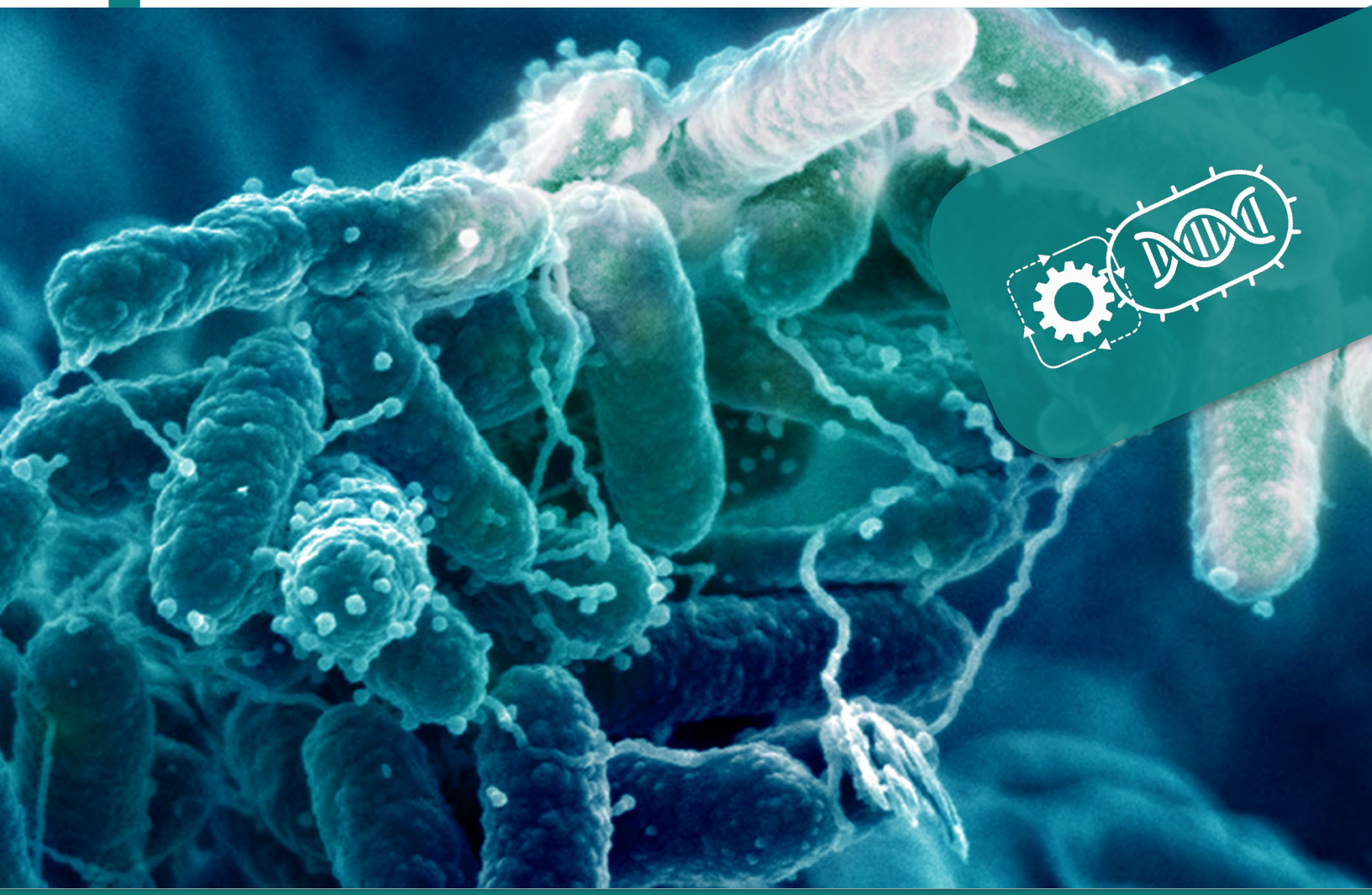
DOE/SC-2027

Microbial Design for a Developing Bioeconomy

Frontier Science for the Bioeconomy Workshop Series



U.S. DEPARTMENT of ENERGY | Office of Science
Biological and Environmental Research Program

Microbial Design for a Developing Bioeconomy Workshop

January 28–30, 2025

Convened by

U.S. Department of Energy Office of Science

Biological and Environmental Research Program

Workshop Co-Chairs

James Carothers
University of Washington

Davinia Salvachúa Rodríguez
National Laboratory of the Rockies

Organizer

Dawn Adin
U.S. Department of Energy

Writing Team

James Carothers (co-chair), University of Washington; Davinia Salvachúa Rodríguez (co-chair), National Laboratory of the Rockies; Hector García Martín, Lawrence Berkeley National Laboratory; Adam Guss, Oak Ridge National Laboratory; Pamela Peralta-Yahya, Georgia Institute of Technology

About BER

The Biological and Environmental Research (BER) program's mission is to support transformative science and scientific user facilities to achieve a predictive understanding of complex biological, Earth, and environmental systems. This research supports the U.S. Department of Energy's (DOE) vision to advance innovative solutions for U.S. energy expansion and national security challenges. BER's fundamental research, conducted at DOE national laboratories and other research institutions, plays a unique role in ensuring national leadership in biotechnology and the ability to understand and predict the interdependencies involving energy and the environment over a wide range of conditions.

Front cover image credit: Microscopic images, such as this scanning electron microscopy image of a biofilm from the Environmental Molecular Sciences Laboratory (EMSL), provide insight into how microbes influence larger reactions and processes. Courtesy Pacific Northwest National Laboratory.

Suggested citation: U.S. DOE. 2026. *Microbial Design for a Developing Bioeconomy: Frontier Science for the Bioeconomy Workshop Series,* DOE/SC-2027. U.S. Department of Energy Office of Science. https://doi.org/10.2172/3013055.

DOE/SC-2027

Microbial Design for a Developing Bioeconomy

Workshop Report

Frontier Science for the Bioeconomy Workshop Series

April 2026

U.S. DEPARTMENT of ENERGY | Office of Science
Biological and Environmental Research Program

Frontier Science for the Bioeconomy

The *Microbial Design for a Developing Bioeconomy* workshop report is one of a four-part Frontier Science for the Bioeconomy series hosted by the DOE Biological and Environmental Research program. The series defines frontiers of plant science, agriculture, synthetic biology, biobased materials, and environmental microbiome science that will unlock the promise of an innovative, resilient U.S. bioeconomy.

The bioeconomy is economic activity driven by research and involving innovation in the life sciences, engineering, and biomanufacturing. It includes industries, products (e.g., chemicals, biomass and seed oil feedstocks, and bioproducts), services, and processes derived from plant and microbial resources.

Microbial Design for a Developing Bioeconomy

Vision: Harness and leverage the diverse genetic and metabolic potential of microbes as platforms to efficiently produce biofuels, bioproducts, and biomaterials for a thriving bioeconomy.

Engineering Microbial Communities

Vision: Understand the assembly, function, and behavior of microbiomes and how to manipulate them to facilitate microbial energy solutions and advance their utility across the bioeconomy.

Plant Design for a Developing Bioeconomy

Vision: Understand and manipulate potential biomass crops to efficiently generate biofuels, bioproducts, and biomaterials under variable abiotic stresses.

Resilient Crop Production for the Bioeconomy

Vision: Determine, predict, and improve the functioning of plants and their microbiomes in the environment to optimize biomass feedstock production.

Contents

Executive Summary

A significant shift is underway in microbial design, from exploratory biology to intentional, function-driven design. This transition is powered by foundational advancements in synthetic biology, automation, and data-guided workflows that enable precise control over microbial behavior. These advancements position microbial biosystems design as an emerging cornerstone of the $1 trillion U.S. bioeconomy, offering transformative production potential for fuels, chemicals, and materials.

Workshop on Microbial Design for a Developing Bioeconomy

To assess the current landscape in biosystems design and identify future directions, the U.S. Department of Energy (DOE) convened the Microbial Design for a Developing Bioeconomy workshop January 28–30, 2025. Approximately 40 scientists and engineers from academia, industry, and national laboratories explored foundational principles, enabling technologies, and strategic opportunities to advance microbial design supporting DOE mission needs and broader national interests, including energy security and a robust bioeconomy.

The workshop explored the potential for DOE's Biological and Environmental Research (BER) program's foundational science priorities in discovery, mechanistic understanding, and microbial systems engineering to transform microbial design. Participants identified critical scientific and infrastructural challenges and outlined opportunities for DOE to leverage its unique capabilities and resources, offering a clear pathway from foundational science to mission impact.

The workshop is one of a four-part Frontier Science for the Bioeconomy series hosted by DOE's Biological Systems Science Division (BSSD) within the BER program. The series defines frontiers of plant science, agricultural sustainability, synthetic biology, biobased materials, and environmental microbiome science that can unlock an innovative, resilient U.S. bioeconomy.

Workshop participants explored the potential for BER's foundational science priorities in discovery, mechanistic understanding, and microbial systems engineering to transform microbial design.

Knowledge Gaps and Advancement Opportunities

Several critical gaps in microbial design currently impede bioeconomic advancement. The inability to reliably predict microbial behavior and understand how molecular or cellular architectures govern structure–function relationships in cellular activities limits design precision and slows development of effective biocatalysts. Biocatalysts are defined in this context as engineered organisms or enzymes that drive specific chemical reactions with high efficiency and selectivity. Commonly used model microbial platforms often lack the adaptability and scalability required for broader applications, underscoring the need to invest in microbial biodiversity and develop modular, well-characterized chassis. Finally, persistent disconnects between academic innovations and industrial needs, which are compounded by limited cross-disciplinary expertise, slow development of reliable biological systems that function effectively in real-world industrial settings.

To bridge these gaps and accelerate development of predictive biosystems and bioprocesses, several identified areas need improvement.

- Understanding and designing efficient metabolic pathways for product synthesis.

- Managing complex regulatory and metabolic networks.
- Advancing genome editing and regulation using emerging molecular tools.
- Using computational modeling to enhance strain performance and streamline pathway design.
- Using enzymes and other cell-free systems (e.g., synthetic cells) as tools to expand the biochemical reaction space and improve understanding of microbial processes.
- Designing robust biocatalysts capable of performing under industrially relevant cultivation conditions.
- Developing emerging technologies, such as high-throughput assembly, screening, and integration of artificial intelligence and machine learning (AI/ML), to accelerate novel biocatalyst design.
- Implementing infrastructure, standards, and community resources to support microbial design, including data sharing, method benchmarking, workforce training, and public engagement.

Several key themes emerged from workshop discussions and are covered in the chapters of this report, outlining takeaways essential for advancing a more comprehensive approach to biosystems design.

- **Metabolic Pathway Engineering:** Approaches to accelerate efficient pathway engineering include predicting gene function, designing efficient metabolic pathways, and controlling complex, dynamic regulatory and metabolic networks. Strategies include using orthogonal biological systems or new-to-nature pathways that minimize unintended biosystem interference and integrating microbial metabolism with nonbiological chemical reactions.
- **Microbial Genome Engineering:** Detailed investigation of genome editing and regulation using emerging molecular tools and computational modeling can enhance strain performance and streamline pathway design.
- **Cell-Free Systems:** Improved understanding and use of enzymes and cell-free platforms can expand the biochemical reaction space, accelerate prototyping, and improve mechanistic understanding of microbial processes.
- **Microbial Interactions with Engineered Environments:** Identifying the molecular factors that drive consistent microbial production relies on improved knowledge of biocatalyst behavior under industrially relevant cultivation conditions. This approach includes developing tools and systems to predict microbial performance in engineered environments.
- **Technologies To Accelerate Microbial Design:** Identifying limitations and opportunities in high-throughput assembly, screening, systems biology tools, and integration of AI/ML can accelerate development of novel biocatalysts and predictive design frameworks.
- **Shared Tools and Resources for a Microbial Bioeconomy:** Strategies to strengthen infrastructure, standards, and community resources can support enhanced microbial design. Examples include creating data-sharing platforms, establishing method benchmarking, expanding workforce training, and promoting public engagement.

Chapter 1

Microbial Design and Innovation To Unlock the Bioeconomy

The U.S. bioeconomy—currently valued at nearly $1 trillion—is expected to grow exponentially in coming years. Advances in genome editing, artificial intelligence (AI), automation, and miniaturization have set the stage for a transformation that could expand its value to $4 trillion within the next two decades (Hodgson et al. 2022). These emerging technologies are not only accelerating scientific discovery but are enabling entirely new capabilities in how biology is designed, scaled, and deployed for practical applications. The future of a robust, resilient bioeconomy no longer depends on whether it is possible but on how quickly and intentionally it can be built. Across the United States, biotechnology is advancing through the efforts of academic institutions, national laboratories, companies, industry alliances, workforce development programs, and policy initiatives that drive the future of this rapidly growing field (Hodgson et al. 2022; NSCEB 2025).

To support these efforts, DOE's Biological and Environmental Research program (BER) convened a workshop titled Microbial Design for a Developing Bioeconomy from January 28–30, 2025 (see Appendix A, p. 53). The workshop goal was to assess how DOE and BER missions can accelerate the ability to harness and leverage the diverse genetic and metabolic potential of microbes (see Section 1.5: Workshop Themes: Microbial Design for a Developing Bioeconomy, p. 6). Bacteria, fungi, viruses, archaea, microalgae, and protists demonstrate potential as biocatalyst platforms to efficiently produce biofuels, bioproducts, and biomaterials that will form the backbone of a thriving and resilient U.S. bioeconomy (see Fig. 1.1, p. 2).

1.1 Realizing the Potential of Biocatalysis

Microbial design is no longer confined to the laboratory. It can potentially reshape global supply chains and redefine how to source, build, and sustain the energy and materials that support society. Designed enzymes and other microbial components serve as biocatalysts that address the growing demand for fuels, chemicals, materials, and therapeutics.

In human health, for example, cell-based technologies are driving a revolution in medicine. CRISPR-Cas9 systems have opened new avenues for gene therapy by providing precise and accessible gene editing techniques for microbial biocatalyst engineering. In addition, microbial systems serve as platforms for producing antibiotics and vaccines, phage display techniques accelerate discovery of new antibodies, cell-free diagnostics enable fast and portable disease detection, and engineered enzymes are increasingly central to pharmaceutical synthesis.

The impact of engineered biology extends beyond healthcare into industrial biotechnology. By converting low-cost, alternative feedstocks into a broad portfolio of valuable products, biocatalysts can help meet societal needs while enhancing the resilience of supply chains. These biological systems still face challenges in achieving the titers, rates, and yields required for commercial viability, but biocatalysis has already demonstrated its

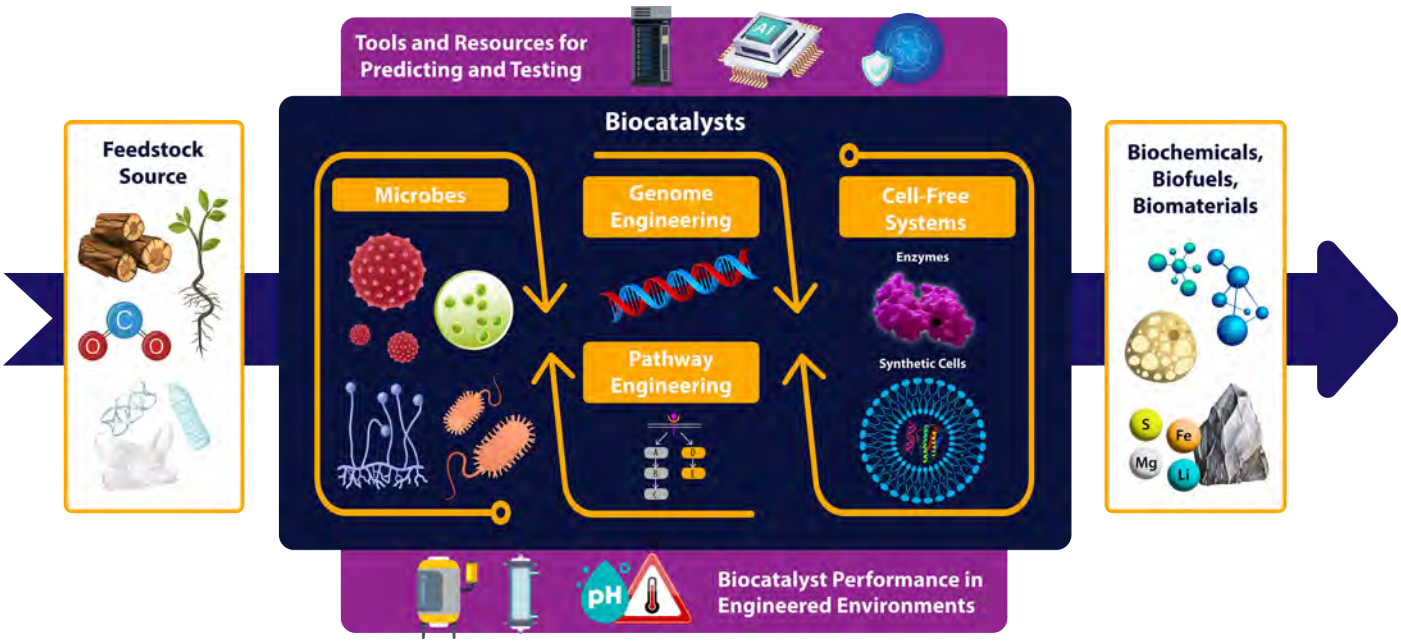


Fig. 1.1. Workshop Framework for Microbial Design for a Developing Bioeconomy. The main themes explored during the workshop, highlighted in orange and purple, aimed to identify foundational knowledge gaps and strategic opportunities in microbial design for advancing robust biological conversion of diverse feedstocks into biochemicals, biofuels, and biomaterials.

transformative power across multiple sectors, laying the groundwork for future breakthroughs.

Traditional microbial fermentations continue to support brewing and winemaking industries, while emerging companies are introducing alternative proteins and nondairy products through precision fermentation. Another example is the enhanced performance of detergents and additives by incorporating biological products like enzymes and biosurfactants.

Industrial biotechnology to produce biofuels and chemicals has also achieved significant milestones (Lee et al. 2019; Keasling et al. 2021). Examples in the industrial sector include microbial conversion of waste gases (e.g., carbon dioxide and carbon monoxide) into ethanol, acetone, and isopropanol, as well as conversion of lignocellulose-related sugars into other value-added chemicals (e.g., 1,4-butanediol and 1,3-propanediol).

Industrial biotechnology is also rapidly transforming materials design and production. One example is mycelium-derived materials production (Vandelook et al. 2021). Filamentous fungi mycelia can be produced on agricultural residues and shaped into biodegradable packaging materials. Once dried, the resulting product is lightweight, durable, and compostable. Advances in strain selection and process control focus on improving growth rates and material properties to enable applications beyond packaging, including building insulation and structural composites. Fungal mycelia are also being used to produce leather-like materials. Through controlled cultivation and post-processing, fungal biomass can be transformed into flexible sheets with leather-like properties. Ongoing biological efforts aim to improve strength, texture, elasticity, and water resistance, which could be achieved by further understanding and tuning mycelial networks.

In chemical manufacturing, engineered microbes are increasingly used to produce polymer precursors. For instance, microbial catabolic intermediates (e.g., muconic acid and β-ketoadipate) can be converted into key precursors for nylon and other materials from lignocellulose-derived monomers (Johnson et al. 2019).

Another promising potential industrial application is the development of living materials (Johnston et al. 2020; McBee et al. 2022; An et al. 2023). Living

materials are composites that incorporate engineered cells capable of sensing and responding to their environment and performing functional tasks. Through synthetic biology, organisms can be programmed to produce structural components, self-heal when damaged, or adapt to environmental changes.

Recent advances have demonstrated that microbes immobilized in shear-thinning hydrogels can be spatially organized and preserved for long-term reuse and on-demand bioproduction. These hydrogel-based systems outperform traditional liquid cultures in stability and repeatability, enabling plug-and-play bioproduction that can be activated when needed (Johnston et al. 2020).

Another example of living materials is biocement, generated by microbes via an induced calcium carbonate precipitation process (Gebru et al. 2021). When properly produced, calcium carbonate forms strong bonds between soil particles, resulting in high, unconfined compressive strength and low water permeability. Improving metabolic stability and environmental resilience will be key to broader implementation. Overall, these materials hold promise for applications in responsive architecture, wearable technology, regenerative infrastructure, and durable and flexible biomanufacturing platforms, although future work is needed to fully realize their potential for field-ready deployment.

