver innovative solutions to market. Prior to joining Shell in 1981, Clarke graduated from Imperial College, London, after which he studied as an elected University Research Fellow at Cambridge University, and as a Royal Society European Research Fellow at the University of Grenoble, France. During this period he published numerous papers and a book and received various publication awards. Within Shell he has worked extensively, taking ideas from lab to market at the interface between fuels and engines, including the worldwide removal and replacement of leaded gasoline and the introduction of cleaner and improved performance fuels in developed and developing markets. Working with the Brazilian fuels market for a number of years gave him early first-hand experience of the potential, as well as practical issues, associated with the use of biofuels. Clarke has been responsible for facilitating the planning and delivery of strategic research programs across the biodomain within Shell for more than 10 years. Clarke chaired the U.K. Synthetic Biology Roadmap coordination group during 2012 and is now Co-Chairman of the U.K. Synthetic Biology Leadership Council.

Andrew Ellington received his B.S. in biochemistry from Michigan State University in 1981 and his Ph.D. in biochemistry and molecular biology from Harvard in 1988. As a graduate student he worked with Dr. Steve Benner on the evolutionary optimization of dehydrogenase isozymes. His postdoctoral work was with Dr. Jack Szostak at Massachusetts General Hospital, where he developed methods for the *in vitro* selection of functional nucleic acids and coined the term "aptamer." Ellington began his academic career as an assistant professor of chemistry at Indiana University in 1992 and continued to develop selection methods. He has previously received the Office of Naval Research Young Investigator, Cottrell, and Pew Scholar awards. In 1998 he moved to the University of Texas at Austin and is now the Fraser Professor of Biochemistry in the Department of Molecular Biosciences. Ellington was a member of the Defense Sci-

Copyright © National Academy of Sciences. All rights reserved.

ence Studies Group of the Institute for Defense Analysis and has actively advised numerous government agencies on biodefense and biotechnology issues, including serving on the BioChem 2020 panel of the Defense Intelligence Agency. Most recently he was named a National Security Science and Engineering Faculty Fellow and a Fellow of the American Association for the Advancement of Science (AAAS). He has served on the boards of numerous companies and helped found the aptamer company Archemix. Ellington's lab work centers on developing nucleic acid circuitry for point-of-care diagnostics and on accelerating the evolution of proteins and cells through the introduction of novel chemistries.

Nathan Hillson is a biochemist staff scientist at Lawrence Berkeley National Laboratory (LBNL), Director of Synthetic Biology at the Joint Bioenergy Institute, and Program Lead of Genome Engineering at the Joint Genome Institute. His responsibilities are to develop and demonstrate experimental wetware, software, and laboratory automation devices that facilitate, accelerate, and standardize the engineering of microbes. He earned a Ph.D. in biophysics from Harvard Medical School and was a postdoctoral research fellow at the Stanford University School of Medicine. He joined LBNL in 2009.

Richard Johnson is the CEO and founder of Global Helix LLC, a thought leadership and strategic positioning consulting firm based in Washington, D.C. Johnson has worked extensively on the linkage of global scientific developments, law, and policy with fundamental research, innovation, and entrepreneurship. After 30 years, he retired as Senior Partner in Arnold & Porter LLP, where he represented many research universities, foundations, and innovative companies. His current interests include (1) synthetic biology and the engineering of biology to enable bio-economic growth; (2) neuroscience and brain health, especially Alzheimer's and dementia; (3) global research collaborations and Big Data; (4) intellectual assets for value creation; and (5) rethinking organizational models for innovation policy and knowledge-based capital. Johnson is a member of the Board on Life Sciences at the National Academy of Sciences and the NAS Synthetic Biology Forum. He serves as the Chairman of the OECD/BIAC Technology and Innovation Committee and, recently, was named one of the 12 global members of the new OECD Global Advisory Council for Science, Technology, and Innovation. In addition, Johnson is the Chairman of Brown's Biology & Medicine Council and the International Council of the Innovation Knowledge Centre at Imperial College (London). He also is a member of boards for UC-Berkeley SynBerc; the BioBricks Foundation and Stanford BioFab, Brown Institute of Brain Sciences and BRAIN initiatives; and the INCF at the Karolinska Institute. For many years, he

Copyright © National Academy of Sciences. All rights reserved.

served on the MIT Corporation Committee and numerous university visiting committees. Johnson received his Juris Doctor degree from the Yale Law School where he was Editor of the *Yale Law Journal*, his M.S. from the Massachusetts Institute of Technology where he was a National Science Foundation National Fellow, and his undergraduate degree with highest honors from Brown University.

Jay D. Keasling is a professor of chemical engineering and bioengineering at the University of California, Berkeley. He is also Acting Deputy Laboratory Director of the Lawrence Berkeley National Laboratory, the Founding Director of the Synthetic Biology Department at UC Berkeley, and chief executive officer of the Joint BioEnergy Institute. He co-founded Codon Devices Inc. in 2004 and Amyris, Inc. (formerly Amyris Biotechnologies, Inc.) in 2003. He is considered one of the foremost authorities in synthetic biology, especially in the field of metabolic engineering. Other, related research interests include systems biology and environmental biotechnology. Keasling's current research involves the metabolic engineering of the *Escherichia coli* bacterium to produce the antimalarial drug artemisinin. Although it is an effective, proven treatment for malaria, current methods of producing artemisinin (found naturally in the plant *Artemisia annua*) are considered too expensive to cost-effectively eliminate malaria from developing countries. Keasling received his bachelor's degree at the University of Nebraska-Lincoln. He received his Ph.D. from the University of Michigan in 1991. He did postdoctoral work at Stanford University in biochemistry from 1991 to 1992.

Stephen Laderman is Director, Agilent Laboratories. He directs R&D programs aimed at inventing and developing leading-edge measurement solutions for research and diagnostics. His lab applies biology, chemistry, and computer science expertise to the investigation and development of novel reagents, assay protocols, and computational methods that enable new methods in emerging fields within molecular cellular biology, molecular medicine, and synthetic biology. After receiving his A.B. from Wesleyan University in physics and his Ph.D. from Stanford University in materials science and engineering, Laderman joined Hewlett-Packard (HP) Laboratories in 1984 as a member of the technical staff, subsequently holding a variety of research and management positions there and in technology-intensive businesses.

Pilar Ossorio is professor of law and bioethics at the University of Wisconsin, Madison (UW), where she is on the faculties of the Law School and the Department of Medical History and Bioethics at the Medical School. In 2011, she became the inaugural Ethics Scholar-in-Residence at

Copyright © National Academy of Sciences. All rights reserved.

the Morgridge Institute for Research, the private, nonprofit research institute that is part of the Wisconsin Institutes of Discovery. She also serves as the co-director of UW's Law and Neuroscience Program, as a faculty member in the UW Masters in Biotechnology Studies program, and as program faculty in the Graduate Program in Population Health. Prior to taking her position at UW, she was Director of the Genetics Section of the Institute for Ethics at the American Medical Association and taught as adjunct faculty at the University of Chicago Law School. Ossorio received her Ph.D. in microbiology and immunology in 1990 from Stanford University. She went on to complete a postdoctoral fellowship in cell biology at Yale University School of Medicine. Throughout the 1990s, Ossorio also worked as a consultant for the federal program on the Ethical, Legal, and Social Implications (ELSI) of the Human Genome Project, and in 1994 she took a full-time position with the Department of Energy's ELSI program. In 1993 she served on the Ethics Working Group for President Clinton's Health Care Reform Task Force. She received her J.D. from the University of California at Berkeley School of Law in 1997. While at Berkeley, she was elected to the legal honor society Order of the Coif and received several awards for outstanding legal scholarship.