Finally, industrial biotechnology is enabling the biological extraction and recovery of critical minerals (NSCEB 2025). Certain microbes can solubilize metals from low-grade ores or electronic waste through processes such as bioleaching and biosorption. Enzymes and transport proteins can be engineered to improve selectivity and efficiency, offering a more decentralized alternative to conventional mining practices. This approach may play a key role in securing supplies of rare earth elements and other essential materials for electronics, energy storage, and defense.

1.2 Factors Critical for Success in Microbial Design and Biotechnology

A set of common, foundational factors has driven successful microbial design and biotechnologies across these diverse sectors. Together, the most resilient and scalable biotechnologies are those that blend technical excellence with strong market demand, process integration, and social trust.

Industrial biotechnology is enabling the biological extraction and recovery of critical minerals.

- First and foremost, technical success has required significant investment in research and engineering. Moving from milligram production titers in the laboratory to industrial-scale production has often taken years of development and collaboration among large, multidisciplinary teams. For many platforms, including biofuels and specialty chemicals, scale-up has only been possible through deep understanding of both biological systems and chemical engineering principles as well as rigorous process optimization.
- Certain products, such as ethanol, have benefited from inherent process advantages, such as antimicrobial properties that reduce contamination in the bioconversion process and relatively simple downstream processing steps. These features have accelerated adoption and enabled the bioethanol industry to scale rapidly.
- Biological tools like cell-free systems have become more accessible, simplifying early research stages and enabling fast iteration and design cycles toward their implementation.
- Clearly defined performance metrics and access to foundational data have made it possible to develop predictive models and reduce experimental trial-and-error.
- Economic incentives have played a critical role in developing and deploying microbial-based applications. In addition, in cases with a strong market pull, biotechnologies have gained traction despite technical complexity.
- Regulatory clarity and public acceptance have further smoothed the transition from laboratory to market.

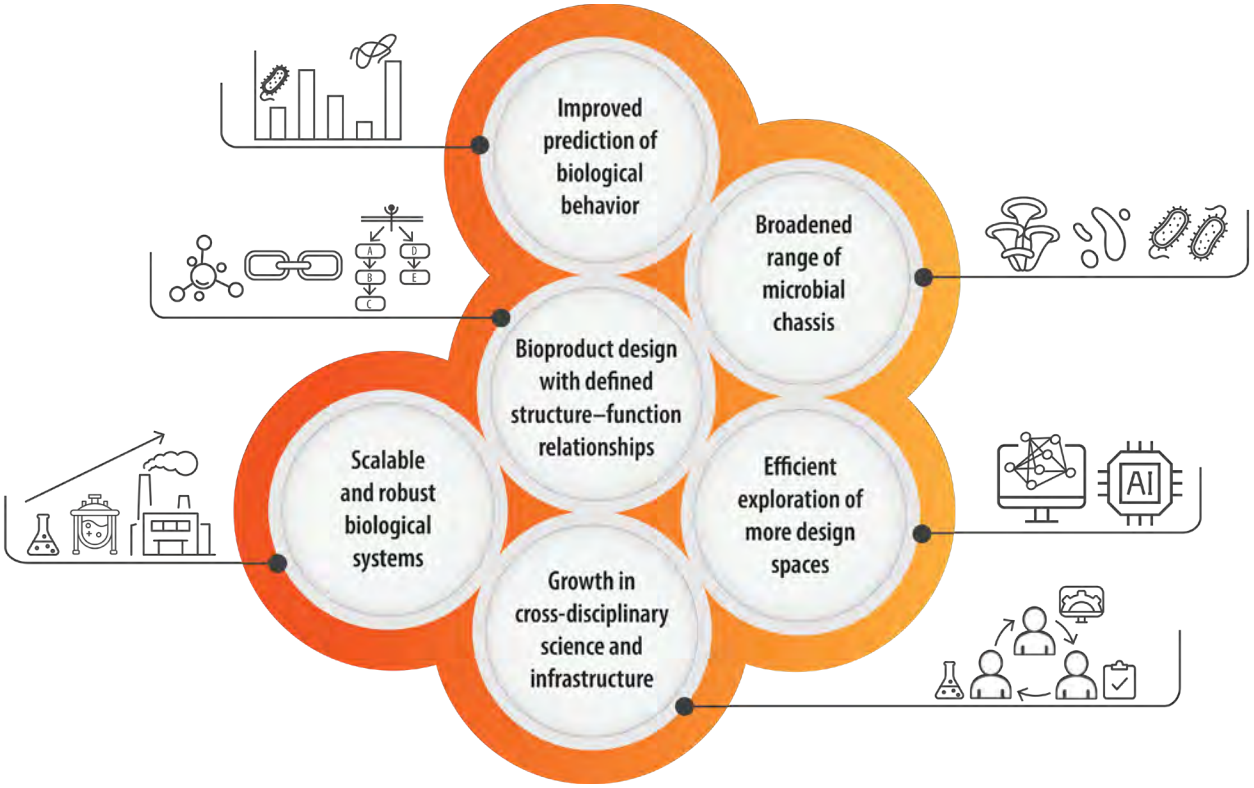


Fig. 1.2. Microbial Design Gaps. Advances are needed to fill critical knowledge and technical gaps in microbial design to achieve a thriving bioeconomy. Addressing current limitations in prediction, scalability, and cross-disciplinary integration is essential to accelerate the design of reliable, high-performing biological systems.

1.3 Key Gaps in Microbial Design for a Deployable Bioeconomy

Despite progress, several challenges continue to limit the widespread deployment of microbial systems in the bioeconomy. As microbial design becomes increasingly central to industrial biotechnology, addressing key knowledge gaps and technical limitations will be critical for building a bioeconomy that is not only innovative but also transferable to industrial scales (see Fig. 1.2, this page).

- **Biological Behavior Prediction:** A primary barrier to deployment is the limited ability to accurately predict microbial behavior in complex systems. Unlike traditional engineering disciplines, biology still lacks the tools to fully forecast how genetic or metabolic modifications will impact microbes under diverse conditions or scales. This unpredictability hampers design precision and increases development time and cost.
- **Structure–Function Relationships:** The field still struggles to design biological products with clear structure–function relationships that provide predictable links between how enzymes build molecules and what the molecules do.
- **Complex Design Spaces:** Progress is further slowed by the difficulty of efficiently exploring complex design spaces with multiple variables. For example, metabolic pathway optimization involves multiple tunable genes and regulatory elements; enzyme engineering entails numerous possible amino acid substitutions; and strain design is complicated by growth conditions, gene expression levels, and host context that all interact to affect performance.

- **Microbial Chassis:** Microbial platforms themselves also present limitations. Historically, metabolic engineering efforts have relied on model organisms, such as *Escherichia coli* or *Saccharomyces cerevisiae*, which may not be optimal for a broad range of target products. It is therefore essential to expand the range of available microbial chassis and ensure they are well-characterized and modular. Doing so will require investment in infrastructure, genetic tool development, data sharing, and microbial biodiversity exploration, as well as community engagement to build shared resources.
- **Scalable and Robust Biological Systems:** A persistent disconnect also exists between academic research, which often focuses on proof-of-concept demonstrations, and industrial innovation, which requires robust biological systems operating in large, industrial volumes. Bridging this gap between academic research and industrial applications will require new models for collaboration and a stronger emphasis on scale-aware design even at early stages of research.
- **Cross-Disciplinary Efforts:** Workforce and team structures in synthetic biology often lack the cross-disciplinary expertise needed for success. Integrating biology with engineering, computation, regulatory science, and business not only requires diverse talent but also training models that promote crosstalk and systems thinking.

1.4 A Vision for the Future of Microbial Design

Microbial systems offer powerful tools for understanding biological design and converting a wide range of feedstocks into valuable fuels, chemicals, and materials. Systems include bacteria, fungi, viruses, archaea, microalgae, and protists, either individually or in cocultures, as well as alternative systems such as cell-free platforms and synthetic cells. The workshop did not center on any specific biocatalyst, feedstock, or product. Discussions focused on the broader advantages and limitations of utilizing biocatalysis across different applications. Key themes included selecting biocatalysts, pairing microbial hosts with feedstocks, and matching biocatalysts to products.

Microbial systems offer powerful tools for understanding biological design and converting a wide range of feedstocks into valuable fuels, chemicals, and materials.

Selecting Biocatalysts

As industrial processes grow more complex, the environments and substrates involved demand equally sophisticated microbial traits. To meet these needs, the biotechnology field is moving beyond well-studied model organisms and turning to nonmodel microbes that bring unique advantages, such as metabolic activity at extreme pH or temperature; tolerance to toxic substrates, products, or byproducts; outstanding high growth rates; consumption of unusual feedstocks; or simplified product recovery.

Nonmodel organisms could significantly improve industrial bioconversion processes, but they are often poorly understood and difficult to engineer. Bioprospecting remains a key strategy for finding promising new candidates by drawing from nature's diversity and evolutionary processes to discover extremophiles, enzymes, or chemistries not found in conventional hosts.

Additionally, biological and chemical catalysis should not be viewed as distinct approaches; biocatalyst selection could include hybrid systems that integrate both. Realizing this vision will require close integration across disciplines and AI-guided workflows for the discovery, development, and deployment of the next generation of microbial catalysts.

Pairing Microbial Hosts with Feedstocks

The workshop explored a critical question of how to best match microbial hosts to feedstocks such as lignocellulosic biomass, post-consumer waste (e.g., synthetic polymers like plastics), and one-carbon

gases (e.g., carbon dioxide, carbon monoxide, and methane) but did not center on any specific feedstock.

One foundational discussion centered on whether biocatalyst development should be driven by precisely pairing specific microbes with defined substrates or by designing systems with maximum substrate flexibility from the outset. Both approaches offer distinct advantages, and the consensus leaned toward the importance of flexibility in system design. One path aims to optimize microbe–substrate pairs to achieve maximum efficiency under controlled conditions, especially for low-diversity feedstocks like syngas or methane. The other favors building modular, tunable systems capable of operating across a wide range of feedstocks. The latter approach provides resilience and adaptability, which are particularly valuable in distributed biomanufacturing settings where input streams can vary regionally.

Ultimately, deep understanding of microbe–substrate interactions, which are shaped by both ecological context and metabolic capacity, is essential for engineering efficient microbial systems that are also adaptable to real-world conditions.

Investments are needed in technologies that detect novel compounds and illuminate the cellular networks responsible for their synthesis.

Matching Biocatalysts to Products

Another foundational discussion centered on identifying the advantages and limitations of matching specific bioproducts (e.g., biofuels, biochemicals, and biomaterials) with specific biocatalysts from the outset when developing next-generation microbial systems. This consideration echoes the earlier discussion on matching hosts to feedstocks but introduces distinct product-driven implications.

The benefits of early matching include a more focused development path, streamlined screening, and potentially faster scale-up. Tailoring the microbial system to a defined product reduces the need for engineering ancillary pathways or components, thereby shortening development timelines. However, this strategy comes with trade-offs. Systems designed around a single product may lack flexibility and portability, potentially resulting in highly specialized pathways that are difficult to adapt to other hosts or production targets.

In the context of matching biocatalysts to products, the product class (e.g., fuel, chemical, material, or more complex output) plays a central role in shaping development strategies. For small molecules (both intracellular and secreted), success often centers on achieving high metabolic flux, managing byproduct formation, and ensuring effective transport across cell membranes. For macromolecules (e.g., polysaccharides or proteins), key considerations include proper folding, secretion efficiency, and minimizing cellular burden during synthesis. Living materials, which embed biological function into structural roles, introduce new challenges related to stability, controlled growth, and self-repair. Lastly, for inorganic products (e.g., bioprecipitated metals or minerals), opportunities exist to better understand precipitation pathways.

Critically, the landscape of possible products is still expanding. Metabolomics and other high-resolution tools continue to uncover previously unknown molecules whose biological roles remain unclear. As discovery pipelines broaden, investments are also needed in technologies that not only detect novel compounds but also illuminate the cellular networks responsible for their synthesis.

1.5 Workshop Themes: Microbial Design for a Developing Bioeconomy

The Microbial Design for a Developing Bioeconomy workshop brought together scientists and engineers (see Appendix B, p. 57) to identify key fundamental knowledge gaps and opportunities in designing biological systems to produce fuels, chemicals, and materials. Discussions were structured around six core themes (see Appendix C, p. 58):

- **Metabolic Pathway Engineering:** Exploring strategies to better understand and design efficient

metabolic pathways, manage complex regulatory and metabolic networks, and improve microbial metabolism for product synthesis (see Ch. 2: Filling Key Knowledge Gaps in Metabolic Pathway Engineering, p. 9).

- **Microbial Genome Engineering:** Focusing on advancing genome editing and regulation through emerging molecular tools and computational modeling to enhance strain performance and streamline pathway design (see Ch. 3: Filling Key Knowledge Gaps in Microbial Genome Engineering, p. 17).
- **Cell-Free Systems:** Examining enzymes as tools to expand the biochemical reaction space and improve understanding of microbial processes (see Ch. 4: Leveraging Cell-Free Systems for Bioproduct Synthesis and Microbial Design, p. 27).
- **Biocatalyst Performance in Engineered Environments:** Addressing challenges to designing robust biocatalysts capable of performing under large-scale, relevant cultivation conditions (see Ch. 5: Basic Science Investigation into Biocatalyst Performance in Engineered Environments, p. 31).
- **Technologies To Accelerate Microbial Design:** Discussing limitations and opportunities in emerging technologies and high-throughput assembly, screening, and AI and machine learning (ML) integration to accelerate development of novel biocatalysts (see Ch. 6: Technologies To Accelerate Microbial Design, p. 39).
- **Shared Tools and Resources To Accelerate Microbial Design:** Discussing the infrastructure, standards, and community resources needed to support microbial design, including data sharing, method benchmarking, workforce training, and public engagement (see Ch. 7: Shared Tools and Resources To Accelerate Microbial Design, p. 45).

Chapter 2

Filling Key Knowledge Gaps in Metabolic Pathway Engineering

An important workshop theme was metabolic pathway engineering—the process of using rational or nontargeted approaches to understand or optimize metabolic flux from a feedstock to a desired product within a microbial host. Typically, this process first involves introducing and overexpressing heterologous microbial enzymes and pathways, suppressing certain native pathways, and modifying native carbon and electron flux in the organism.

Once a functional metabolic pathway is established, a combination of approaches can be employed to optimize pathway flux and maximize product yield. Approaches include targeted gene deletion and overexpression, multiomic profiling, metabolic modeling, and genome-scale perturbation techniques such as CRISPR interference (CRISPRi) and transposon mutagenesis. Adaptive laboratory evolution (ALE) further enables microbial performance refinement under selective pressures, revealing beneficial mutations and emergent strains (Sandberg et al. 2019; Hu et al. 2024; Hren et al. 2025). Together, these approaches form the foundation of functional genomics, offering a systems-level view of cellular behavior and guiding effective strain engineering (see sidebar, The Tools of Functional Genomics, p. 10).

Recent advances in molecular biology, genetics, omics technologies, bioinformatics, and computational sciences have significantly accelerated progress in metabolic pathway engineering. However, substantial knowledge gaps remain, limiting the ability to predictably and efficiently engineer microbial metabolism. This chapter highlights workshop-identified gaps in metabolic pathway engineering, as well as opportunities and strategies aimed at reducing complexity and improving control and predictability in microbial design. The discussion also includes prospects in pathway orthogonality, new-to-nature pathways, and integration with nonbiological chemical reactions.

Recent advances in molecular biology, genetics, omics technologies, bioinformatics, and computational sciences have significantly accelerated progress in metabolic pathway engineering.

2.1 Current Knowledge Gaps in Metabolic Pathway Engineering

Metabolic pathways remain inherently challenging to engineer despite decades of progress in metabolic engineering leading to production improvements for biofuels, amino acids, and other valuable chemicals (Ling et al. 2022). The primary cause is a limited fundamental understanding of biological system complexity. Addressing several critical knowledge gaps would significantly advance the field and enable more precise and efficient metabolic pathway engineering.

The Tools of Functional Genomics

The field of functional genomics examines how genes and their products influence biological processes. Various experimental methods and computational analyses are used to systematically perturb genes or pathways to understand their roles, interactions, and impact on cellular phenotypes. This systems-level approach enables deeper insight into metabolic control, stress responses, and emergent traits, laying the groundwork for predictive engineering.

CRISPR: CRISPR is a genome engineering technology that uses a guide RNA to direct a Cas protein (e.g., Cas9) to specific DNA sequences for targeted modification. Traditional CRISPR systems create double-strand breaks in DNA to enable gene insertions, deletions, or replacements. Advanced forms include nucleotide base editing, which enables single-nucleotide conversions without cutting DNA, and prime editing, which uses a Cas9 nickase fused to a reverse transcriptase and a specialized guide RNA to introduce a wide range of precise edits without creating double-strand breaks.

CRISPR Interference (CRISPRi) and CRISPR Activation (CRISPRa): CRISPRi and CRISPRa employ a catalytically inactive Cas protein (e.g., dCas9) to reversibly repress or enhance gene expression without altering the DNA sequence. These tools enable tunable, multiplexed control of gene regulation, supporting high-throughput functional genomics and precision reprogramming of cellular networks.

Transposon Mutagenesis: This genetic technique randomly inserts mobile DNA elements (transposons) into a host genome, disrupting gene function at the site of integration. The approach enables generation of genome-wide mutant libraries and is broadly used to identify essential genes, map functional pathways, and study gene–phenotype relationships.

Adaptive Laboratory Evolution (ALE): This method improves microbial traits by cultivating populations under defined environmental or selective pressures over many generations. Advantageous mutations accumulate through natural selection, enabling researchers to evolve strains with enhanced fitness, productivity, or stress tolerance. The resulting genotypes can then be analyzed to identify causal mutations and adaptive mechanisms.

Multiomic Analyses: This approach integrates multiple layers of biological information (e.g., genomics, transcriptomics, proteomics, metabolomics, and lipidomics) to provide a comprehensive, systems-level understanding of cellular function. Multiomic analysis enables detailed characterization of phenotypic traits and helps reveal complex interactions between genetic, regulatory, and metabolic networks.

Biological System Complexity and Unpredictability

A primary barrier to microbial metabolic pathway engineering is the inherent complexity and unpredictability of biological systems. Most microbes encode thousands of genes, the majority of which have unknown functions and therefore unpredictable behavior in engineered pathways. This complexity hinders rational biosystem design, with genetic modifications often creating unexpected results, which impact cellular function across multiple levels. Even when nontargeted approaches yield improved strains (e.g., ALE or high-throughput mutagenesis), the underlying mutations rarely provide clear mechanistic insights, limiting the ability to generalize findings or apply them predictively.

Unlike physical systems, biological systems generally lack comprehensive models capable of accurately forecasting system behavior, leading to time-consuming trial-and-error approaches to strain

development. While these challenges are substantial even in well-studied model organisms, they are further amplified in nonmodel microbes. Gene function in nonmodel organisms is primarily inferred based on homology to characterized systems, but metabolic regulation and interdependencies remain poorly understood. These factors further complicate efforts to engineer desired phenotypes.

Unidentified Design Principles Governing Biological Systems

Biological complexity also contributes to a limited understanding of the design rules governing biological systems. While natural systems typically operate with multiple layers of regulatory control, current engineered systems often incorporate no regulatory control or only a single layer.

Tools and strategies to implement multilayered control are underdeveloped, and whether such complexity would tangibly improve bioconversion is not yet well-understood. Engineered pathways often face toxicity and regulatory feedback, which are difficult to manage. Effective approaches to finely tune and dynamically control metabolism at the levels of transcription, translation, post-translation, and metabolic flux are also lacking.

Although transcriptional control of biological systems is fairly well understood, science generally lacks a mechanistic understanding of these processes. For instance, while microbes naturally employ allosteric regulation to adjust enzyme activity in response to cellular conditions (Xu et al. 2012), current ability to engineer or repurpose allosteric mechanisms remains severely limited. In addition, the lack of well-characterized and diverse enzyme orthologs for many key metabolic reactions restricts the ability to select or design enzymes with optimal catalytic properties and limits AI and machine learning (AI/ML) approaches to enable accurate predictions.