Kristala Jones Prather is the Theodore T. Miller Associate Professor of Chemical Engineering at Massachusetts Institute of Technology (MIT), and an investigator in the multi-institutional Synthetic Biology Engineering Research Center (SynBERC) funded by the National Science Foundation. She received an S.B. degree from MIT in 1994 and Ph.D. from the University of California, Berkeley (1999) and worked 4 years in BioProcess Research and Development at the Merck Research Labs (Rahway, New Jersey) prior to joining the faculty of MIT. Her research interests are centered on the design and assembly of recombinant microorganisms for the production of small molecules, with additional efforts in novel bioprocess design approaches. Research combines the traditions of metabolic engineering with the practices of biocatalysis to expand and optimize the biosynthetic capacity of microbial systems. A particular focus is the elucidation of design principles for the production of unnatural organic compounds within the framework of the burgeoning field of synthetic biology. Prather is the recipient of a Camille and Henry Dreyfus Foundation New Faculty Award (2004), an Office of Naval Research Young Investigator Award (2005), a Technology Review "TR35" Young Innovator Award (2007), a National Science Foundation CAREER Award (2010), and the *Biochemical Engineering Journal* Young Investigator Award (2011). Additional honors include selection as the Van Ness Lecturer at Rensselaer Polytechnic Institute (2012) and a Young Scientist of the World Economic Forum Annual Meeting of the New Champions (2012). Prather has been

Copyright © National Academy of Sciences. All rights reserved.

recognized for excellence in teaching with the C. Michael Mohr Outstanding Faculty Award for Undergraduate Teaching in the Department of Chemical Engineering (2006) and the MIT School of Engineering Junior Bose Award for Excellence in Teaching (2010).

Reshma Shetty graduated from Massachusetts Institute of Technology with a Ph.D. in biological engineering in 2008, during which she worked on building digital logic in cells. Shetty has been active in synthetic biology for several years and co-organized SB1.0, the first international conference in synthetic biology in 2004. In 2008, *Forbes* magazine named Shetty one of Eight People Inventing the Future, and in 2011, *Fast Company* named her one of 100 Most Creative People in Business. Shetty and colleagues have founded synthetic biology company Ginkgo Bioworks, Inc., which makes and sells engineered microorganisms for food, fuels, and pharmaceuticals production.

Christopher Voigt is an associate professor in the Department of Biological Engineering at the Massachusetts Institute of Technology. He holds a joint appointment as a chemist scientist at Lawrence Berkeley National Laboratory, is an adjunct professor of chemical engineering at the Korea Advanced Institute of Science and Technology (KAIST), and an Honorary Fellow at Imperial College. Prior to joining MIT, he received his B.SE in chemical engineering from the University of Michigan (1998), a Ph.D. in biochemistry/biophysics at the California Institute of Technology (2002), performed postdoctoral work in the Bioengineering Department of the University of California, Berkeley (2003), and was a faculty member in the Department of Pharmaceutical Chemistry at the University of California, San Francisco (2003-2011).

Huimin Zhao is the Centennial Endowed Chair Professor of chemical and biomolecular engineering, and professor of chemistry, biochemistry, biophysics, and bioengineering at the University of Illinois at Urbana-Champaign and the visiting principal investigator of the Metabollic Engineering Research Laboratory (MERL) in the Agency for Science, Technology and Research (A*STAR) of Singapore. Prior to joining the faculty at the University of Illinois in 2000, Zhao worked at the Dow Chemical company for 2 years. Zhao's primary research interests are in the development and application of synthesis biology tools to address society's most daunting challenges in health, energy, and sustainability, and in the fundamental aspects of enzyme catalysis, cell metabolism, and gene regulation. Zhao has received numerous research and teaching awards and honors, including a Guggenheim Fellowship (2012), Fellow of the American Association for the Advancement of Science (AAAS) (2010),

Copyright © National Academy of Sciences. All rights reserved.

Fellow of the American Institute of Medical and Biological Engineering (2009), and others. He has authored and co-authored 170 research articles and 20 issued and pending patent applications, several of which are being licensed by industry. In addition, he has given plenary, keynote, or invited lectures in more than 200 international meetings and institutions. Zhao received his B.S. in biology from the University of Science and Technology of China in 1992. He earned a Ph.D. in chemistry from the California Institute of Technology in 1998.

NATIONAL RESEARCH COUNCIL STAFF

Douglas Friedman is a Senior Program Officer with the Board on Chemical Sciences and Technology at the National Research Council (NRC) of the National Academy of Sciences in Washington, DC. His primary scientific interests lie in the fields of organic chemistry, organic and bio-organic materials, chemical and biological sensing, and nanotechnology, particularly as they apply to national and homeland security. Friedman has supported a diverse array of activities since joining the NRC. He served as study director or co-study director on *Transforming Glycoscience: A Roadmap for the Future, Determining Core Capabilities in Chemical and Biological Defense Science and Technology, Effects of Diluted Bitumen on Crude Oil Transmission Pipelines, and Responding to Capability Surprise: A Strategy for U.S. Naval Forces*. Additionally, he has supported activities on *The Role of the Chemical Sciences in Finding Alternatives to Critical Resources, Opportunities and Obstacles in Large-Scale Biomass Utilization*, and *Technological Challenges in Antibiotics Discovery and Development.* Friedman is currently supporting studies on safety culture in academic research laboratories, security implications of advancing technologies in the life sciences, and synthetic biology. Prior to joining the NRC Friedman performed research in physical organic chemistry and chemical biology at Northwestern University; the University of California, Los Angeles; the University of California, Berkeley; and Solulink Biosciences. He received a Ph.D. in chemistry from Northwestern University and a B.S. in chemical biology from the University of California, Berkeley.

India Hook-Barnard came to the National Academies in 2008 from the National Institutes of Health where she was a postdoctoral research fellow. She earned her Ph.D. in microbiology-medicine from the Department of Molecular Microbiology and Immunology at the University of Missouri. Her primary interests are in the areas of emerging science, technology, and medicine, from fundamental research to translational application. While at the National Academies, Hook-Barnard was first a senior program officer with the National Academy of Sciences' Board on

Copyright © National Academy of Sciences. All rights reserved.

Life Sciences and most recently with the Institute of Medicine's Board on Health Sciences Policy. Hook-Barnard was study director for the consensus reports: *Sharing Clinical Trial Data: Maximizing Benefits, Minimizing Risk (2015), Determining Core Capabilities in Chemical and Biological Defense Science and Technology (2012), Toward Precision Medicine: Building a Knowledge Network for Biomedical Research and a New Taxonomy of Disease (2011), Sequence-Based Classification of Select Agents: A Brighter Line (2010),* and the workshop summary, *Technologies to Enable Autonomous Detection for BioWatch: Ensuring Timely and Accurate Information for Public Health (2013).* She directed the U.S. National Committee to the International Brain Research Organization from 2008 to 2012 and has served as staff officer for multiple activities, including the Standing Committee on the Department of Defense's Programs to Counter Biological Threats; the workshop *Convergence: Safeguarding Technology in the Bioeconomy;* the Six Party Symposia on Synthetic Biology, and the study, *Animal Models for Assessing Countermeasures to Bioterrorism Agents (2011).*

Copyright © National Academy of Sciences. All rights reserved.

Appendix D

Workshop Agenda and Attendees

WORKSHOP AGENDA

National Academy of Sciences Building
2101 Constitution Ave., N.W.
Washington, DC 20418

Day 1: May 28, 2014

8:00AM **Arrival and registration**
(Breakfast will be provided for committee and panelists)

8:30AM **SESSION 1: WELCOME AND OPENING PRESENTATION**
Introduction to the goals and context of the workshop
Committee Chair:
Tom Connelly
Executive Vice President & Chief Innovation Officer, DuPont

9:00AM **KEYNOTE: Achievements and Future Promise**
Doug Cameron
Co-President and Director, First Green Partners

9:45AM **SESSION 2: PERSPECTIVES ON CHEMICAL INDUSTRY PROCESS**
Panel Moderator:
Lionel Clarke
Co-Chair, UK Synthetic Biology Leadership Council

Copyright © National Academy of Sciences. All rights reserved.

Panel Objectives/Key Questions:

- What are the main drivers for adoption of bio-based processes in industry now and in the future?
- What are the main lessons from recent experience in developing industrial (bio) processes?
- How may feedstock options evolve and impact the supply chain?
- How might the adoption of bio-based processes change the nature of the chemical and energy industries?
- What are the particular considerations for commodity versus specialty chemicals?
- What are the greatest barriers to process development and what is most needed to overcome them?