Challenges in Pathway Orthogonality and Integration

Another challenge in metabolic engineering is how to balance and leverage pathway orthogonality—where a desired product pathway functions independently from pathways for essential cellular processes—and its integration within a broader cellular network. Currently, science lacks a clear understanding of how to design truly orthogonal systems, how to effectively integrate synthetic pathways into the whole-cell context of host metabolism, and under which circumstances orthogonality or integration may be preferable.

This challenge is further complicated by a limited understanding of how pathway integration intersects with cellular stress responses and regulatory networks, particularly in nonmodel organisms. Evolution also constrains metabolism, often in unpredictable ways. Metabolic pathways that decrease cellular fitness tend to perform poorly and inconsistently, and not all desired traits are easily linked to growth.

2.2 Opportunities and Strategies To Reduce Complexity and Enhance Control and Predictability in Microbial Design

Workshop discussions centered around three primary approaches for addressing knowledge gaps in metabolic pathway engineering. The first of these, reducing complexity and enhancing control and predictability in microbial design (see Fig. 2.1, p. 12), highlighted three strategies: deepen understanding of cellular metabolism, improve control of gene expression, and accelerate comprehensive characterization of nonmodel microbes.

Deepen Understanding of Cellular Metabolism

A key strategy for reducing microbial design complexity is to deepen the fundamental understanding of cellular metabolism across a wide range of microorganisms. Achieving this will require innovative approaches to characterize genes with unknown or poorly characterized functions. High-throughput genetics and strain phenotyping combined with advancements in bioinformatics, biochemistry, analytical chemistry, imaging, and AI/ML offer significant opportunities to accelerate this discovery process.

In parallel, AI models capable of predicting enzyme activity and substrate specificity directly from sequence data are critically needed. These models would benefit from large, high-quality datasets through systematic enzyme characterization efforts. Such efforts would not only support deep learning approaches but also yield curated collections of well-characterized orthologous enzymes, expanding the toolkit available for pathway engineering.

Insights gained from these efforts can be integrated into advanced computational models to guide metabolic engineering and improve predictability. Ultimately, these advancements could enable development of cellular digital twins—comprehensive, dynamic models of entire cells—that could simulate cellular behavior under various conditions and guide rational engineering with unprecedented precision (see sidebar, Digital Twins To Revolutionize Microbial Design, p. 13).

Controlling metabolic pathways and associated fluxes will be critical to future metabolic pathway engineering.

Improve Control of Gene Expression

Controlling metabolic pathways and associated fluxes will be critical to future metabolic pathway engineering. Static expression systems can increase metabolic burden and decrease robustness, whereas dynamic regulation offers the potential to create cells that sense and respond to changing environments (Dahl et al. 2013).

Transcriptional control is the most explored area, where transcription factors (i.e., biosensors) regulate gene expression in response to chemical or physical signals. Characterizing novel transcription factors and improving the ability to engineer new biosensors will lay the foundation for rapid deployment of controlled gene expression in metabolic pathways.

Additional opportunities also exist for finer control by targeting post-transcriptional regulation mechanisms, including tools for engineering post-translational protein modifications, small molecule–based allosteric control of enzyme activity, and controlled protein degradation. Predictive models that incorporate these dynamic regulatory layers are also needed to identify new control points and enable enhanced strain performance.

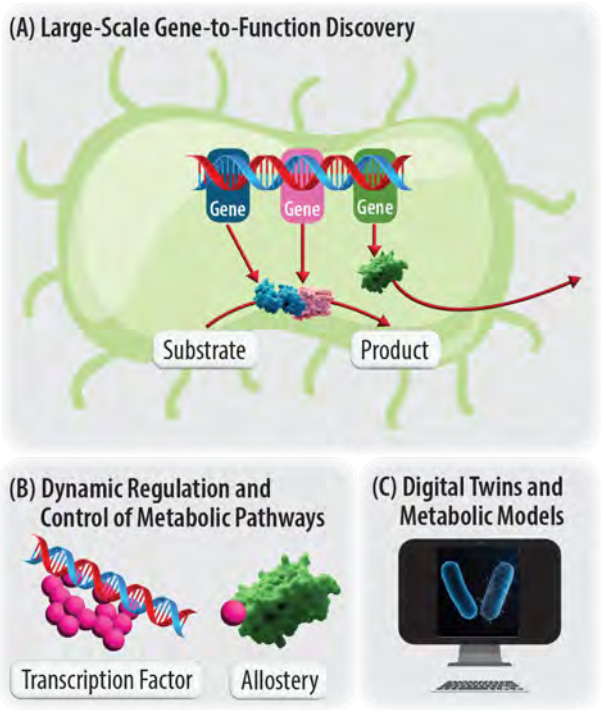


Fig. 2.1. Three Primary Strategies Help Reduce Biological Complexity and Speed Microbial Design. **(A)** High-throughput determination of the biological role of genes from their DNA sequence to their translation into proteins that affect cell function. **(B)** Metabolic pathways can be dynamically controlled with transcription factors that bind to DNA and control gene expression (e.g., biosensors) or through allostery in which molecules bind to proteins to control their activity. **(C)** Microbial digital twins and metabolic models offer a complete computational representation of a microbe and every aspect of its metabolism from gene expression to producing bioproducts, providing a platform for metabolic redesign.

Accelerate Comprehensive Characterization of Nonmodel Microbes

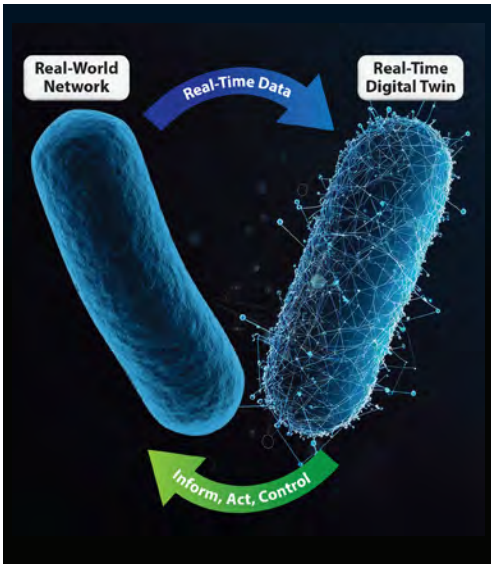
Model organisms have provided a platform for significant fundamental research and engineering in metabolic pathway design. However, nonmodel microbes often exhibit complex phenotypes that promise bioprocessing

Digital Twins To Revolutionize Microbial Design

Digital twins are virtual models or replicas of a physical object, system, or process that simulate real-world behavior and characteristics in real time. The concept was introduced at the beginning of the 21st century in the manufacturing field but did not gain traction until the mid-2010s. Digital twins can revolutionize microbial design by enhancing predictive power, accelerating design cycles, improving systems understanding, and enabling dynamic control.

A microbial cell digital twin is a computational model that mirrors the cell's behavior and internal processes in real time, leveraging both data-driven and mechanistic approaches. It can be designed to simulate and predict how a specific microbe or cell type responds to genetic modifications, environmental changes, or other perturbations. When integrated with bioprocess systems, digital twins may enable real-time monitoring and adaptive control of microbial bioprocesses.

While a digital twin itself is a virtual model, it can be continuously fed by real-time data streams from sensors and instruments monitoring the physical microbial system or bioprocess. This live data would update the digital twin's state, allowing it to reflect the current condition of the real system instantaneously.



Digital Twins Could Accelerate Prediction and Control of Bioprocesses. Real-time data from microbial cultivations is used to update and refine the digital twin, enabling more accurate predictions of microbial behavior. These insights can then inform and optimize real-world cultivation processes through active monitoring and control.

advantages. A major challenge lies in determining which insights from model systems are transferable to these less-characterized organisms. Regardless, nonmodel microbes present opportunities to discover potentially enabling novel biological mechanisms in the field of metabolic pathway engineering. Therefore, accelerating comprehensive genetic and physiological characterization of diverse microbes will be essential for advancing the future of microbial design.

2.3 Opportunities for Orthogonal, Integrated, and New-to-Nature Pathways

The second approach for addressing knowledge gaps in metabolic pathway engineering is exploring orthogonal, integrated, and new-to-nature pathways (see Fig. 2.2, p. 14). The inherently programmable and modular nature of biology lends itself to incorporating synthetic monomers and novel chemistries into bioproduction. However, research is needed to understand how to effectively decouple engineered metabolic pathways from the microbial host's cellular context.

Designing orthogonal biological systems isolated from essential cellular processes offers a promising strategy to minimize crosstalk between pathways, reduce unintended interference, and enable precise control over engineered biosynthetic routes. Orthogonal processes have the potential to simplify cellular complexity and improve pathway performance.

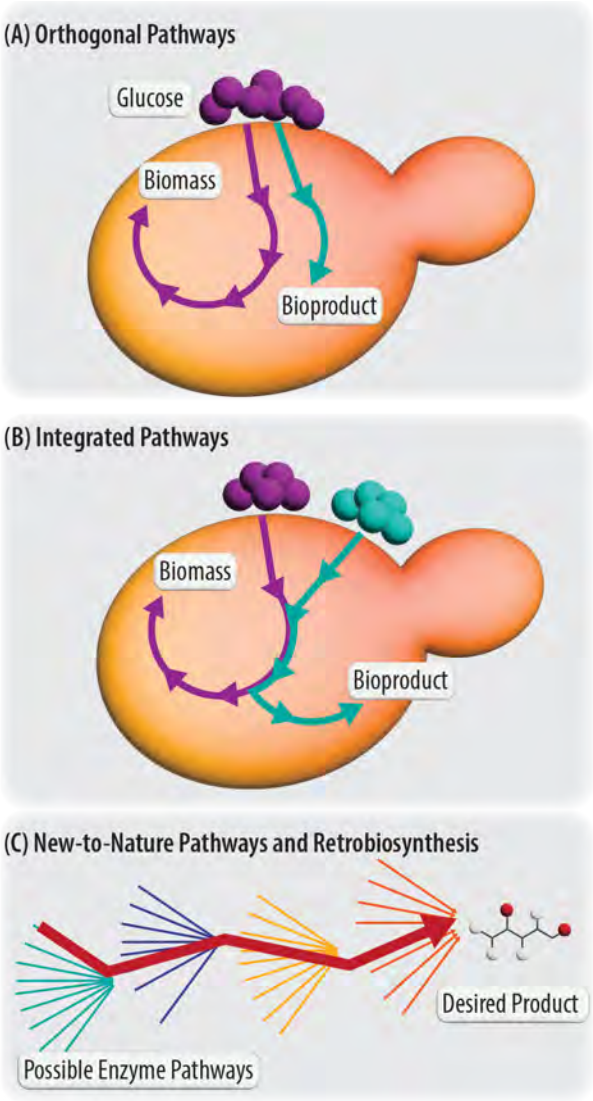


Fig. 2.2. Several Strategies for Engineering Metabolic Pathways Leverage the Modular Nature of Biology. Pathway engineering strategies create novel cellular functions capable of producing proteins, enzymes, cofactors, and final bioproducts. **(A)** Orthogonal approaches create pathways for producing target chemicals that are separate from pathways involved in cell growth. The approach aims to optimize production without interfering with the rest of the cellular context. **(B)** Integrated approaches add new steps and products while leveraging existing cellular pathways. This approach can harness native complexity to control and enhance bioproduction. **(C)** Retrobiosynthesis builds new-to-nature pathways by borrowing and combining steps from existing pathways. Pathway engineers work backward to find a chain of chemical reactions that will produce the enzymes needed to build a chemical targeted for bioproduction.

Research opportunities in orthogonal pathway design include creating and developing orthogonal ribosomes capable of incorporating noncanonical amino acids into proteins or polymerizing novel materials. Another strategy is to uncouple viability from bioproduction, potentially through cell-free systems that isolate and optimize specific reactions.

An opposite approach to orthogonality is deep integration of metabolic pathways into the native or existing cellular context. Understanding design rules for such integration requires extensive fundamental research into the mechanisms underlying metabolic interconnectedness across diverse microbes. This research challenge presents an opportunity to learn from and emulate nature's strategies for coordinating complex metabolic networks. Novel modeling approaches will critically guide this integration.

New-to-nature pathways present unique challenges and corresponding opportunities for metabolic engineering. Improved tools for retrobiosynthetic pathway predictions could enable more effective pathway designs and better metabolic integration into the cellular context. Advances in computational protein design and high-throughput biochemical characterization are essential for creating novel enzymes that will expand the diversity and feasibility of new-to-nature pathways. Thermodynamic and kinetic bottlenecks in these novel pathways may be addressed through advancements in *in vivo* enzyme clustering, including using scaffolds or synthetic organelles to colocalize reactions, shield them from competing reactions, and minimize the toxicity of pathway intermediates.

2.4 Engineering Microbial Pathways To Leverage Nonbiological Chemical Reactions

The third approach for addressing knowledge gaps in metabolic pathway engineering involves leveraging nonbiological chemical reactions in microbial pathways or biotechnological processes (see Fig. 2.3, p. 15). In optimized bioprocesses for target chemical production, bioconversion often outperforms chemical conversion because biocatalysis occurs under mild

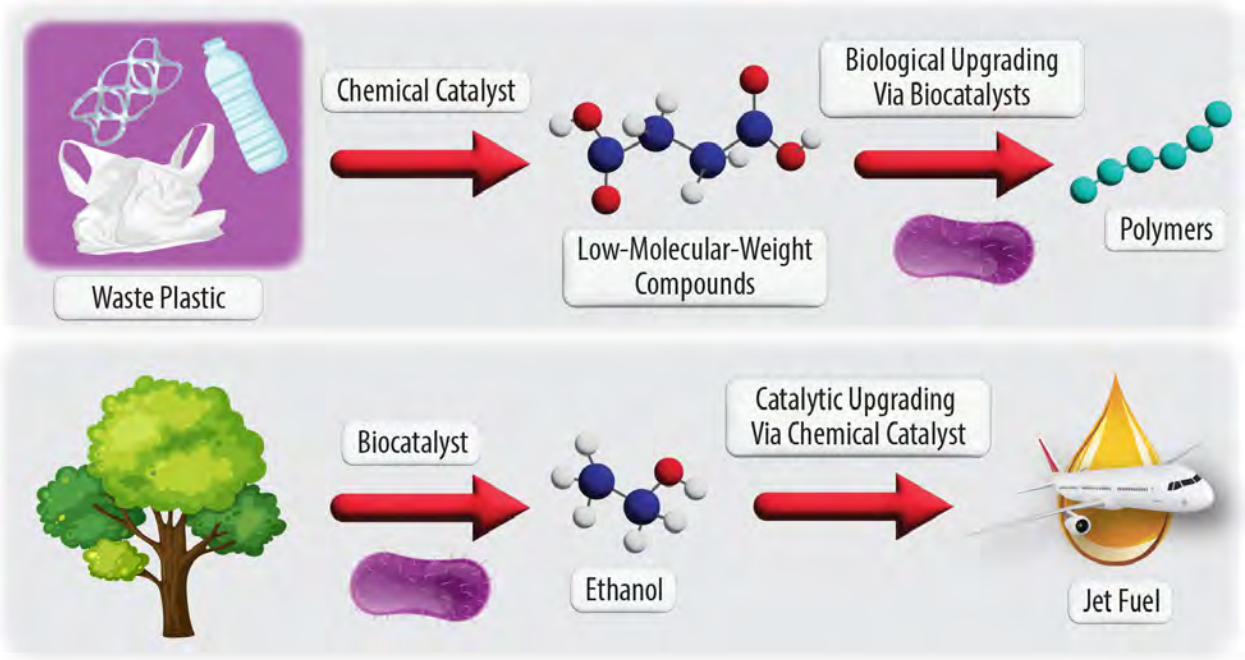


Fig. 2.3. Nonbiological Chemical Reactions Can Synergize with Microbial Pathways in Designed Hybrid Systems. Biology and chemistry excel at different reactions, such as in the different phases of the two hybrid-system examples shown here. Hybrid systems can combine steps in a bioproduct pathway using the microbe or nonbiological catalyst that most efficiently and effectively produces a desired reaction. For example, while chemical catalysts can degrade complex polymers like waste plastic better than microbes, microbes can perform well at converting the resulting low-molecular-weight compounds to final bioproducts **(top)**. Conversely, chemical catalysts can convert biological intermediates into final products like jet fuel **(bottom)**.

conditions (e.g., low temperatures and pressures) and can display exquisite stereochemical specificity. Nonetheless, chemical catalysis offers distinct advantages and can synergize with biological approaches in various ways (Sullivan et al. 2022). Several examples of challenges and opportunities highlight the need for interdisciplinary innovation at the interface of chemistry and biology.

In hybrid systems, integrating chemical steps may shorten biosynthetic pathways and reduce overall process complexity. For instance, catalytic deconstruction of polymeric feedstocks (e.g., waste plastic and lignocellulose components) can proceed much faster and more efficiently than biological depolymerization mechanisms, producing compound mixtures that are well suited for biological upgrading. Biosystems can also produce chemical intermediates rather than the final target molecule, which are then converted to final products via chemical catalysis, such as muconic acid conversion into adipic acid (Vardon et al. 2016) and ethanol conversion to jet fuel (Hannon et al. 2020). This division of labor enables biological processes to be applied to specific tasks that are most compatible with microbial metabolism, such as processing targeted deconstruction products or synthesizing key catabolic intermediates that feed into downstream catalytic steps.

However, to fully realize the potential of these hybrid approaches, fundamental research is needed to understand how best to integrate biology and chemistry into cohesive, efficient workflows that do not interfere with each other. This approach can be informed by combining chemical and enzymatic catalysis databases, including kinetic data, with AI-driven design tools and by developing hybrid processes to identify key innovation opportunities.

Chemical catalysis can be especially valuable when biological systems face thermodynamic barriers or pathway intermediates are too toxic for biological systems. For instance, electrochemistry can harness electrical energy to drive thermodynamically unfavorable reactions, such as carbon dioxide reduction into single carbon compounds or molecules with carbon–carbon bonds. The resulting electrochemically derived products can serve as substrates for bioconversion, yet the impact of integrating electrochemical processes into biological processes remains underexplored.

Another promising avenue for chemical catalysts overcoming thermodynamic barriers is the development of enzymes or other biological systems that are active and stable in organic solvents. Water is essential to living systems, but many reactions (e.g., dehydration and polymerization) are typically thermodynamically unfavorable in water. Advancing the ability to rationally design solvent-tolerant enzymes, biological systems, or entire organisms will be a key enabler for expanding biocatalysis beyond aqueous environments. Similarly, embedding enzymes or microbes in solid-state environments—or engineering localized aqueous microenvironments within nonaqueous systems—could allow biological functions to persist where water-based systems are otherwise limiting.

Chapter 3

Filling Key Knowledge Gaps in Microbial Genome Engineering

A second core workshop theme centers on microbial genome engineering, which involves deliberate DNA modification to endow microbial cells with novel functions or improved traits. This foundational technology for fundamental research and real-world applications underpins advances in areas such as biomanufacturing, biomaterials development, and systems biology (Nielsen et al. 2016; Lee et al. 2019).

Recent progress reflects a convergence of genetic tool development, data-driven discovery, and application-driven innovation, which enable more scalable and predictive engineering strategies (see Fig. 3.1, p. 18). Tools like recombinase-based genome editing, CRISPR gene regulation, high-throughput screening platforms, and biosensor-coupled selections support iterative strain improvement and pathway optimization. Advances in DNA synthesis and genome assembly enable large-scale combinatorial testing of regulatory architectures, while methods like massively parallel reporter assays and metabolic flux analysis provide fine-grained insights into genotype–phenotype relationships.

Collectively, these innovations enable researchers to link experimental results with predictive models, enhancing the capacity to design and deploy microbial systems with increasing complexity. This chapter presents current challenges and new paradigms in microbial genome engineering, as well as an extensive discussion of pathways to advance genome engineering as applied in microbial design.

3.1 Current Challenges in Microbial Genome Engineering

Despite major advancements in microbial genome engineering, persistent challenges remain in predicting genetic intervention outcomes, engineering multiple genomic loci simultaneously, and generalizing design strategies across microbial species. These limitations underscore the need for coordinated innovation in genetic tool development, data integration, and experimental standardization to achieve greater scalability and predictability across diverse biological contexts (see sidebar, Scalability and Predictability in Microbial Genome Engineering, p. 19).