Markus Pompejus
Head of Research, Bioactive Materials and Biotechnology, BASF
Mark Burk
Executive Vice President and Chief Technology Officer, Genomatica
Guo-ping Zhao
Director, Laboratory of Synthetic Biology, Institute of Plant Physiology and Ecology (IPPE), Shanghai Institutes for Biological Sciences (SIBS)
Jennifer Holmgren
Chief Executive Officer, LanzaTech

Panel discussion: ~45 min

11:30AM **Lunch** *(Lunch will be provided for committee and panelists)*

12:30PM **SESSION 3 : TECHNICAL CHALLENGES IN SAFETY AND BIOCONTAINMENT**

Panel Moderator:

Pilar Ossorio
Professor of Law and Bioethics, University of Wisconsin Law School, Madison

Panel Objectives/Key Questions:

- How should we characterize, measure, and minimize the different types of risks that could arise from different types of industrial biology (production of commodity chemicals vs. "fine chemicals")?

Copyright © National Academy of Sciences. All rights reserved.

- What types of risks might be shared across a variety of different types of chemical production?
- Are traditional risk assessment approaches adequate for understanding and responding to risks posed by biological production of commodity chemicals and "fine chemicals"?
- Into what existing risk regulation and governance frameworks will different types of industrial biology fall?
- What factors will influence the ways various publics understand the risks and uncertainties associated with biological chemical production?
- How do we communicate the risks and uncertainties of continuing on the course of traditional chemical production (i.e., there may be risks to doing something new, but there are also environmental and health risks to doing the same old thing)?
- Given that risks cannot be reduced to zero, and attempting to do so could be counterproductive and unjustifiably expensive, how do we develop systems that can quickly identify and appropriately mitigate adverse events when they do occur?

Mark Segal
Senior Microbiologist, Risk Assessment Division (RAD), U.S. Environmental Protection Agency
Eleonore Pauwels
Program Associate, Science and Technology Innovation Program, Woodrow Wilson International Center for Scholars
Dietram Scheufele
John E. Ross Professor in Science Communication, University of Wisconsin, Madison
Ed You
Special Agent, WMD Directorate, Federal Bureau of Investigation

Panel discussion: ~30 min

2:00PM **Break**

Copyright © National Academy of Sciences. All rights reserved.

2:15PM **SESSION 4: SYNTHESIS AND GENOME-SCALE ENGINEERING**

Panel Moderator:

Andy Ellington
Wilson M. and Kathryn Fraser Research Professor in Biochemistry, University of Texas at Austin
Chris Voigt
Associate Professor of Biological Engineering, Massachusetts Institute of Technology

Panel Objectives/Key Questions:

- How do we coordinate system design (parts, genes, regulation) with whole genome engineering methods, including random and directed DNA changes?
- How can we measure the impact of synthetic genetics on the host, insulate these effects, and use genome-wide information to inform the design process?
- Can we utilize part interactions with the host as part of the design, or is there a push toward ever more orthogonal systems?
- To what extent are programmed genomic interventions of any sort likely to undergo further modification as a result of selection? Can we make interventions that are robust to changes in environment and evolution?

Todd Peterson
Chief Technology Officer, Synthetic Genomics Inc.
Jennifer Doudna
Professor of Biochemistry and Molecular Biology, University of California, Berkeley
Harris Wang
Assistant Professor in Systems Biology, Columbia University
Timothy Lu
Associate Professor of Biological Engineering and Electrical Engineering, Massachusetts Institute of Technology

Panel discussion: ~30 min

3:45PM **SESSION 5: MEASURMENT (OF ENGINEERED ORGANISMS, PATHWAYS, SYSTEMS)**

Panel Moderator:

Steve Laderman
Director, Molecular Tools Laboratory, Agilent Technologies, Inc.

Copyright © National Academy of Sciences. All rights reserved.

Panel Objectives/Key Questions:

- What role does *in vitro* and *in vivo* measurement play today in pursuing the science, technology, and practice of synthetic biology?
- What are the primary methods used today?
- What do you see as the most promising emerging *in vitro* and *in vivo* measurement methods?
- What gaps in sample handling, measurement methods, data analysis and interpretation, and/or standards will still remain compared to what is desirable?
- What breakthroughs would be needed to close those gaps?
- What research would be needed to achieve those breakthroughs?
- What would be the benefits? What would that future state look like?