At present, predictive models struggle to generalize across complex or nonmodel microbial systems, particularly when introducing large numbers of simultaneous genomic modifications. Engineering many genes or pathways at once often leads to context-dependent effects, unintended interactions, and reduced strain robustness.

Context dependence remains a significant barrier to genome engineering, even for the conceptually simple task of identifying stable genomic hot spots for consistent gene expression. Within complex nonlinear gene and metabolic networks, the broader engineering challenge of decreasing strain performance emerges as additional genes, circuits, or pathways are introduced. Current models often inadequately anticipate interactions between introduced components and host physiology, leading to unpredictable outcomes such as reduced viability or product yields.

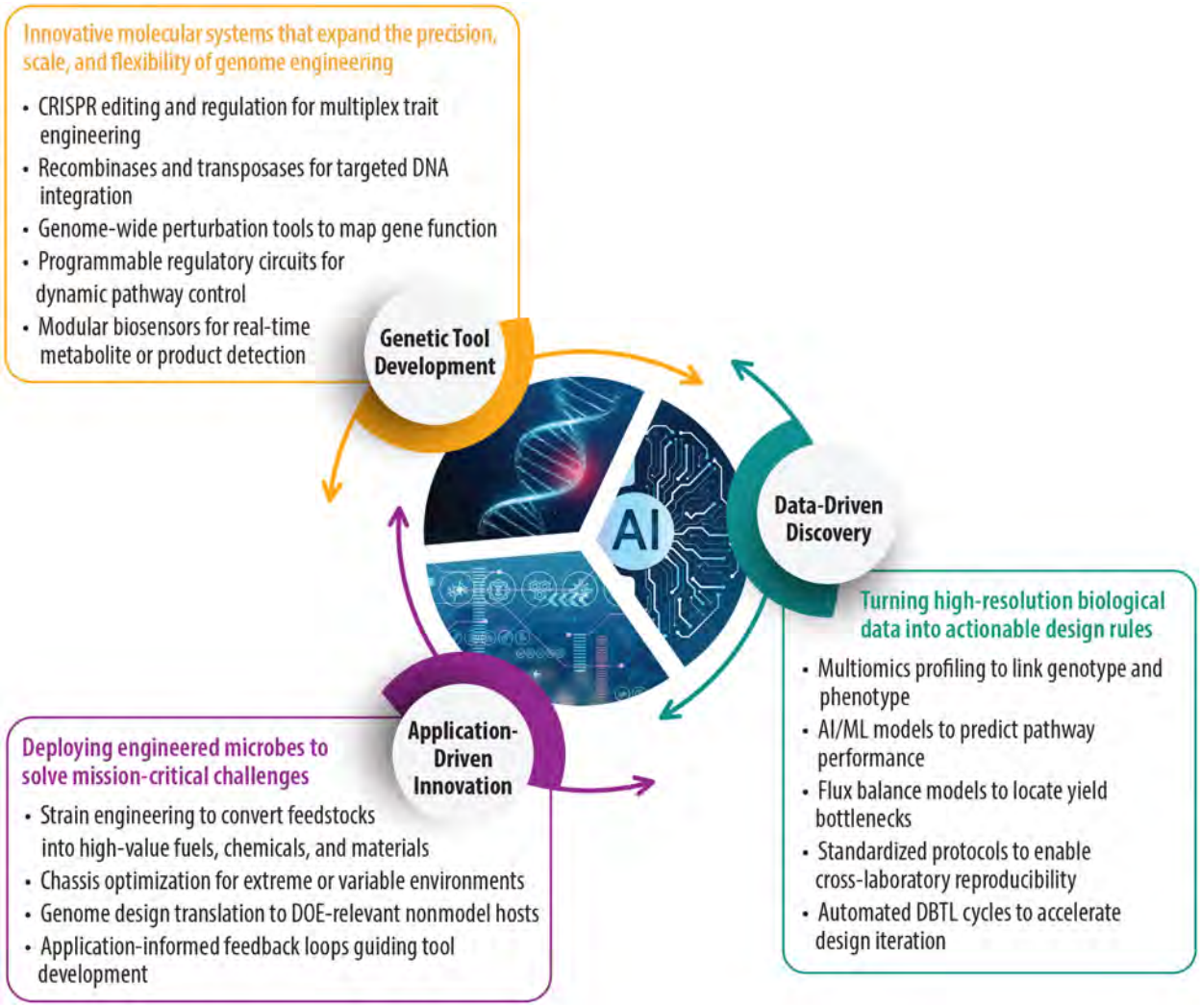


Fig. 3.1. A Diverse Toolbox Drives Microbial Genome Engineering. A convergence of progress in genetic tool development, data-driven discovery, and application-driven innovation underpins recent advances and future opportunities in microbial genome engineering (AI/ML: artificial intelligence and machine learning; DBTL: design-build-test-learn).

3.2 Emerging Paradigms in Microbial Genome Engineering

Several developments in microbial genome engineering collectively signal a transformative shift toward more scalable, predictive, and generalizable approaches, with profound implications for the future bioeconomy:

- Genome engineering is shifting from single-gene edits to multigene perturbations. This shift in outlook can enable more precise control over function, enhanced biocontainment, and industrial robustness.
- Integration of automation, multiomics analysis, and machine learning (ML) is driving development of AI-powered, self-driving experimentation to rapidly iterate through design-build-test-learn (DBTL) cycles and reduce the need for trial-and-error approaches.
- A growing emphasis on cross-organism generalization is producing tools and models that function beyond traditional model organisms, expanding the potential for production in new microbial hosts.
- Predictive, high-resolution engineering is becoming feasible by combining large-scale perturbation datasets and mechanistic models, enabling

Scalability and Predictability in Microbial Genome Engineering

Increasing the scalability and predictability of microbial genome engineering is key to advancing from exploratory experiments to robust, industrial-ready microbial platforms that meet the performance demands of the emerging bioeconomy.

Scalability in microbial genome engineering refers to the ability to perform many genetic modifications at once, enabling efficient construction of complex phenotypes. Genome modification strategies include multiplex editing, introducing genome-wide perturbations to gene expression, and combinatorial testing of regulatory elements and pathways. As researchers build increasingly sophisticated microbial systems, the capacity to scale genome engineering efforts becomes essential, not only for prototyping but for transitioning designs into robust production strains.

Predictability in microbial genome engineering centers on anticipating the outcomes of genetic interventions, such as how changes to gene content or regulation will affect cell growth, product yield, or system stability. A lack of predictability can result in costly trial-and-error cycles and reduce the reliability of engineered strains in real-world conditions.

Central to scalability and predictability challenges is a need for integrated modeling and data collection frameworks. Current predictive models often fail to account for context-dependent and nonlinear behaviors that emerge as more genetic components are introduced. For example, gene expression levels may vary depending on genomic insertion sites or interactions between synthetic and native regulatory networks. These effects are further complicated by dynamic environmental conditions found in bioprocessing settings.

To address these issues, the microbial genome engineering community is embracing design-build-test-learn (DBTL) cycles supported by high-throughput experimentation, multiomics datasets, and machine learning–guided design. Emerging engineering strategies include using CRISPR interference (CRISPRi) and activation (CRISPRa) for large-scale, tunable perturbations; self-driving biofoundries that rapidly iterate across design space; and mechanistically informed models that blend stoichiometric, thermodynamic, and kinetic constraints with data-driven inference.

multiplexed modifications with greater efficiency and genetic stability.

- Work on recoded, minimal, and synthetic genomes is revealing emergent behaviors (e.g., unanticipated regulatory dynamics) and opening new avenues for programmable chassis design.

3.3 Pathways To Advance Genome Engineering for Microbial Design

Advancing genome engineering for microbial design requires predictive, context-aware frameworks that integrate biological insight, computational modeling, and scalable experimentation. While foundational technologies such as CRISPR editing and gene regulation have transformed genetic manipulation, the next frontier lies in harnessing these tools for rational, multiplexed genome design that performs reliably across diverse hosts and environmental conditions (Jinek et al. 2012; Cong et al. 2013).

To realize these goals, future efforts must address persistent knowledge gaps in genome function, regulation, and system-level interactions. This includes developing methods to perturb and interpret gene networks at scale, constructing synthetic genomes with modular

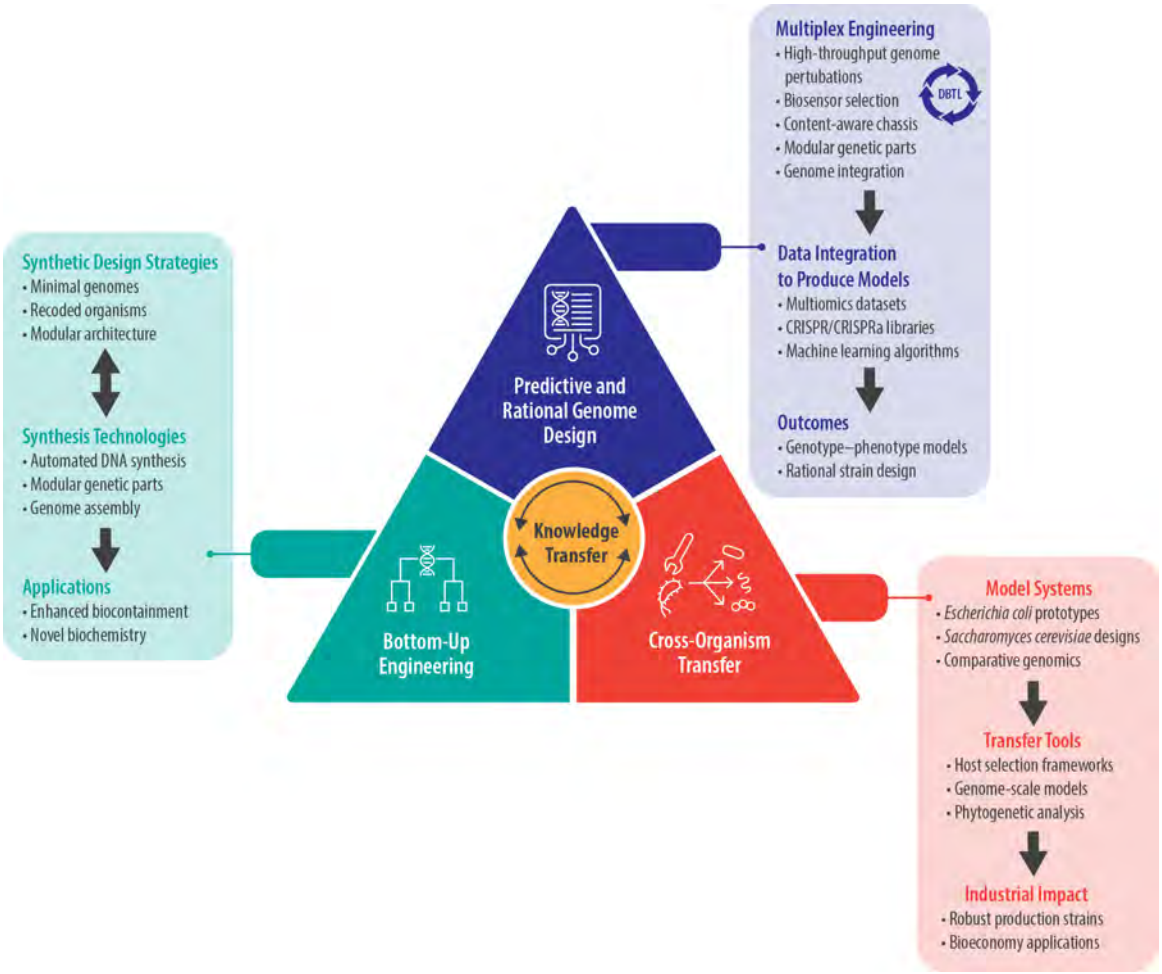


Fig. 3.2. Three Entry Points To Advance Genome Engineering for Microbial Design. Predictive and rational genome design **(A)** is fueled by knowledge gained from endeavors in bottom-up strain engineering **(B)** and developing tools that can be transferred from model organism systems in the laboratory to nonmodel organisms in industrial systems **(C)**. Predictive and rational genome design combines computational modeling with high-throughput perturbation to enable hypothesis-driven strain optimization through iterative design-build-test-learn (DBTL) cycles. Bottom-up genome engineering constructs synthetic genomes using a first principles approach to create modular chassis with programmable traits and enhanced biocontainment. Cross-organism transfer translates designs from model organisms to industrially relevant hosts through comparative analysis and predictive tools.

architectures, and translating design principles to industrially relevant nonmodel organisms.

Genome engineering advancements can be achieved through three key pathways that share the common goal of creating designable microbial systems supporting bioeconomy innovation (see Fig. 3.2, this page). The pathways are: developing predictive and rational genome design frameworks, constructing synthetic genomes through bottom-up engineering, and extending design principles to nonmodel microbial systems. All of these approaches will benefit from strengthening feedback loops between tool developers and end users to align innovation with application needs.

Predictive and Rational Genome Design

Microbial genome engineering is advancing toward integrated frameworks that combine computational

modeling, high-throughput genome perturbation, and systems-level analysis. This approach focuses on enabling rational, multiplexed genome design, where "rational" refers to deliberate, hypothesis-driven selection of genetic modifications based on predictive models of genotype–phenotype relationships. Rather than relying on trial-and-error, rational design uses mechanistic insight and computational guidance to target specific genes, pathways, or regulatory elements for modification.

ML models trained on multiomics datasets help prioritize modifications, while massively parallel perturbation platforms (e.g., CRISPRi and CRISPRa) and variant libraries support systematic exploration of genetic interactions (Dong et al. 2018). Multiplex genome editing technologies enable simultaneous introduction of dozens to hundreds of changes, accelerating trait optimization and revealing nonintuitive design rules. Control-theoretic approaches and feedback-regulated circuits add dynamic responsiveness, enabling engineered strains to adapt to changing environmental or process conditions (Boo et al. 2019).

As discussed in more detail below, these emerging capabilities form a foundation for enhancing precise, predictive genome engineering across diverse microbial hosts and industrial applications.

Emerging capabilities form a foundation for enhancing precise, predictive genome engineering across diverse microbial hosts and industrial applications.

Computational Genome Design

Computational models play an increasingly central role in microbial genome engineering by enabling *in silico* prototyping of genetic constructs, regulatory architectures, and metabolic pathways. These tools assist in selecting candidate modifications before empirical testing, which reduces reliance on trial-and-error approaches. Models also enable simulations that account for gene regulatory dynamics and emergent properties of engineered strains. Some groups are integrating cybergenetic and control-theoretic principles into design tools. This approach enables implementation of feedback-regulated circuits or adaptive genetic programs that respond dynamically to environmental or process cues. Though computational genome design is less mature than protein engineering, continued development and benchmarking are expected to improve prediction accuracy and make genome-scale design more tractable.

Predictive Metabolic Modeling

Genome-scale metabolic network models represent the complete set of biochemical reactions encoded by an organism's genome, linking genes to proteins and reactions (King et al. 2015; Gu et al. 2019). These models simulate how metabolic fluxes respond to genetic or environmental perturbations, offering a powerful foundation for rational strain engineering. When calibrated with multiomics datasets, these models can forecast the effects of gene knockouts, pathway insertions, or environmental changes, thereby informing rational design strategies.

However, the predictive accuracy of these models depends heavily on high-quality baseline information, including reaction stoichiometry, thermodynamic constraints, and regulatory interactions. Traditional genome-scale models are often based on steady-state assumptions (e.g., flux balance analysis), which can be limiting in dynamic or regulatory contexts.

To address these limitations, researchers are integrating kinetic constraints, enzyme capacity, and thermodynamic feasibility into expanded modeling frameworks that better capture non-steady-state behaviors, such as transient gene expression or adaptive metabolic shifts. Because many kinetic parameters are difficult to obtain experimentally, especially across diverse organisms, hybrid approaches are gaining traction. These approaches combine stoichiometric backbones with limited kinetic detail and are increasingly supplemented by ML algorithms trained on high-throughput datasets (Zhang et al. 2020). The fusion of mechanistic and data-driven methods aims to produce more robust and adaptable models capable of guiding multiplexed genome engineering and supporting a broader shift

from trial-and-error workflows to predictive and design-driven synthetic biology.

Multiplex Genome Modifications

Multiplex genome engineering strategies, where dozens to hundreds of edits or perturbations are introduced simultaneously, are critical for constructing complex phenotypes, rewiring regulatory networks, and building robust production strains. Emerging technologies aim to minimize off-target effects while increasing the number of simultaneous modifications.

Achieving high-efficiency multiplexing requires improving DNA synthesis, delivery, and recombination efficiency. In some cases, adaptive laboratory evolution (ALE) could be combined with multiplex edits or perturbations to stabilize engineered genotypes or uncover beneficial epistatic interactions. High-throughput, massively parallel experimentation provides the platform necessary to implement and evaluate large multiplex libraries efficiently, enabling rapid screening of variant combinations under controlled conditions. The ability to multiplex at scale could also support systematic exploration of genotype–phenotype landscapes, which would enhance understanding of cellular systems while accelerating the predictive accuracy of genome design and engineering efforts.

Developing a suite of chassis strains tailored to different bioprocesses may offer a way to accelerate industrial deployment of new bioproduction platforms.

Nonintuitive Targets

Advancing genome engineering requires systematic strategies to uncover inconspicuous genetic targets that drive valuable phenotypes. While canonical pathway genes often take center stage in metabolic engineering, many traits of interest emerge from complex regulatory or stress response networks.

Global genetic screens and genome-wide perturbation strategies can redirect focus from the engineer to the organism, enabling cells to reveal which genes are important. This approach is especially valuable when product traits can be linked to fitness, allowing selection pressure to expose beneficial variants that may not have been considered through rational design.

Genome-Wide Tools Connecting Genotype to Phenotype

Technologies such as CRISPRi, CRISPRa, Tn-seq, DAP-seq, and massively parallel reporter assays are now routinely used to probe the regulatory architecture of microbial genomes. These high-throughput perturbation platforms enable researchers to quantify the phenotypic impact of hundreds or even thousands of genetic changes in parallel, often under diverse conditions or selective pressures. When integrated with transcriptomic, proteomic, and metabolomic datasets, these approaches map gene essentiality and the broader context-dependent interactions that shape performance in production strains. Standardizing and expanding access to these tools, especially for non-model organisms, remains a key research need.

Chassis Optimization and Context-Aware Engineering

While identifying promising genetic targets is a critical step in microbial design, ensuring that those targets behave predictably in desired environments requires robust microbial chassis. Genome reduction combined with ALE and regulatory rewiring can potentially create chassis strains with reduced variability, streamlined regulation, and improved compatibility with engineered pathways. However, this process is far from routine. Fitness remains highly context-specific, with trade-offs between robustness, productivity, and versatility. Developing a suite of chassis strains tailored to different bioprocesses, with each tuned for specific environmental stresses, substrates, or production goals, may offer a way to accelerate industrial deployment of new bioproduction platforms.

Data Integration and Predictive Modeling

Even the best experimental approaches produce only fragments of insight for rational genome design unless the resulting data can be aggregated into predictive models. A recurring workshop theme was the need

Growing Role of Biosensors in Microbial Genome Engineering

Biosensors are genetic constructs that couple intracellular metabolite concentrations to easily measurable outputs like fluorescence or growth. By linking desired phenotypes to selectable or screenable signals, biosensor-coupled selection enables high-throughput screening of large variant libraries. This approach accelerates identification of beneficial microbial genotypes during strain optimization and pathway engineering.

Workshop discussions emphasized the importance of developing robust, modular biosensors that are tunable for sensitivity, dynamic range, and specificity. In particular, biosensors that report on intermediate or end-product accumulation are key to guiding adaptive laboratory evolution, enabling rapid enrichment of strains with improved productivity. Expanding the library of well-characterized biosensors, especially for nonmodel organisms and industrially relevant metabolites, remains a major opportunity to increase the pace and precision of microbial engineering efforts.

to incorporate multidimensional data into ML frameworks to discover emergent patterns and guide rational strain design. These AI-driven approaches can help overcome the limitations of purely empirical or heuristic workflows, especially when used in combination with standardized metadata. By linking perturbation results with mechanistic models and preexisting datasets, ML can help prioritize genetic targets for genome engineering and forecast downstream effects on metabolism, regulation, and viability.

New Assays and Selection Strategies

Selection-based enrichment strategies that link target phenotypes to measurable outputs, such as fluorescence, survival, or reporter signals, are important. These enrichment strategies can dramatically accelerate the discovery process by filtering top performers from large libraries. However, many traits of interest, especially those related to materials properties or complex metabolic profiles, lack simple screening methods. This technical limitation could be met by developing new biosensors and high-throughput assays (see sidebar, Growing Role of Biosensors in Microbial Genome Engineering, this page) or even leveraging cell-free systems for pathway prototyping (see Ch. 4, p. 27), which could also bring new classes of phenotypes into the fold of programmable genome engineering.

Bottom-Up Genome Engineering of Synthetic Microbes

Bottom-up genome engineering aims to construct microbial genomes using a first principles approach, enabling rational control over genetic architecture and whole-cell system behavior, including metabolism, regulation, and responsiveness to changing conditions. This approach applies to designing minimal genomes, recoded organisms, and fully synthetic chromosomes that support their modularity, biocontainment, and expanded biochemical capabilities (Gibson et al. 2010; Isaacs et al. 2011; Ostrov et al. 2016; Richardson et al. 2017).

By stripping genomes to their functional core or redesigning them for programmable traits, researchers can explore emergent properties, optimize regulatory logic, and build chassis tailored to specific production goals. These synthetic platforms offer testbeds for validating genetic tools, probing design constraints, and accelerating strain development. As synthesis technologies improve, including advances in DNA synthesis, genome assembly, and error correction, bottom-up engineering may unlock new routes to biomanufacturing and broaden the chemical space accessible to synthetic biology. These technologies include high-fidelity, large-scale DNA writing; enzymatic and chip-based synthesis platforms; and automated workflows for assembling and debugging complex genetic constructs. Together, they could reduce the time and cost needed to build custom genomes, enabling more ambitious designs and a more rapid transition from concept to functional microbial chassis.

Redesigning Organisms from First Principles

Bottom-up approaches allow organism redesign using a first-principles approach that starts with a target list of genetic parts and tailors a microbe's genetic architecture to serve specific functions in industrial, biomedical, and environmental contexts. One key approach is replacing redundant or ambiguous genetic codes with recoded genomes containing expanded systems that incorporate nonstandard amino acids or orthogonal replication and translation machinery. These systems not only enable new biochemical functionality but also offer enhanced biocontainment through genetic isolation, which limits the risk of horizontal gene transfer.

Minimal Genomes and Cells

Minimal genomes, generated through systematic genome reduction, distill an organism down to genes essential for persistence and growth. These streamlined systems serve as testbeds for understanding gene essentiality and the functional consequences of large-scale gene deletions. Reducing the metabolic load of microbial biocatalysts by eliminating genes related to nonessential growth or stress responses can shift cellular resources toward bioproduct synthesis, which improves product yield and production efficiency. Minimal cells also provide a controlled environment to investigate genetic robustness, mutation effects, and chromosome evolution.

However, these designs may exhibit emergent phenotypes reflecting previously hidden regulatory or metabolic interactions that genome content alone could not predict. Genome reduction may also eliminate adaptive pathways and stress responses important for long-term robustness, especially under the variable and stressful conditions of industry-scale production.

Synthetic Genomes and Cells

Building on the concepts of genome redesign and minimal genomes, synthetic genome design goes further by assembling entirely new microbial chromosomes *de novo*, often from chemically synthesized DNA without relying on a natural genome as a starting point. Synthetic genomes offer a route to fully customized organisms where metabolic pathways, regulatory networks, and genome organization are all subject to design. These synthetic constructs are not limited by evolutionary legacy and can include features such as genome rearrangements, modular control systems, and non-natural biochemistry.

A closely related concept to synthetic chromosome design is synthetic cell design. This approach constructs minimal cells or entirely artificial cells from engineered genetic components. The cells are designed to perform targeted functions, sometimes outside the context of a complete living organism. Synthetic cells may lack full reproductive capacity but are engineered to execute core metabolic tasks or serve as programmable biosensing and production platforms.

Improving Design Through Comparative Analysis and Modular Strategies

Progress in bottom-up genome engineering will depend on deep comparative analysis across diverse microbes and modular approaches that enable rapid prototyping. Together, these strategies can bridge current knowledge gaps and unlock methods for constructing novel, programmable, and robust synthetic organisms.

Phenotypic characterization of diverse microbial species needs to be expanded to uncover new traits and gene functions relevant to the bioeconomy. Comparative genomics and functional profiling can help identify key design rules and components that are transferable across organisms.

Some efforts have begun exploring the influence of chromosomal organization and spatial genome architecture on expression stability and circuit robustness, with the aim of integrating synthetic operons or modular control regions more predictably. This may support the development of a minimal set of organisms that collectively span a broad range of biochemical capabilities, facilitating modular pathway deployment and combinatorial designs. The use of cell-free systems to prototype chassis functions outside the constraints of full cells is accelerating this work, enabling iterative DBTL cycles without needing to construct full organisms.

Remaining Challenges and Long-Term Potential

Bottom-up genome engineering holds great potential in synthetic microbe engineering but significant hurdles remain. These include the scalability and

fidelity of genome synthesis, stability of transplanted genomes, and ability to rationally predict system-wide outcomes from modular components. Nevertheless, bottom-up approaches open new possibilities for rational chassis design, robust bioproduction, and deeper insights into the principles of life, which can provide a foundation for a new generation of synthetic biology and microbial engineering.

Applying Microbial Design Insights to Nonmodel Systems

Extending microbial design principles to nonmodel organisms is critical for translating genome engineering from laboratory prototypes to industrial applications. While model microbes like *Escherichia coli* and *Saccharomyces cerevisiae* enable rapid experimentation and analysis, their limitations in production settings demand porting functional genome designs into robust, fit-for-purpose hosts. This porting is often nontrivial, requiring an array of predictive models, comparative genomics, host selection frameworks, and effective genetic tools to guide cross-organism design and ensure functional deployment in diverse, DOE-relevant systems.

Translating Microbial Designs from Models to Applications

Effectively translating biological insights from model organisms to industrially relevant microbes is essential for expanding the impacts of microbial genome engineering beyond research settings. Model chassis organisms offer well-characterized platforms for prototyping genetic constructs, evaluating pathway designs, and debugging gene and metabolic regulatory architectures. These systems provide a testbed for rapid microbial design iteration, enabling researchers to optimize designs before transferring them into more complex or less predictable production strains.

However, this porting is rarely straightforward. Genetic parts, pathways, and regulatory elements often behave unpredictably across hosts due to differences in metabolism, stress response, and global regulation. As a result, many engineered functions that work well in model organisms fail when ported to nonmodel systems without further adaptation.

Genome-scale models can serve as a bridge to transfer insights learned from well-studied model organisms to poorly characterized or industrially relevant microbes.

Predictive Tools for Cross-Organism Genome Design

In microbial genome engineering, genome-scale models can serve as a bridge to transfer insights learned from well-studied model organisms to poorly characterized or industrially relevant microbes. Various strategies support informed transfer of genetic insights across the microbial tree of life.

For nonmodel systems, researchers increasingly utilize genome-scale models derived from phylogenetically related organisms or apply transfer learning techniques to adapt existing models when direct experimental data is limited. Expanding and validating genome-scale models across a wider range of microbial species is a pressing need, particularly for hosts with industrial relevance. In parallel, data-driven host selection tools are being developed to guide microbial strain choice based on process-specific requirements, helping to move beyond default reliance on traditional chassis organisms.

Closing the Loop Between Tool Developers and End Users

Effective genome engineering depends on sustained feedback between tool developers and end users. This ensures that emerging capabilities address real-world constraints and industrial priorities. Industrial needs, such as efficient bioconversion of low-cost feedstocks into high-value products, drive demand for targeted, scalable solutions. At the same time, advancements in platforms like CRISPR-based editing and regulation expand the design space across diverse microbial hosts. These and other emerging technologies are reshaping microbial genome engineering. Table 3.1 (p. 26) outlines key tools and approaches for expanding experimental capabilities, streamlining workflows, and accelerating strain development.

Table 3.1 Pivotal and Emerging Tools in Microbial Genome Engineering

Tool/Approach	Key Advantage	Impact on Genome Engineering
CRISPR-Based Editing and Regulation	Enables precise edits and programmable gene control	Broadly adopted for editing, silencing, and tuning gene function
Automated DNA Synthesis	Provides rapid, scalable creation of genetic constructs	Facilitates modular design and combinatorial libraries
Adaptive Laboratory Evolution (ALE)	Evolves strains under selection pressures for improved traits	Complements rational design in strain optimization
High-Throughput Screening	Screens large libraries for functional phenotypes	Enables data-rich inputs for machine learning–guided strain development
Metabolic Flux Analysis	Tracks pathway dynamics via isotopic labeling	Informs pathway engineering and metabolic optimization
Whole-Genome Sequencing	Detects unintended mutations or structural changes	Ensures validation and quality control
Computational Genome Design	Predicts desired phenotypic outcomes of edits and pathways	Reduces trial-and-error in engineering workflows
Self-Driving Biofoundries	Integrates automation, analytics, and AI and machine learning in iterative DBTL cycles	Accelerates DBTL loops at scale

Various experimental and computational tools accelerate genome design by enabling precise edits, predictive modeling, and automated strain optimization. These approaches support iterative design-build-test-learn (DBTL) workflows and expand the design space for microbial biocatalyst development.

Maximizing the impact of these tools requires interdisciplinary integration of biological insight, engineering design, and computational modeling. Predictive genome engineering will benefit from approaches that couple systematic perturbation of gene regulatory networks (e.g., through CRISPRi and CRISPRa) with multiomics data streams feeding ML models. Standardized fermentation systems and metadata ontologies—organized frameworks that define and relate experimental variables, biological entities, and data types—further enhance reproducibility and enable cross-study comparability.

Together, these capabilities form the foundation for multiplex-capable, predictive genome engineering frameworks that perform reliably across varied genetic architectures and environmental conditions, closing the loop from tool creation to practical deployment.

Chapter 4

Leveraging Cell-Free Systems for Bioproduct Synthesis and Microbial Design

Another core workshop theme was identifying opportunities to leverage cell-free systems in bioproduct synthesis and microbial design. Cell-free expression systems and single or multienzyme systems are biocatalysts with wide-ranging applications. For example, they can reconstitute known pathways, including:

- Biological machinery for cellular respiration and ATP synthesis (Biner et al. 2020)
- Signaling networks for two-component systems (Peruzzi et al. 2023)
- Metabolic pathways for formate-to-commodity chemicals (Chowdhury et al. 2024; Cardiff et al. 2025)

They can also create new-to-nature circuitry, including:

- Biomimetic cofactors (Richardson et al. 2020)
- CRISPR-based multilayer circuit control (Tickman et al. 2022)
- Biosensors (Jung et al. 2022)
- Controlled exchange of biological parts between synthetic cells (i.e., cell-free systems encapsulated by liposomes) in systems with low competition from cellular metabolism and regulation (Heili et al. 2024)

Cell-free systems enable the study of biological processes in a reductionist system. Such systems can reveal intrinsic limitations of the biological parts (e.g., enzyme activity) and networks (e.g., balancing of enzyme levels in metabolic pathways, cofactor utilization, and recycling) comprising chemical bioproduction pathways. These limitations ultimately affect biosynthetic productivity.

Understanding biological parts and networks has enabled successful engineering of microbes for improved chemical production. Workshop discussions on cell-free systems therefore focused on proposing strategies to leverage nonmicrobial or alternative biocatalysts to generate products and gain fundamental understanding of microbial design (see Fig. 4.1, p. 28). Strategies include using cell-free systems for high-throughput determination of enzyme properties, AI-driven metabolic engineering, and accelerating metabolic pathway design.

4.1 High-Throughput Determination of Enzyme Properties

Cell-free expression systems enable rapid enzyme production and characterization, which accelerates the elucidation of enzyme–function relationships beyond the pace of traditional methods that rely on cellular overexpression and purification. Characterization

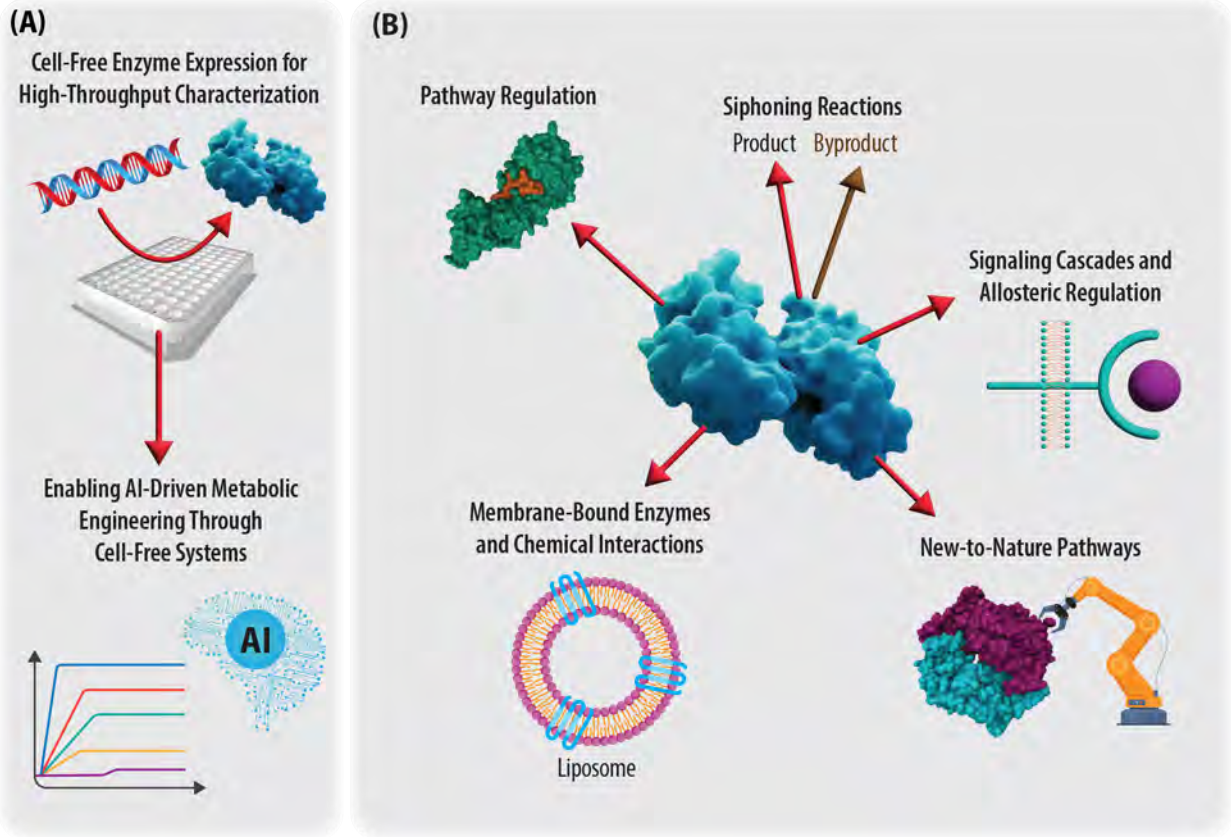


Fig. 4.1. Opportunities To Leverage Cell-Free Systems for Bioproduct Synthesis and Understanding Microbial Design. **(A)** High-throughput determination of enzyme properties using cell-free systems as a platform can advance artificial intelligence (AI)–driven metabolic engineering. **(B)** Cell-free prototyping provides opportunities to understand and explore pathway regulation, siphoning reactions, signaling cascades and allosteric regulation, membrane-bound enzyme activity and their chemical interactions, and new-to-nature pathways to accelerate metabolic pathway design before their implementation *in vivo*.

can now draw on complementary tools. Bioinformatics pipelines can assign putative functions to vast numbers of gene sequences based on homology and domain architecture. Large numbers of crystal structures have enabled AlphaFold to generate accurate structural models for enzymes without any structural data (Abramson et al. 2024). Multiomics data enables near-prediction of *in vivo* enzymatic activity of known enzymes (Heckmann et al. 2020). Machine learning (ML) approaches are getting closer to predicting enzyme function (Yang et al. 2024).

The ease and speed of experiments in cell-free systems could enable a leap in biocatalyst design by generating an expanded repertoire of enzymes characterized by differences in their catalytic activities, expression efficiencies, cofactor use, stability, and substrate and product profiles. This can be achieved while minimizing interference from background cellular reactions.

Coupling cell-free systems with mass spectrometry to expand the number of detectable and identifiable bioproduced compounds could permit rapid validation of enzymes whose substrate and product profiles are predicted by computational models (Yu et al. 2023). The resulting data could be standardized and integrated into AI/ML pipelines to accelerate the engineering of enzymes with user-specified properties (Landwehr et al. 2025).

If AI/ML enzyme models prove transferable across enzyme classes, the resulting data could help

deorphanize experimentally untested enzymes. Information on characterized enzymes would be accessible through chemical reaction and enzyme databases, such as the DOE Systems Biology Knowledgebase (KBase; kbase.us; Arkin et al. 2018). These detailed enzyme profiles would also support annotation and expansion of mathematical models of microbial metabolism (Shin et al. 2023). Looking forward, combining characterized enzymes with computational protein design could help refine enzyme properties via directed evolution.

4.2 AI-Driven Metabolic Engineering Through Cell-Free Systems

A key success of cell-free systems has been their use for rapid pathway prototyping leading to accelerated engineering of microbial systems (Liew et al. 2022). Metabolic pathway engineering in a simplified cell-free system has enabled kinetic pathway modeling and understanding of intrinsic system constraints (Martin, J. P., et al. 2023). Looking ahead, cell-free systems will accelerate applied AI/ML approaches to metabolic pathway engineering.

Vast amounts of standardized experimental data are needed to train ML models to generate successful microbial designs. Experimental microbial data not only take a relatively long time to acquire but are confounded by cellular background reactions and regulation of targeted microbial pathways. Cell-free systems produce data that are not convoluted by cellular background reactions and are potentially valid across multiple microbial systems. They can also accelerate the acquisition of data required to train the AI/ML algorithms used in microbial design and drive the next generation of microbial bioproduction platforms.

4.3 Metabolic Pathway Design Via Cell-Free Prototyping

Complete metabolic pathways can be prototyped in cell-free systems before the pathways are implemented *in vivo*. Cell-free systems are amenable to high-throughput screening using robotic platforms. For example, acoustic liquid handling platforms that use acoustic deposition, in which focused sound waves precisely transfer nanoliter-scale liquid droplets without physical contact, can dispense reaction volumes as small as 25 nanoliters.

This approach enables rapid evaluation of pathway performance, enzyme combinations, and cofactor dependencies. Such pathway prototyping builds mechanistic understanding of substrate and product inhibition and any in-pathway allosteric regulations that may limit biosynthetic productivity of the ultimate microbial biocatalyst.

A key success of cell-free systems has been their use for rapid pathway prototyping leading to accelerated engineering of microbial systems.

Cellular Siphoning

Metabolic pathway prototyping in cell-free systems can reveal how metabolic intermediates are diverted, or siphoned, by cellular metabolism, reducing overall biosynthetic productivity. The fate of in-pathway intermediates can be tracked using labelled substrates and microbial extracts or fractions thereof. This approach enables detection of bottlenecks where intermediates accumulate, competing reactions that siphon flux away, and steps where enzyme inefficiency or cofactor imbalances stall conversion.

This information facilitates targeted interventions. For instance, enzymes can be overexpressed or engineered to relieve bottlenecks, competing pathways can be knocked out or suppressed, and pathways can be redesigned to improve flux toward desired products. Metabolic intermediates could also be used as substrates to deorphanize putative enzymes by supplementing them with cell extracts (e.g., plant and nonmodel microbial extracts) and using metabolomics to identify resulting compounds.