Drew Endy
Associate Professor, Stanford University
John McLean
Stevenson Associate Professor of Chemistry, Vanderbilt University
Johnathan Sweedler
James R. Eiszner Family Chair in Chemistry, University of Illinois at Urbana-Champaign
Marc Salit
Leader, Genome Scale Measurements Group, NIST Material Measurement Laboratory

Panel discussion: ~30 min

5:15PM **OVERVIEW FOR TOMORROW**
Committee Chair:
Tom Connelly
Executive Vice President & Chief Innovation Officer, DuPont

5:30PM **Adjourn for Day**

6:00PM **Committee will reconvene for closed session discussion over dinner**

Copyright © National Academy of Sciences. All rights reserved.

Day 2: May 29, 2014

8:00AM *(Breakfast will be provided for committee and panelists)*

8:30AM **DAY 2 WELCOME AND OPENING PRESENTATION:**
Committee Chair:
Tom Connelly
Executive Vice President & Chief Innovation Officer, DuPont

8:45AM **SESSION 6: COMPUTER-AIDED DESIGN, MANUFACTURING, AND TESTING**
Panel Moderator:
Nathan Hillson
Biochemist Staff Scientist, Berkeley Lab, Lawrence Berkeley National Laboratory
Panel Objectives/Key Questions:

- What are the bottlenecks to integrating the diaspora of bioinformatics tools and biological knowledge into a coherent industrially relevant workflow (e.g., not invented here syndrome, software licensing, software documentation, unaligned incentive structures, data structure standardization)?
- What are the current hurdles to the development of predictive genome-scale metabolic models that are accurate under industrially relevant conditions (e.g., sufficiently detailed comprehensive experimental measurements, non-steady-state mathematical frameworks, lack of fundamental biological knowledge)?
- Are there specific technical/knowledge/infrastructure challenges that, if overcome, would dramatically improve the chemical space accessible to retrosynthetic design and the accuracy thereof?
- In a world where any biological analytical measurement can be readily output in DNA (sequencing readout only required, no mass-spec, etc.), and there is a deluge of test data, what would be the resulting bottlenecks and infrastructure challenges?

Eric Klavins
Associate Professor of Electrical Engineering, University of Washington

Copyright © National Academy of Sciences. All rights reserved.

Bernhard Palsson
Galletti Professor of Bioengineering, Professor of Pediatrics, and Principal Investigator of the Systems Biology Research Group, University of California, San Diego
Chris Anderson
Assistant Professor of Bioengineering, University of California, Berkeley
Sriram Kosuri
Assistant Professor, University of California, Los Angeles

Panel discussion: ~30 minutes

10:15AM **Break**

10:30AM **SESSION 7: ADVANCED MOLECULES—WHAT DOES THE FUTURE LOOK LIKE?**

Panel Moderator:

Kristala Jones Prather
Theodore Miller Career Development Associate Professor, Massachusetts Institute of Technology

Panel Objectives/Key Questions:

- How can we expand the range of molecules and/or materials that can be industrially produced through biology?
- How do we/can we significantly increase the range of elements incorporated into biologically produced chemicals beyond carbon, hydrogen, oxygen, and nitrogen?
- What new methods are required for pathway discovery and design, enzyme discovery and design, and implementation to bring these new molecules/materials to market?

Michelle Chang
Associate Professor of Chemistry, University of California, Berkeley
Jeffrey Moore
Senior Investigator, Process Research, Merck and Company, Inc.
Mike Jewett
Assistant Professor of Chemical and Biological Engineering, Northwestern University

Panel discussion: ~30 min

Copyright © National Academy of Sciences. All rights reserved.

12:15PM **Lunch** *(Lunch will be provided for committee and panelists)*

1:00PM **SESSION 8: SCALE UP AND SCALE OUT**

Panel Moderator:

Huimin Zhao
Professor and Centennial Endowed Chair of Chemical and Biomolecular Engineering, University of Illinois, Urbana-Champaign

Panel Objectives/Key Questions:

- Process scale-up and scale-out are critical aspects of commercializing industrial processes for production of chemicals and fuels. In this panel, we will highlight several case studies such as the large-scale production of 1,3-propanediol, and 3-hydroxypropionic acid and discuss the key challenging issues in process scale-up and scale-out.
- What are the key lessons we learned from the few successful case studies?
- What are the key challenging issues in process scale-up and scale-out?
- How can we ensure the engineered organisms will behave similarly under large-scale process conditions as in small-scale laboratory conditions?
- How does technoeconomic analysis help process scale-up and scale-out?