Signaling Cascade and Allosteric Regulation

The build phase of metabolic engineering relies overwhelmingly on transcriptional regulation of pathway

genes and enzymes with desired biochemical properties (Ko et al. 2020). Adding a layer of nontranscriptional regulation often improves the productivity of these bioproduction platforms. For example, mimicking nature by increasing carbon flux through metabolic pathways, including enzyme scaffolding and subcellular compartment localization, can improve the biosynthetic productivity of a microbial biocatalyst (Liu et al. 2023).

Similarly, cell-free systems can elucidate cellular regulation, including signaling cascades and post-translational modifications, that could be confounded in a cellular environment. This information can be ported to engineered microbes to further improve their design.

Cell-free systems uniquely offer the ability to prototype new-to-nature metabolic pathway architecture.

Membrane-Bound Enzymes and Membrane–Chemical Interactions

Limited standardized information exists for membrane-bound proteins due to their location, variable expression, and difficulty to characterize experimentally. Cellular membranes contain transporters to pump toxic compounds and secrete bioproducts, receptors to collect input from the environment and modulate bioproduct formation, and enzymes for energy generation and chemical tailoring (e.g., ATP synthase, glycosyl transferases, and P450s).

Increased knowledge about these proteins, including their activity, substrate specificity, and stability, would support the design of next-generation metabolic pathways for chemical bioproduction. In particular, synthetic cells could be designed to express and experimentally evaluate these membrane proteins in isolation. Their liposome composition could be varied to mimic cellular stress or the membrane composition of different microbes.

Innovative Pathway Architecture for New-to-Nature Pathways

Cell-free systems uniquely offer the ability to prototype new-to-nature metabolic pathway architecture. This feature is unattainable in cellular environments where pathway regulation would hinder evaluation. Cell-free systems can prototype pathways that require toxic substrates or substrates that do not cross the cell membrane. They can also compare the activity of pathways requiring exotic cofactors not present in model organisms against standard pathways.

Adding Stepwise Challenges To Understand Intracellular Microbial Environments

Cell-free systems facilitate precise characterization of in-pathway metabolites, proteins, and mRNA levels with limited interference from microbial background reactions. This isolation makes cell-free reactions a great tool to understand microbial metabolism by adding one component at a time. For example, enzymes thought to siphon intermediates away from a pathway could be added one at a time to calculate their impact on overall bioproduct yield. In addition, the redox balance of a cell-free expression pathway can be disturbed to evaluate different cofactor recycling systems and cofactor siphoning to other fates.

Chapter 5

Basic Science Investigation into Biocatalyst Performance in Engineered Environments

Another core workshop theme covered basic science approaches to address challenges in microbial biocatalyst design for industrially relevant biotechnological setups (see sidebar, Engineered Environments, p. 32). Resilient and stable bioeconomies rely on a steady influx of new biotechnologies and bioprocesses. While these emergent tools frequently succeed in the laboratory, scale-up has proven a considerable challenge. The failure of a bioprocess to transition from the laboratory to an industrial scale or engineered environment, called the "valley of death," often results from uncertainties in the translatability of microbial performance (Crater and Lievense 2018; Biggs et al. 2021). In addition, traits such as product tolerance, stress resilience, and metabolic versatility often manifest only under process-relevant conditions that may differ significantly from laboratory environments.

The valley of death in biotechnology development represents a critical knowledge gap between basic scientific research and the testing of biological systems at large scales. This gap, which is often driven by proprietary constraints limiting publication, must be addressed to reduce the risks associated with scaling bioprocesses (Kampers et al. 2021).

Workshop participants discussed biocatalyst performance challenges in engineered environments, opportunities to address these challenges, and strategies for investigating and predicting biocatalyst responses within these environments.

5.1 Microbial Performance Challenges in Engineered Environments

In natural environments, microbial heterogeneity and evolution support community resilience, enabling adaptation to changing conditions, promoting metabolic cooperation, and helping maintain function over time through natural selection. In contrast, these traits often become challenges in engineered systems, where the focus is on stability, consistency, and achieving a specific production or performance outcome (see figure in sidebar, Engineered Environments, p. 33).

Microbial biocatalysts operating in engineered environments are not only exposed to system heterogeneities but also extended cultivation periods and accumulation of extracellular or intracellular products. These factors can significantly impact microbial function and enzyme activity, resulting in inconsistent biocatalyst performance across time and scale. Several biological factors contribute to inconsistent microbial performance in engineered environments.

- **Intrinsically Variable Cellular Function and Growth Kinetics:** As cells transition through different growth phases (e.g., exponential and stationary), their metabolic activities and stress responses can shift dramatically. This variability

Engineered Environments

An engineered environment is a human-designed system with abiotic conditions tailored to support and enhance microbial and enzyme biocatalyst performance for biotechnological applications. Understanding the impacts of abiotic conditions on biocatalysts, both at the outset of an innovation cycle and throughout bioprocess development, fundamentally guides effective microbial design efforts for bioproduct generation and other biotechnological applications. Additionally, improving the predictability of microbial performance in engineered environments will critically accelerate biotechnological process development and deployment.

Engineered environment systems take various forms and can be designed to enable specific biological activities, such as biopolymer deconstruction, bioproduct generation, and engineered living material development.

- **Bioreactors:** Bioreactors are common hardware systems used to sustain microbial activity and range in scale from milliliter volumes to hundreds of thousands of liters. They enable monitoring and control of key parameters such as pH, gas composition, and temperature, and vary in geometry and configuration to accommodate different types of substrates (e.g., liquids, gases, and solids).
- **Solid-State Cultivation Platforms:** These engineered environments support biocatalysts that function optimally in static conditions and high solids content. Filamentous fungi cultivations provide a typical example, where substrate deconstruction can occur simultaneously with mycelia production.
- **Membrane Systems:** Equipment that enhances product recovery from microbial cultivations includes membrane systems. These engineered systems can introduce additional abiotic conditions that may impact biocatalyst performance (Salvachúa et al. 2021).

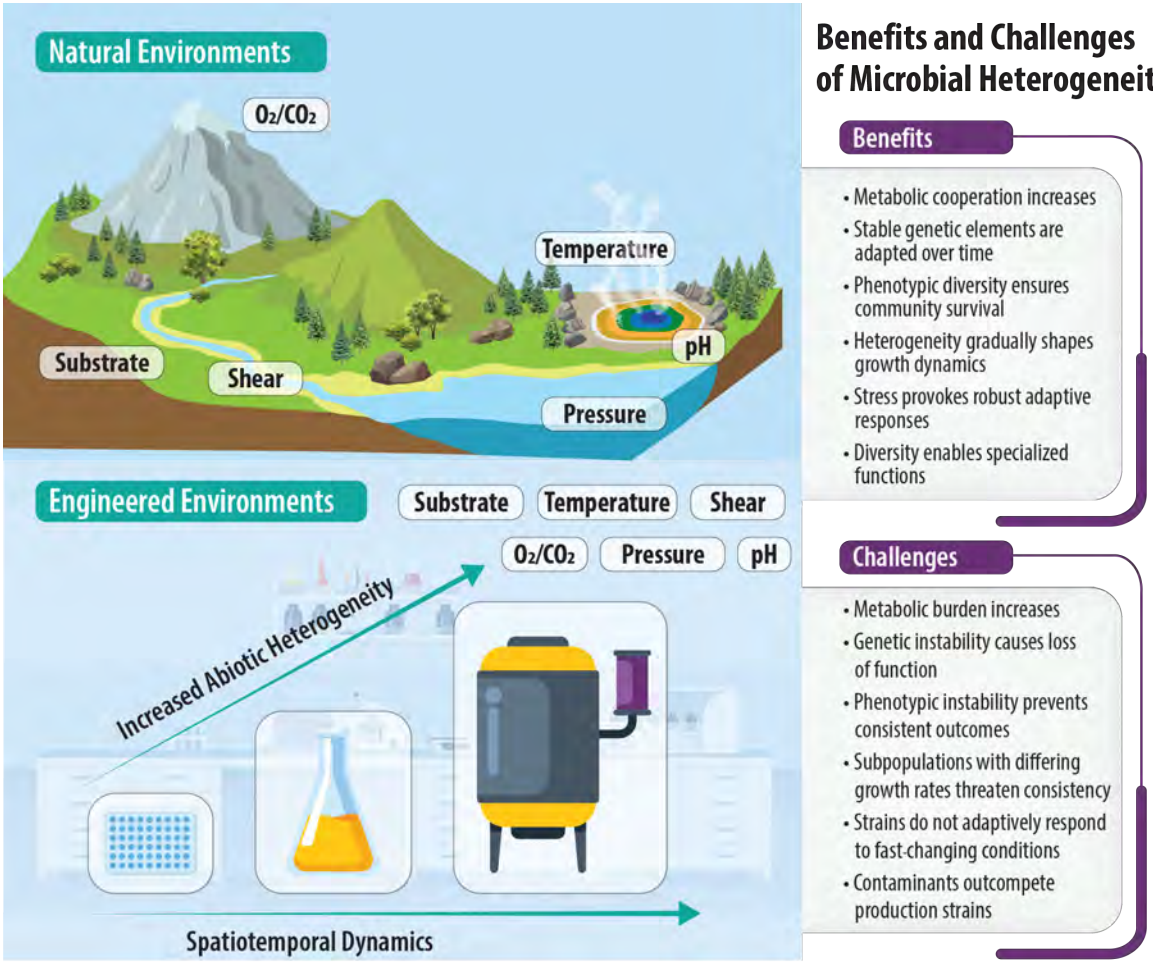
Although bioreactors, solid-state cultivation platforms, and membrane systems have been used for decades, design challenges persist, particularly managing nonhomogeneous regions such as gradients in nutrients, biological products, pH, gas composition, pressure, temperature, and shear (see figure, p. 33). These gradients often arise from inefficient mixing and pressure differentials and become more pronounced as cultivation scales up. Nonhomogeneous regions can alter the kinetics of biological processes in unpredictable ways, ultimately reducing bioconversion efficiency and increasing manufacturing costs (Lara et al. 2006; Delvigne et al. 2014, 2017).

Continued on next page

introduces selective pressures and uneven performance across the cell population, ultimately undermining microbial strain stability and reliability.

- **Genetic Instability:** Microbial strains often accumulate mutations during long-term cultivation or scale-up processes and sometimes lose engineered functions or undergo unanticipated metabolic changes.
- **Phenotypic Instability:** Genetically identical cells can display diverse behaviors due to random fluctuations in gene expression and local environmental differences.
- **Lack of Adaptive Regulatory Mechanisms:** Engineered microbes often lack dynamic systems that enable adjustment to environmental shifts, disadvantaging them when key parameters change such as nutrient levels, pH, and oxygen

Continued from previous page



Abiotic Heterogeneity Can Produce Opposite Outcomes in Natural and Engineered Environments. Similar abiotic factors influence both natural and engineered environments. However, microbial heterogeneity in natural systems buffers heterogeneous abiotic factors by enabling microbial communities to adapt to diverse stresses and conditions in the long term. In contrast, abiotic heterogeneity in engineered systems can hinder consistent microbial performance and productivity in the short term, which are key requirements in bioeconomic contexts.

concentration. Without such flexibility, production efficiency often declines under relevant bioprocessing conditions.

- **Metabolic Burden:** Engineered heterologous pathways can divert vital resources away from microbial growth and maintenance pathways, resulting in reduced productivity, pathway shutdown, and accumulation of toxic intermediates.
- **Contamination:** Bacteriophages, wild-type strains, and other microorganisms present in engineered environments can disrupt production processes or outcompete engineered strains, presenting significant risk in large-scale systems by compromising yield and reproducibility.

5.2 Opportunities To Advance Consistent Microbial Performance in Engineered Environments

Understanding the basis of inconsistent microbial performance in engineered environments and addressing associated challenges requires an integrated strategy. Mechanistic insights can be combined into microbial systems with systems-level understanding while also leveraging strain engineering and bioprocess design. Workshop discussions centered on ways to identify and solve sources of inconsistency that alter cell resource allocation and microbial performance (see Fig. 5.1A, p. 35). Four topics emerged: generating evolutionarily stable genetic backgrounds, achieving phenotypic homogeneity, developing strategies for dynamic microbial control, and understanding cellular energy balances.

Ensuring the long-term genetic stability of engineered microbes is critical for maintaining their designed functions and achieving reliable product yields.

Evolutionarily Stable Genetic Backgrounds

Ensuring the long-term genetic stability of engineered microbes is critical for maintaining their designed functions and achieving reliable product yields. Various approaches can be explored to develop stable microbial genetic backgrounds.

- Understanding how synthetic pathways interact with host genomes over extended cultivation periods.
- Elucidating emerging evolutionary dynamics under selective pressures such as metabolic burden, nutrient limitation, and environmental fluctuations.
- Determining how chromosomal integration sites influence pathway stability and performance.
- Building a comprehensive map of safe harbor genetic insertion sites that support long-term function without compromising host fitness.
- Investigating genome stability from the perspective of transposons and insertion sequences.
- Assessing the frequency and mechanisms of gene exchange, particularly in open or semi-open bioprocessing environments where horizontal gene transfer may generate loss or acquisition of engineered traits and potentially destabilize production strains.
- Understanding genetic resistance to phage infection.

Phenotypic Homogeneity

Achieving phenotypic homogeneity, with uniform cellular responses and minimal emergence of divergent subpopulations, is essential for maintaining consistent microbial performance in engineered environments. However, several knowledge gaps remain in understanding the design principles that govern phenotypic robustness at the cellular level.

- Understanding how environmental heterogeneity (e.g., gradients in oxygen, pH, nutrients, and product concentrations) interacts with stochastic gene expression to generate diverse cellular states.
- Identifying specific thresholds and cellular triggers that drive phenotypic divergence.
- Understanding how stress response pathways influence the distribution of metabolic fluxes under dynamic conditions.
- Exploring the role of quorum sensing, a mechanism of cell-to-cell communication that can significantly shape phenotypic outcomes in dense cultures (Darch et al. 2012; Moreno-Gámez et al. 2023). Fundamental questions remain about how quorum sensing systems interact with synthetic metabolic pathways, whether they can be tuned predictably in non-native hosts, and how they contribute to or mitigate population-level heterogeneity.

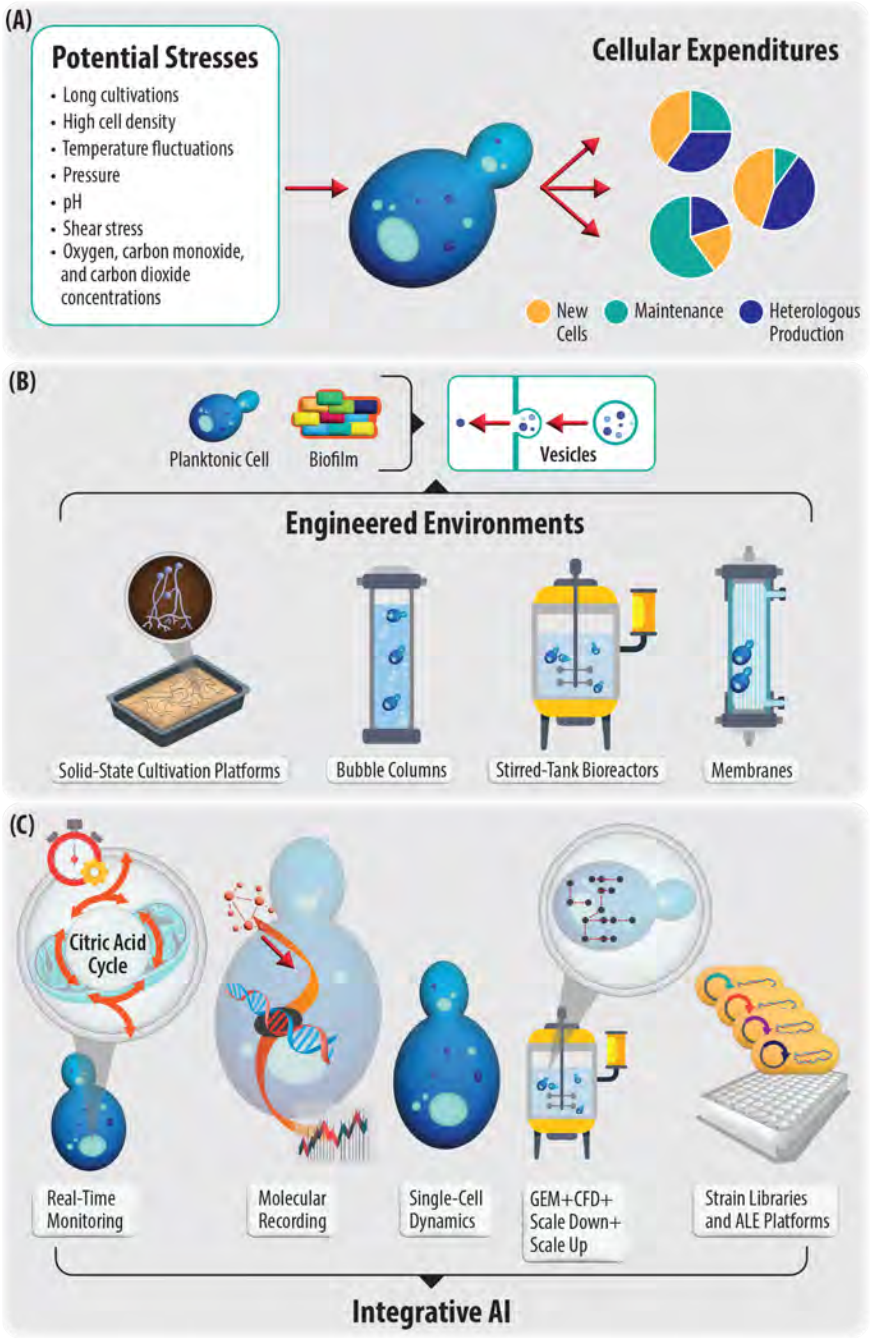


Fig. 5.1. Challenges and Strategies in Achieving Consistent Microbial Performance in Engineered Environments. **(A)** Inconsistent performance under different stresses can change how cells allocate resources toward pathways for new cell generation, cellular maintenance, and heterologous production (i.e., desired products; Cordell et al. 2023). **(B)** Microbes can adopt different lifestyles (e.g., planktonic or biofilm) depending on their environmental context. Both lifestyles can produce extracellular vesicles, which serve diverse functions such as communication, nutrient exchange, and stress response. Engineered environments may significantly influence both microbial lifestyle and vesicle production dynamics, but these effects remain largely unexplored. **(C)** Strategies for analyzing, controlling, and predicting microbial responses to stress in engineered environments include real-time monitoring of genomic and phenotypic changes, molecular recording of cellular events, single-cell dynamics, genome-scale metabolic modeling (GEM) combined with computational fluid dynamics (CFD) and scale down, and strain libraries and adaptive laboratory evolution (ALE) platforms.

Understanding how metabolic pathways influence microbial phenotypic and genotypic stability is essential to optimizing microbial performance, particularly under fluctuating conditions.

Dynamic Microbial Control

Systems-level dynamic microbial control strategies (e.g., feedback-regulated gene circuits and environmental sensors) provide powerful approaches for real-time modulation of microbial behavior, enabling adjustments to fluctuating bioprocess conditions. However, widespread implementation of dynamic control mechanisms remains constrained by fundamental gaps in understanding the complex interplay between engineered control elements and a host's native regulatory networks. A deeper, systems-level understanding of microbial regulatory networks is needed, particularly how microbes sense and respond to engineered environments and intracellular perturbations during cellular growth and production phases.

Cellular Energy Balances

ATP-positive pathways are metabolic processes that produce a net increase in available cellular ATP beyond what is required for basic cellular maintenance and survival. These pathways can enhance microbial productivity by providing additional energy for biosynthetic activities, stress responses, and overall cellular function. In contrast, cellular energy balance maintenance involves ensuring sufficient ATP supply to meet energy demands during bioprocesses and is fundamental to achieving consistent microbial performance. Understanding how metabolic pathways that generate varying levels of ATP influence microbial phenotypic and genotypic stability, as well as the cell's overall energetic and redox states, is essential to optimizing microbial performance, particularly under fluctuating conditions.

5.3 Opportunities To Decipher Microbe–Environment Dynamics Using Engineered Environments

A foundational knowledge gap in microbial design is understanding how the physical and operational structure of engineered environments impacts the consistency of microbial performance. Addressing this gap requires multidisciplinary insights across multiple dimensions.