Bill Provine
Director of Science & Technology External Affairs, DuPont
Bruce Dale
University Distinguished Professor, Michigan State University
Joel Cherry
President, Research & Development, Amyris

Panel discussion: ~30 min

2:30PM **Final Discussion and Closing Remarks**

3:00PM **Adjourn Workshop**

3:15PM ***Committee will meet for 2 hrs in closed session***

5:30PM ***Adjourn***

Copyright © National Academy of Sciences. All rights reserved.

WORKSHOP ATTENDEES

Committee Members

Michelle Chang, University of California, Berkeley
Lionel Clarke, UK Synthetic Biology Leadership Council
Thomas Connelly, E. I. du Pont de Nemours & Company
Andrew Ellington, University of Texas at Austin
Nathan Hillson, Lawrence Berkeley National Laboratory
Richard Johnson, Global Helix LLC
Stephen Laderman, Agilent Technologies, Inc.
Pillar Ossorio, University of Wisconsin Law School, Madison
Kristala Prather, Massachusetts Institute of Technology
Christopher Voigt, Massachusetts Institute of Technology
Huimin Zhao, University of Illinois, Urbana-Champaign

Speakers

Chris Anderson, University of California, Berkeley
Henry Bryndza, E. I. du Pont de Nemours & Company
Mark Burk, Genomatica
Doug Cameron, First Green Partners
Joel Cherry, Amyris
Parag Chitnis, National Science Foundation
Bruce Dale, Michigan State University
Jennifer Doudna, Howard Hughes Medical Institute/UCB
Jennifer Holmgren, LanzaTech
Mike Jewett, Northwestern University
Eric Klavin, University of Washington
Sriram Kosuri, University of California, Los Angeles
Tim Lu, Massachusetts Institute of Technology
John McLean, Vanderbilt University
Jeffrey Moore, Merck & Co.
Bernhard Palsson, University of California, San Diego
Eleonore Pauwels, Woodrow Wilson International Center for Scholars
Todd Peterson, Synthetic Genomics, Inc
Markus Pompejus, BASF Corporation
William Provine, E. I. du Pont de Nemours & Company
Marc Salit, National Institute of Standards and Technology
Dietram Scheufele, University of Wisconsin-Madison
Mark Segal, U.S. Environmental Protection Agency
Jonathan Sweedler, University of Illinois
Harris Wang, Columbia University

Copyright © National Academy of Sciences. All rights reserved.

Edward You, Federal Bureau of Investigation
Guo-ping Zhao, Shanghai Institutes for Biological Sciences (SIBS)