One important dimension is the influence of growth and production chambers on microbial behavior. For example, microbial physiological responses can vary significantly between solid-state platforms, bubble columns, and stirred-tank bioreactors, each presenting distinct environmental conditions (see Fig. 5.1B, p. 35).

Substrate form also plays a critical role. In some systems, microbial biocatalysts interact directly with complex solid substrates (e.g., lignocellulose and synthetic polymers like plastics) during simultaneous deconstruction and conversion. In others, microbes encounter soluble intermediates produced through upstream chemical pretreatment, which present different metabolic demands.

Auxiliary equipment designed to support cellular activity and facilitate product recovery can also affect microbial stability and long-term performance. For instance, membranes that retain solid substrates or microbial biomass can enhance conversion efficiency and concentrate cells, while also integrating with downstream separation systems (e.g., liquid–liquid extraction units and adsorption columns). The impact of this artificial cell concentration mechanism on microbial metabolism remains poorly understood.

Interactions between microbes and engineered environments become even more complex when factoring in specific cellular lifestyles. For example, under typical bioprocessing conditions, some cells may grow as biofilms or in planktonic form. These distinct growth modes offer valuable opportunities to explore differences in transporter activity, metabolic cooperation, and stress responses across varied conditions within engineered systems.

Additionally, many bacteria produce extracellular vesicles, which can serve roles in selective product separation, metabolite compartmentalization, and enzyme localization. These functions can become relevant in engineered systems where spatial organization can enhance metabolic efficiency and overall system performance.

5.4 Strategies To Analyze, Control, and Predict Microbial Responses in Engineered Environments

Gaining a comprehensive understanding of how microbes respond to engineered environments requires examining their physiological behavior across molecular, cellular, and population scales. In addition, developing predictive and controllable systems will require an integrated suite of biological, analytical, and computational tools that not only advance fundamental understanding but also enable real-time decision making within bioprocesses (see Fig. 5.1C, p. 35). Although many of these tools were developed in the context of engineered environments, they hold value for investigating microbial behavior in natural ecosystems more broadly.

Key strategies and toolsets for predicting and controlling microbial responses in engineered environments identified during the workshop include:

- **Real-Time Monitoring of Genomic and Phenotypic Changes:** Understanding and addressing microbial performance consistency requires tools that detect genome instability and phenotypic shifts during cultivation. Real-time genomics and phenomics coupled with adaptive process controls (see sidebar, Digital Twins to Revolutionize Microbial Design, p. 13) offer promising approaches to fill knowledge gaps in microbial performance consistency.
- **Molecular Recording:** Molecular recording tools present a promising frontier in understanding microbial behavior in engineered environments (see sidebar, Molecular Recording, p. 38). This approach would enable lineage tracking and retrospective insight into single-cell experiences during cultivation.
- **Integrated Single-Cell and Systems Biology Approaches:** Phenotypic heterogeneity is a major driver of performance inconsistency. Bridging this gap will require high-resolution, single-cell tools and high-throughput phenotyping integrated with systems biology and bioprocess data to map the origins and consequences of microbial population variability.
- **Integrated Genome-Scale Metabolic Models with Computational Fluid Dynamics:** Computational fluid dynamics remains a well-established method for visualizing spatial and temporal gradients in bioreactors and other engineered systems. Integrating computational fluid dynamics in relevant engineered environments with genome-scale metabolic models represents a promising yet underexplored approach to enhance the predictive power of computational models. This integration can yield deeper insights into how spatial and temporal fluctuations influence microbial behavior and metabolic activity, and it may serve as a key component of future digital twin frameworks (see sidebar, Digital Twins to Revolutionize Microbial Design, p. 13).
- **Scaled-Down Engineered Systems and Systematic Physiological Profiling:** Small-scale, high-throughput platforms replicating the environmental complexity of industrial bioprocesses are essential for generating relevant, translatable data. These systems enable robust model development, facilitate hypothesis testing, and improve prediction accuracy under realistic process conditions.
- **Scaled-Up Engineered Systems:** Access to highly instrumented pilot-scale cultivation platforms capable of tracking a wide range of biological and operational parameters is essential for linking microbial performance to bioprocess conditions. These systems can provide data to inform and validate predictive models and digital twins.

Molecular Recording

Molecular recording systems in microbes are typically synthetic biology tools that encode information as molecular changes, often using programmable enzymes like CRISPR-associated systems, recombinases, or DNA polymerases. These systems act like biological black boxes, allowing microbes to record exposure to environmental signals (e.g., toxins, nutrients, and pH shifts), track gene expression or metabolic states, log evolutionary or lineage events over time, and store digital or analog biological data directly in DNA. Examples of different systems include:

- **CRISPR-based memory systems** record the presence of specific molecules by integrating short DNA sequences into a cell's genome (Shipman et al. 2016).
- **Recombinase-based systems** flip or excise DNA segments in response to a signal, creating a heritable memory of the event (Roquet et al. 2016).
- **Synthetic transcriptome/logging tools** capture temporal information using barcoded RNA expression patterns (Bhattarai-Kline et al. 2022).

- **Strain Libraries Coupled to Adaptive Laboratory Evolution (ALE) Platforms:** Microbial or strain variant libraries can be used in miniaturized systems to systematically explore boundary conditions such as pH, substrate concentration, and stress tolerance. These libraries enable interrogation of diverse phenotypes to identify traits associated with robust or poor performance and can aid in functional gene annotation. When coupled with ALE platforms, particularly those designed to simulate repeated bioreactor passaging, this approach provides a powerful framework for tracking evolutionary trajectories and uncovering underlying factors that influence biocatalyst behavior at scale.
- **Integrative Artificial Intelligence Models:** Developing models that incorporate the factors outlined above would generate a robust and comprehensive dataset that enables more accurate predictions of population behavior in relevant engineered environments. Capturing and analyzing negative data from environments that trigger microbial inconsistencies are critical steps to producing predictive models.

Chapter 6

Technologies To Accelerate Microbial Design

Remaining technology gaps and opportunities to accelerate microbial design emerged as another central workshop theme. Rapid technological change in the last decade has brought significant progress and potential applications in microbial design. Recent advances include new molecular biology techniques enabling precise gene editing, automation platforms for scaling DNA synthesis and assembly, omics for high-throughput phenotyping, and AI for efficient big-data analysis and design. Strategic integration of such technologies into research workflows could solve many current obstacles to accelerated microbial design.

Workshop discussions focused on identifying key gaps and exploring emerging technologies with the potential to transform microbial design. The discussions investigated limitations and opportunities in high-throughput platforms and design automation, omics and data-rich phenotyping, visualization and tools for design interpretation, and AI and machine learning (AI/ML) integration for predictive modeling and workflow optimization.

6.1 High-Throughput Technologies

High-throughput technologies relevant to the microbial design cycle include experimental and computational platforms that enable rapid, parallelized data generation and analysis. These are tools for miniaturized and automated assays, real-time sensing, dynamic control, and scalable phenotyping. When coupled with AI/ML approaches, such technologies could dramatically accelerate the design-build-test-learn (DBTL) cycle by optimizing experimental workflows, reducing iteration time, and enabling predictive design.

The following subsections explore specific opportunities to advance microbial design through high-throughput data collection, AI/ML integration, assay standardization, and self-driving biofoundries.

Expediting Design-Build-Test-Learn Cycle Data Collection

Several emerging technologies can disruptively accelerate the DBTL cycle (see Fig. 6.1, p. 40).

- **Biosensors** can facilitate real-time monitoring of metabolic states and potentially stress levels, supporting both fundamental insights and improved control of microbial populations in bioprocesses (Rogers et al. 2016).
- **Cell-free systems** comprised of lysates or purified components and used to protype genetic circuits and metabolic pathways outside living cells can reduce biological complexity and accelerate design iterations.
- **Cell imaging** is anticipated to yield insights into cellular behavior and stress responses (e.g., associating cell shapes with cell health).
- **High-throughput fluxomics** that uses isotopic tracers and automated analysis to quantify intracellular metabolic fluxes is expected to provide new insights into pathway activity and bottlenecks across engineered strains.

- ***In vivo* nuclear magnetic resonance** permits real-time, nondestructive monitoring of microbial metabolism by measuring chemical signatures within intact cells.
- **Microfluidics** enables high-throughput screening and phenotyping (e.g., in mimicked engineered environments) by miniaturizing and automating assays (Gach et al. 2017).
- **Optogenetics** uses light signals to control gene expression. It is expected to accelerate variant testing by providing dynamic gene expression control without needing to rebuild strains (Zhao et al. 2018).
- **Robotics platforms** that automate microbial strain engineering, liquid transfers, and experimental workflows are expected to enable scalable and reproducible execution of DBTL cycles.

Accelerating Design-Build-Test-Learn with Artificial Intelligence and Machine Learning

Several AI/ML approaches can leverage experimental data to accelerate the DBTL cycle.

Active Learning

In microbial engineering, design space refers to the multidimensional landscape of tunable parameters, including genetic variants, media compositions, and environments, that impact microbial functions. These spaces are often vast, nonlinear, and sparsely sampled, making it difficult or impossible to perform exhaustive searches.

Generally, the more experimental samples that are sent through a DBTL cycle, the higher the probability of obtaining successfully engineered microbes. However, the exact number of samples needed to achieve success

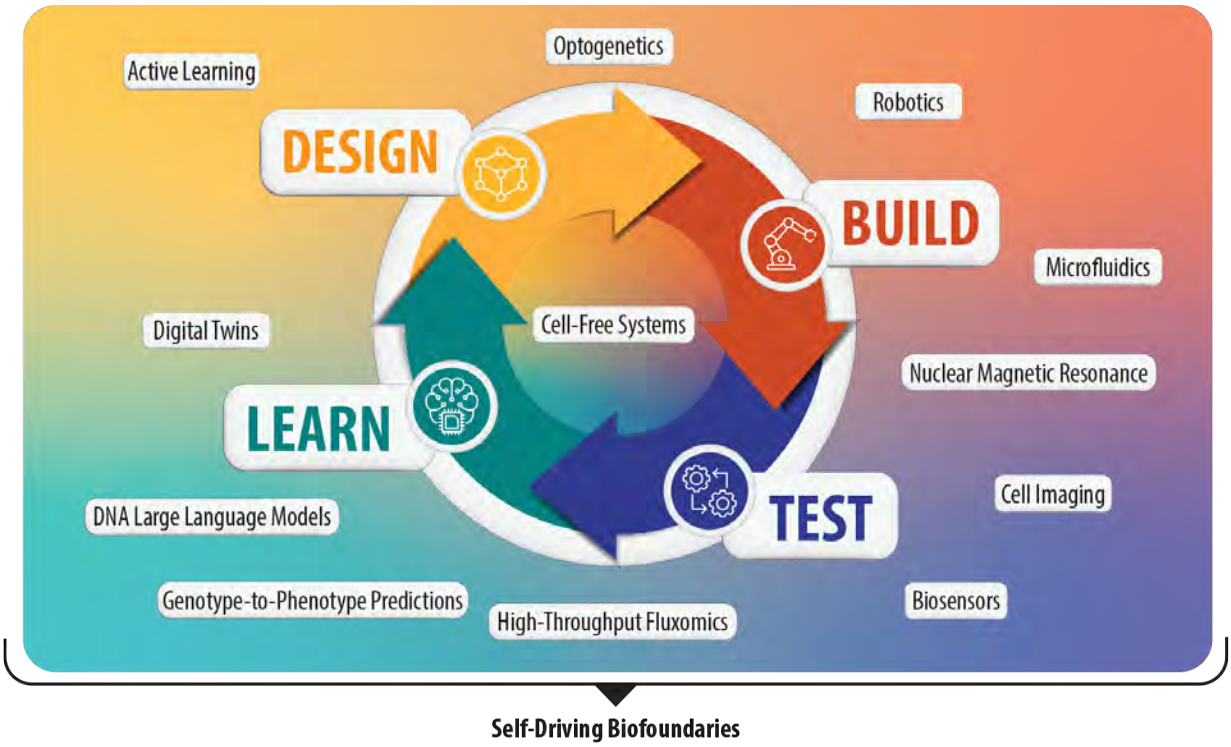


Fig. 6.1. Numerous Opportunities Exist To Advance High-Throughput Microbial Design. New technologies are accelerating the design-build-test-learn (DBTL) cycle by increasing throughput and expanding access to diverse cellular data. Artificial intelligence and machine learning approaches can leverage these data to streamline design workflows and lay the foundation for self-driving biofoundries.

is highly context-dependent and is influenced by both the biological system (e.g., cell-free pathway versus gut microbiome) and the final objective (e.g., ensuring growth versus achieving a 99% theoretical yield of a desired project).

Active learning through Bayesian experimental optimization offers a data-efficient strategy to explore large design spaces (Pandi et al. 2022; Zournas et al. 2025). By iterating through the most informative experiments, active learning reduces the number of required DBTL cycles and can drive convergence toward optimal designs.

Digital Twins

Digital twins are dynamic, computational representations of microbial cells that are continuously updated through feedback loops with experimental data (see sidebar, Digital Twins to Revolutionize Microbial Design, p. 13). Unlike static models, digital twins are designed to simulate cellular behavior across scales, ranging from well plates to bioreactors, and to iteratively refine predictions based on observed outcomes. In the context of microbial design, a digital twin integrates mechanistic knowledge and real-time measurements to forecast how a cell will respond to genetic edits, environmental changes, or process conditions. If digital twins could be quantitatively predictive, then hypotheses could be tested *in silico* before committing resources to physical experiments. When effectively implemented, digital twins could greatly accelerate DBTL cycles, improve design reliability, and support development of self-driving laboratories.

DNA Large Language Models

DNA large language models (LLMs) are AI systems trained on extensive biological datasets including DNA sequences, annotations, and experimental outcomes and are used to predict functional properties and guide design decisions. In microbial engineering, DNA LLMs can accelerate the DBTL cycle by narrowing the design space, identifying promising genetic variants, and reducing exhaustive screening. These models are already demonstrating utility in protein engineering, where initial predictions often yield improved variants without requiring large-scale mutagenesis. However, DNA LLMs are still maturing and must be carefully validated. Their performance depends heavily on the quality and relevance of training data, and their predictions may not generalize across different organisms, experimental contexts, or design goals. Transferability remains a key challenge, especially when shifting from well-characterized systems to novel applications or nonmodel microbes. Despite these limitations, DNA LLMs represent a promising tool for accelerating microbial design by using existing biological knowledge and enabling more efficient exploration of genetic possibilities.

Digital twins could greatly accelerate DBTL cycles, improve design reliability, and support development of self-driving laboratories.

High-Speed Phenotyping

High-speed phenotyping is central to scaling up microbial genome engineering because it links phenotype to genotype, rapidly identifies high-performing strains, and can provide a type of comprehensive understanding needed to predictively model cellular states. Promising targets for increasing phenotyping speed, defined here as measurement capacity or the ability to capture large amounts of diverse, high-resolution biological data across many samples and conditions, include robotics, microfluidics, biosensors, and bioimaging.

Measurement capacity encompasses not only the number of samples that can be simultaneously processed but also the range of traits (e.g., growth rate and metabolite concentrations), the frequency of measurements over time, and the relevance of those measurements to engineering goals. These techniques enable testing of vast combinatorial libraries of genetic variants, often in formats that can be miniaturized and automated.

Another important characteristic of high-speed phenotyping is translatability, which is ensuring that measured phenotypes faithfully represent cell state during large-scale production. Translatability could be improved

by developing miniaturized bioreactors that mimic the conditions cells experience in engineered environments.

Finally, several recent developments can maximize a system's genetic diversity while minimizing the need for thorough phenotyping. Combinatorial approaches can easily produce large and very diverse multigene pathway libraries. Phenotyping such systems *in vivo* can be cumbersome, but cell-free systems can rapidly assess the feasibility of pathway combinations and identify putative bottlenecks (see Ch. 4: Leveraging Cell-Free Systems for Bioproduct Synthesis and Microbial Design, p. 27). Cell-free systems can quickly prototype and validate pathways and genetic circuits before implementing them *in vivo*, enabling faster iterations (Vögeli et al. 2022). In this way, the number of challenging and expensive *in vivo* experiments can be minimized without reducing success rates. Adaptive laboratory evolution (ALE) presents another mechanism for introducing unexpected diversity and identifying beneficial microbial strain mutations under a given environmental condition (e.g., butanol stress).

Making omics assays, both novel and traditional, available at a large scale would release a major bottleneck in microbial design.

High-Throughput Assay Standardization and Data Transferability

Standardized file formats, biological ontologies, and data collection workflows are needed for efficient definition, exchange, and integration of datasets across platforms. Biological ontologies are structured vocabularies defining relationships between biological entities such as genes and metabolites and are particularly important for harmonizing data across labs and ensuring that measurements refer to the same concept. Establishing standardized approaches and benchmarks for key products or techniques, such as assay standardization, is one of the most effective ways to ensure comparable and reproducible results across different laboratories. Reproducibility enhances the transferability and compounding of scientific knowledge.

Working Toward Self-Driving Biofoundries

Self-driving or autonomous biofoundries, platforms that accelerate biological system engineering using high-throughput experimentation and design, represent an aspirational but increasingly realizable vision. These platforms integrate automation, liquid handling, real-time analytics, and ML into closed-loop DBTL cycles. By iteratively refining design hypotheses based on experimental results, biofoundries can optimize strain performance with minimal human intervention.

Key biofoundry components include robotic execution of protocols, automated data capture and analysis, and ML algorithms that recommend the next set of experiments. This approach shortens development timelines and increases throughput, ultimately enabling more sophisticated genome designs that would be impractical using manual methods alone. Self-driving biofoundries are also well suited for exploring undercharacterized organisms and engineering applications that require rapid adaptation to changing performance goals.

6.2 Omics Technologies

Leveraging the complexity of omics data to guide biological engineering is challenging, but ML and real-time measurements present emerging opportunities. AI/ML, for example, can constrain the design space (e.g., in combinatorial pathway design) to combinations that are most likely to succeed (Zhang et al. 2020). ML approaches can predict which designs are more likely to work, reducing the number of costly and time-consuming experiments. Real-time measurement capabilities require development but could enable dynamic adjustment of processes and facilitate data analysis leading to desired outcomes.

Making omics assays, both novel and traditional, available at a large scale (i.e., volume) would release a major bottleneck in microbial design. Several emerging omics technologies could also play key roles in advancing microbial and enzyme biocatalyst design.

- **Single-cell omics** involves measuring omics data for one cell at a time (Kuchina et al. 2021). These data can provide important insights into cellular heterogeneity, cell dynamics, and perhaps even cell stress level.
- **Living-cell omics** refers to nondestructive measurements of cell activity using techniques like *in vivo* nuclear magnetic resonance. The approach would enable real-time measurements, insight into cell metabolism dynamics, and models with more accurate predictions.
- **Post-translational modifications (PTMs)** are critical to understanding both eukaryotic and prokaryotic metabolism. PTMs such as phosphorylation and acetylation can dramatically alter enzyme activity, pathway behavior, and microbial stress responses. While mapping PTMs is increasingly feasible, especially in eukaryotes, systematic PTM characterization in microbes remains underdeveloped.
- **Metabolic flux analysis using labeled isotopes** is the gold standard of metabolic measurement (Antoniewicz et al. 2021) but is limited by its low-throughput nature. Progress toward making this technique high throughput would be invaluable.

6.3 Visualization Tools

Visualization tools help reveal microbial biocatalyst behavior across scales, from molecular interactions and single-cell dynamics to population-level performance in large cultures. Scalable imaging and single-cell technologies can uncover features often missed by bulk measurements, including stress responses, spatial organization, and cell-to-cell variability. When paired with ML, these tools support predictive modeling, early failure detection, and more informed design decisions. To advance such research goals, BER-supported structural biology and imaging resources (berstructuralbioportal.org) provide the scientific community with free access to cutting-edge tools.

Cell Imaging

Cell imaging enables visualizations of cellular morphology and cell–cell interactions that can influence microbial biocatalyst performance. Changes in cell shape, or morphology, can indicate stress and different metabolic states, but they are often subtle and difficult to identify. Increased efforts to develop inexpensive and high-throughput methods to obtain morphological images can provide ML approaches with the input needed to establish such connections (Jin et al. 2022).