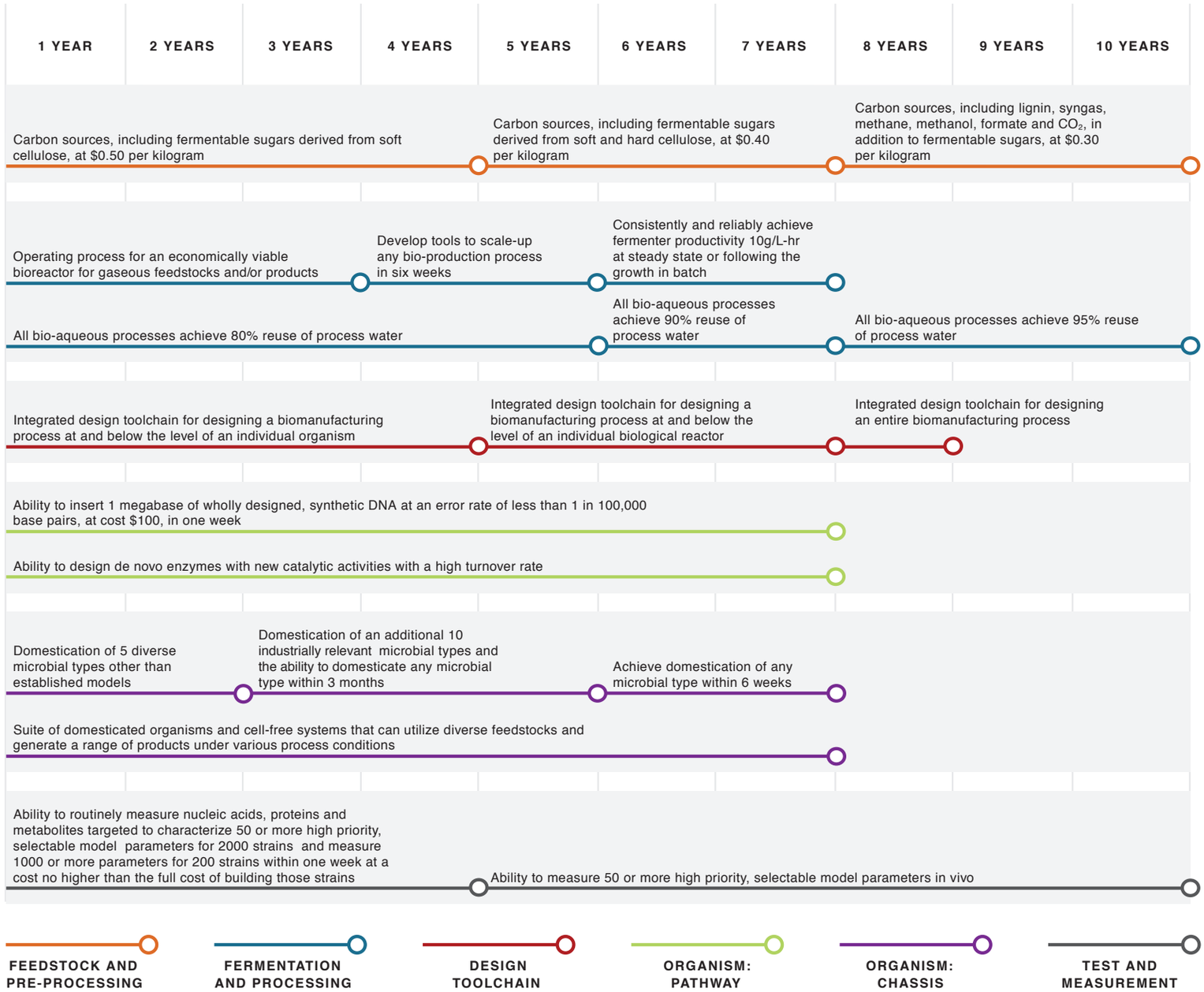
Participants

Jamie Bacher, Pareto Biotechnologies
Lynn Bergeson, Bergeson & Campbell, P.C.
Randall Dimond, Promega Corporation
Jay Fitzgerald, U.S. Department of Energy BER/AAAS
Barbara Gerratana, National Institute of General Medical Sciences/NIH
Theresa Good, National Science Foundation
Joseph Graber, U.S. Department of Energy
Ellen Jorgensen, Genspace NYC Inc.
Devin Leake, Gen9
Malin Young, Sandia National Labs
Dagmar Ringe, National Science Foundation
David Rockcliffe, National Science Foundation
David Ross, National Institute of Standards and Technology
Emily Tipaldo, American Chemistry Council
Walter Valdivia, The Brookings Institution
Susanne von Bodman, National Science Foundation
Kate Von Holle, University of Chicago
Megan Weinshank, BASF
Malin Young, Sandia National Labs

NRC Staff

Douglas Friedman, Senior Program Officer, Board on Chemical Sciences and Technology
India Hook-Barnard, Senior Program Officer, Board on Life Sciences
Carl Anderson, Research Associate, Board on Chemical Sciences and Technology
Nawina Matshona, Senior Program Assistant, Board on Chemical Sciences and Technology
Lauren Soni, Senior Program Assistant, Board on Life Sciences

Copyright © National Academy of Sciences. All rights reserved.



FIGURES S-1 and 5-1 Technical Roadmap to Enable the Industrialization of Biology.

Copyright © National Academy of Sciences. All rights reserved.