Another possible use for cell imaging is detecting microbial contamination, which is a primary risk in large-scale biomanufacturing. Combining the high-throughput nature of cell imaging with ML image recognition approaches can provide rapid early detection of contamination.

Single-Cell Measurements

Single-cell measurements can capture metabolic state at the single-cell level or assess population heterogeneity. Heterogeneity analysis provides insight into whether evolutionary forces may be reducing the fraction of product-synthesizing cells to suboptimal levels (Rugbjerg et al. 2018). Single-cell measurements can also provide information on transporter activity, a critically understudied component of biomanufacturing, and dynamic interactions between cells.

Comprehensive Cell Models

The concept of developing comprehensive cell models that capture molecular behavior is gaining traction, but fully realized 4D spatial–temporal models remain aspirational (Macklin et al. 2020). While tools such as AlphaFold have revolutionized protein structure prediction, integrative models of entire cells reflecting dynamic interactions of molecular machinery over time are still lacking. Building such models will require combining static structural data from sources like AlphaFold, crystallography, and cryogenic electron microscopy and tomography with advanced imaging techniques such as soft X-ray tomography and live-cell imaging. The latter approaches add spatial and, increasingly, temporal resolution. Integrating these data

Rather than treating AI as a post hoc analysis tool, codesigned systems would employ models to dynamically guide experiment selection, interpret outcomes, and adapt to new data in real time.

streams will enable the simulation of cellular processes as they unfold in real time.

6.4 Integrating Artificial Intelligence and Machine Learning into Technology Development

AI and ML are increasingly central to microbial design, offering tools to navigate complex design spaces, prioritize experiments, and extract actionable insights from high-dimensional biological data. Recognizing their transformative potential, BER conducted a workshop in 2022, Artificial Intelligence and Machine Learning for Bioenergy Research (genomicscience.energy.gov/amber-ai-ml), which identified three exemplary research grand challenges that could benefit from AI/ML solutions: designing microbes and microbial communities to specification, enabling closed-loop autonomous design and control for biosystems, and advancing bioprocess scale-up and automation. These approaches can accelerate the DBTL cycle by enabling predictive modeling, guiding experimental design, and integrating diverse data streams. However, key challenges remain in automating laboratory workflows and integrating data using AI/ML approaches, including standardizing data and protocols and capturing contextual data (i.e., metadata).

Biological data is highly contextual. Experimental context (e.g., media, cultivation conditions, and altitude) influences measured phenotype and can complicate the transferability of predictive models across different experimental conditions. Capturing context effectively has proven challenging, often due to hidden variables that commonly influence different measurements. Standardizing metrics and workflows for data collection across different platforms—for example, by adopting data standards inspired by the National Microbiome Data Collaborative standards initiative (microbiomedata.org/overview)—is expected to help alleviate this challenge.

However, file standards and ontologies are sorely needed. Some solutions exist for exchanging certain types of data (Haug et al. 2020; Deutsch et al. 2023), but no generally accepted biology ontologies or file standards exist for synthetic biology and biomanufacturing that combine data types and describe experimental design. Some solutions have been proposed, but general standards are needed, like those the Protein Data Bank contributed to successful AlphaFold development (Burley et al. 2017; Morrell et al. 2017). Although such efforts are critically important, they are very complicated to fund.

Taken together, these insights suggest that the future of microbial design will depend not only on solving data challenges but on forming a new view of AI-integrated experimentation and autonomous discovery. Beyond the fundamental issue of data standardization discussed in Section 6.1: High-Throughput Technologies (see p. 39), AI can accelerate microbial technology development by embedding computational reasoning directly into experimental workflows.

Rather than treating AI as a post hoc analysis tool, codesigned systems would employ models to dynamically guide experiment selection, interpret outcomes, and adapt to new data in real time. To drive innovation in microbial engineering, the field must develop decision-support systems that can operate self-driving biofoundries and automate scientific tasks, such as hypothesis generation, refinement, and consolidation, to build structured knowledge and prioritize promising directions (Martin, H. G., et al. 2023).

Chapter 7

Shared Tools and Resources To Accelerate Microbial Design

A final core workshop theme covered how developing microbes for a robust bioeconomy hinges on the availability of shared tools, resources, and collaborative frameworks that accelerate microbial and enzyme biocatalyst design. Improving this availability requires:

- Building infrastructure that supports advanced technologies, facilities, and capabilities.
- Establishing standardized terminology and data-sharing practices, as well as benchmarking and disseminating innovative experimental and computational methodologies.
- Looking beyond technical advancements to consider potential ethical, legal, and societal implications in an expanding microbial bioeconomy.
- Engaging the public in topics related to microbial science, its applications, and societal implications.
- Training a knowledgeable workforce.

Participants discussed these topics based on diverse biological platforms, from model and nonmodel microbes to synthetic and cell-free systems. While specific needs vary across platforms, several common resources and development needs emerged (see Table 7.1, p. 46).

7.1 Technologies, Facilities, and Capabilities

Progress in microbial design requires an integrated resource ecosystem that seamlessly connects discovery to application. Foundational elements include advanced genomic and metabolic design and engineering tools, state-of-the-art engineering platforms, and collaborative facilities that provide access to specialized experimental infrastructure. When paired with early-stage techno-economic analysis, these resources form a critical backbone for accelerating innovation and delivering next-generation biotechnologies.

Collaborative Facilities

Unlocking the full potential of microbial systems for a growing bioeconomy depends on access to infrastructure and collaborative research facilities. For instance, BER user facilities, including the DOE Joint Genome Institute (JGI; jgi.doe.gov) and Environmental Molecular Sciences Laboratory (EMSL; emsl.pnnl.gov), offer critical state-of-the-art capabilities that accelerate innovation across biological platforms. Establishing biotechnology centers within DOE's existing network of national laboratories could further scale national efforts by coordinating cross-disciplinary resources to address grand research challenges (NSCEB 2025). Centralized user facilities dedicated to genome annotation and model curation could also help avoid duplicating efforts and expand analytical capabilities to researchers nationwide.

Genomic and Metabolic Design and Analytical Tools

Improved tools for genome annotation, metabolic reconstruction, and regulatory network analysis are foundational to microbial design across model and

Table 7.1 Microbial Design Resource and Development Needs Across Biocatalyst Platforms

Technologies, Facilities, and Capabilities (Section 7.1, p. 45)	Standardized Terminology, Data Sharing, and Benchmarking (Section 7.2, p. 47)	Mitigating Risks and Unintended Consequences (Section 7.3, p. 49)	Public Engagement (Section 7.4, p. 50)	Future Workforce Training (Section 7.5, p. 51)
• Collaborative facilities • Genomic and metabolic design and analytical tools • Genetic and molecular engineering capabilities • Accessible relevant experimental infrastructure • Accessible early-stage techno-economic analysis	• Metadata standards and ontologies • Data sharing frameworks • Standardized protocols • Repositories • Benchmarking pipelines and computational standards • Negative results • Incentive structures and best practices	• Biocontainment and biosafety • Cybersecurity and bioassurance • Responsible innovation	• Seeking common language • Clarifying the role of microbial design in the bioeconomy • Raising awareness and accessibility	• Foundational and interdisciplinary training • Specialized technical skills • Recruitment and science communication • Bridging knowledge gaps

Similar resource and development needs span diverse microbial design biocatalyst platforms. With increased focus, five primary categories of needs could accelerate microbial design efforts. Improvements in these spaces can connect discovery to application by enabling predictive genome design; standardizing protocols; increasing access to data, tools, and infrastructure; refining research directions based on changing economic needs; mitigating risk; gaining public support; and training a future workforce.

nonmodel systems. High-quality genome sequences must be paired with curated metabolic reconstructions and growth phenotypes to guide strain development.

Technologies for characterizing strain-level substrate utilization can inform metabolic model constraints and identify intervention points for strain optimization. Essential genome-wide analytical techniques, such as chromatin immunoprecipitation sequencing (ChIP-seq), assay for transposase-accessible chromatin using sequencing (ATAC-seq), and CRISPR Perturb-seq, map gene regulatory elements, particularly in non-model microbes with limited prior characterization. The resulting datasets could enable construction of high-resolution gene regulatory networks. These networks critically inform understanding of physiological responses under varying environmental and process conditions and drive rational design strategies.

In synthetic cell development, a need exists for predictive genome design frameworks that create fully synthetic or minimized genomes with defined functionalities. Metabolic reconstruction efforts benefit from access to rich phenotypic and physiological datasets, which are necessary for calibrating genome-scale metabolic models and more complex kinetic modeling frameworks.

Genetic and Molecular Engineering Capabilities

Critical needs in microbial biosystems design include developing and distributing robust, adaptable genetic tools and expanding access to technologies such as CRISPR interference and activation (CRISPRi/a), gene knockouts, and overexpression libraries across diverse microbial platforms. These capabilities are essential for mapping gene function, validating models, and implementing synthetic regulatory circuits.

However, many researchers, especially those working with nonmodel microbes, face substantial technical and financial barriers in domesticating microbial strains and adapting tools. Establishing collaborative user-accessible strain engineering centers could provide standardized workflows, DNA synthesis services, and expertise for developing and distributing toolkits across a wide range of microbial platforms. A fee-based access model could enable academic research groups to pursue ambitious projects while enabling industry stakeholders to benefit from and contribute to nationwide advances in microbial genetic engineering.

Accessible Relevant Experimental Infrastructure

Translating genetic insights into functional, industrially relevant outcomes requires shared experimental infrastructure. Bioreactor platforms are needed to support scale-down and scale-up studies, enabling systematic evaluation of strain performance under process-relevant conditions. Access to these systems would facilitate knowledge transfer, reduce duplicated efforts, and provide critical feedback loops between strain engineering and process optimization.

Infrastructure requirements for cell-free and synthetic systems are distinct from microbial systems. They include microbial lysate production facilities at scale, lysate quality standardization protocols, and analytical pipelines for real-time reaction monitoring. Improving the generation and testing of lysates from broad microbial chassis (i.e., other than *E. coli*) can directly expand the utility of cell-free systems producing complex and toxic biomolecules. Lysate-based systems also need to be interfaced with downstream chemical processing environments, including organic solvent and nonaqueous conditions.

Early integration of techno-economic analysis can prioritize research directions and improve alignment with real-world constraints.

Accessible Early-Stage Techno-Economic Analysis

The primary focus of microbial genome engineering remains on enabling foundational science, but early integration of techno-economic analysis (TEA) can prioritize research directions and improve alignment with real-world constraints. TEA provides a complementary perspective to biological modeling by enabling early estimates of mass and energy balances, product yields, and process economics.

TEA can guide decisions about strain selection, pathway prioritization, and experimental resource allocation, particularly under constraints of throughput, time, and funding. It also supports strategic trade-offs, such as whether to invest in adapting nonmodel organisms or optimizing legacy chassis. Efforts to more tightly integrate TEA into design-build-test-learn (DBTL) cycles could inform iterative engineering decisions and bridge gaps between basic discovery and commercial deployment.

7.2 Standardized Terminology, Data Sharing, and Benchmarking

Accelerating microbial design across diverse platforms relies heavily on establishing and adopting standardized terminologies, data-sharing frameworks, and benchmarking protocols. Across model microbes, nonmodel microbes, cell-free systems, and synthetic cells, several key themes emerged that underpin a shared infrastructure.

Metadata Standards and Ontologies

Standardized metadata and ontologies are critical for capturing genetic constructs, experimental conditions, phenotypic measurements, and environmental parameters in a format that supports data integration and reuse. Unified identifiers for genes, proteins, strains, and synthetic parts, as well as harmonized naming conventions and curated reaction databases, are needed. Progress is underway, but a broader commitment to metadata standardization is necessary to enable discovery driven by AI and machine learning (ML) and predictive modeling in microbial biosystems design.

Adopting standardized protocols could provide a foundation for interoperable data ecosystems and community-driven standards that accelerate progress in biosystems design.

Data-Sharing Frameworks

Effective data-sharing infrastructure is essential to accelerating progress in microbial genome engineering. However, substantial challenges persist due to concerns about data privacy, intellectual property, inconsistent formatting, and no shared repositories. These challenges limit reuse of high-quality datasets and hinder collaborative efforts across institutions. Technical solutions that support data exchange without requiring direct access to raw data are needed to address this, particularly in contexts involving proprietary or sensitive information.

A promising approach to building data-sharing infrastructure is federated learning, an ML technique in which models are trained across decentralized datasets without moving the data. Instead, participating institutions train local models and only share updates (e.g., model weights or gradients), which are aggregated into a global model. This preserves privacy while developing robust, generalizable models across diverse biological systems. Federated learning is a viable strategy for linking distributed research efforts, such as microbial strain optimization or omics analysis, while maintaining institutional autonomy.

Participants also discussed establishing a centralized "Web of Biological Data" digital platform to unify access to annotated genomes, experimental data, and metadata. Such infrastructure, if appropriately funded and governed, could transform microbial design knowledge sharing and application.

Standardized Protocols

Adopting standardized protocols across microbial biosystems design is essential to ensure consistent experimental procedures, data acquisition, and computational analysis. Standardization in these areas, if widely adopted, could provide a foundation for interoperable data ecosystems and community-driven standards that accelerate progress in biosystems design.

A major challenge to experimental and computational reproducibility and comparability is a lack of shared methods, particularly when engineering diverse organisms across model, nonmodel, and synthetic platforms. This challenge extends to both wet-laboratory workflows (e.g., DNA assembly, transformation, lysate preparation, and expression assays) and dry-laboratory procedures (e.g., data preprocessing, model training, and simulation benchmarking).

Greater consistency in experimental design would improve performance comparisons and reduce method revalidation across laboratories. Centralized, version-controlled repositories for protocols, datasets, and materials are critical infrastructure to support transparent and iterative microbial design.

Another obstacle to protocol standardization is heterogeneous omics data collection, which often varies by instrumentation, software pipeline, and data reporting convention. Aligning data-capture practices can facilitate downstream integration and analysis, particularly in ML applications.

Comprehensive quality assurance and control practices are needed, especially for lysate-based and cell-free systems. The 2024 Engineering Biology Research Consortium (EBRC) report, *Engineering Biology*

Metrics and Technical Standards for the Global Bioeconomy, provides a guiding reference, recommending actions to improve assay reproducibility, harmonize terminology, and facilitate alignment across experimental platforms (Freemont et al. 2024).

Repositories

Accessible repositories for strains, lysates, genetic parts, and standardized protocols are essential infrastructure for collaborative microbial biosystems design. Streamlined material transfer agreements and increased protocol transparency would lower access barriers, particularly for academic users. Curated collections of nonmodel organisms and reference strains are needed to support rational microbial chassis selection and reduce redundant microbial domestication efforts.

Benchmarking Pipelines and Computational Standards

Community-wide benchmarking frameworks are needed to enable consistent evaluation of biological components, engineered strains, and computational tools. Standardized pipelines for data processing and model evaluation, using shared performance metrics and assumptions, will support reproducible, iterative design cycles. These frameworks are especially important as biosystems design expands to increasingly complex, multigene modifications and integrated metabolic networks.

Negative Results

Limited sharing of negative results is a major impediment to progress in biosystems design. Although academic and commercial entities face different incentives, intentional negative findings are valuable for improving model accuracy and avoiding redundant experiments. Lightweight reporting formats, such as structured summaries or checklists, could help address this gap. Benchmarking studies and reproducibility initiatives may provide a platform to share these results without disrupting publication norms.

Incentive Structures and Best Practices

Establishing clear incentive structures is essential for promoting rigor, reproducibility, and transparency. Recognizing activities such as protocol sharing, negative results publishing, and contributing to standardized datasets can help shift academic and industrial norms toward more open and collaborative practices. Journals and funding agencies are increasingly positioned to enforce expectations around data availability, metadata completeness, and methodological transparency. However, such expectations must be accompanied by aligned incentives and practical support to ensure that compliance is feasible and rewarded.

Incentive structures must align with broader national priorities in biotechnology, biosecurity, and data stewardship. As the U.S. government continues to support developing a secure and resilient bioeconomy, coordination among federal agencies, research institutions, and industry stakeholders becomes increasingly important. Best practices should reflect not only the research community's technical needs but also national strategic investment goals in microbial biosystems design and engineering.

7.3 Mitigating Risks and Unintended Consequences

As microbial technologies advance toward broader industrial, environmental, and biomedical applications, proactive identification and management of risks and unintended consequences is essential. A comprehensive approach to risk mitigation spans biosafety, cybersecurity, ethical and societal engagement, and continual improvement of regulatory frameworks. These considerations are particularly important in the context of emerging platforms such as synthetic cells, nonmodel microbes, and cell-free systems.

Biocontainment and Biosafety

Clearly defined biosafety protocols and stewardship frameworks are needed across all microbial platforms. Robust and platform-specific biocontainment strategies critically prevent unintended release and persistence of engineered microbes beyond application

requirements. For both model and nonmodel organisms, improved genetic safeguards and physical containment protocols should mitigate potential risks to the environment. Biocontainment system design must consider the potential of horizontal gene transfer, especially in open environment applications.

Synthetic cells benefit from an inherently nonreplicative nature, significantly reducing uncontrolled propagation risks. Nonetheless, these cells require appropriate containment in contexts involving environmental release. Similarly, cell-free systems, which lack living organisms, are inherently low-risk. Yet biosafety concerns remain for lysate contamination, improper lysate disposal, and unintended use of cell-free kits for unsupervised gene expression.

Cybersecurity and Bioassurance

As microbial biosystems design and engineering become increasingly automated and data-driven, digital infrastructure security and integrity are critically important. Improved collaboration between microbiologists, computer scientists, and cybersecurity experts can ensure that cybersecurity and bioassurance protocols keep pace with emerging technologies in microbial systems design.

Cybersecurity and bioassurance measures are essential to protect not only sensitive genomic and experimental data, but also the hardware and software systems that control biological workflows. These systems may include cloud-based biofoundry platforms, robotic liquid handlers, and DNA synthesis pipelines.

Bioassurance encompasses a broad framework that includes cybersecurity, physical security, and trustworthiness of engineered biological systems. Secure authentication systems, version control protocols, and audit trails are needed to ensure data provenance and design reproducibility. Critical proactive safeguard strategies include implementing access controls, anomalous activity monitoring, and integrating DNA sequence screening tools into design tools.

Responsible Innovation

As microbial biosystems design accelerates toward real-world deployment, research trajectories should align with national priorities for economic competitiveness, supply chain resilience, and robust bioeconomic growth. Rather than treating ethical, legal, and societal factors as afterthoughts, responsible innovation calls for their integration throughout the research and development life cycle.

7.4 Public Engagement

Promoting public trust and understanding is essential to realizing the full potential of microbial design in building a robust bioeconomy. As engineered microbes increasingly contribute to sectors like food, health, and energy, clear communication and inclusive outreach are critical to building societal acceptance.

Seeking Common Language

An important strategy is building public understanding and confidence in microbial technologies that underpin an emerging bioeconomy. Naming conventions and language choices can significantly influence public perception. For example, microorganisms associated with food or fermentation (e.g., *Saccharomyces cerevisiae*) are often viewed more positively than less familiar species. In addition, terms like organic, GMO, and synthetic often carry different meanings for different people, and complex scientific language can contribute to misunderstanding or mistrust.

Transparent communication and clear, accessible explanations of benefits, risks, and safeguards are essential to fostering public trust. Anticipating both positive and adverse societal and economic impacts of microbial systems will ensure that new technologies serve national interests across energy, agriculture, health, and manufacturing sectors.

Clarifying the Role of Microbial Design in the Bioeconomy

Transparent messaging about the opportunities and limitations of engineered microbes helps bridge the gap between innovation and public perception. Collaborations with familiar industries (e.g., breweries and bakeries) can ground abstract concepts in everyday experiences. Educational initiatives that highlight real-world benefits, from sustainable manufacturing to healthcare, reinforce the value of microbial

technologies. However, many researchers feel unprepared to engage with the public, pointing to the need for institutional and funding support for science communication training.

Raising Awareness and Accessibility

Targeted public outreach and early education help demystify microbial science and engage future generations. Activities such as hands-on microbiology experiments and school-based programs make science more tangible. Framing biotechnology within a historical context, such as its relationship to traditional crop and livestock breeding, can also help audiences understand its trajectory. Still, trust can be eroded if messaging appears biased, and apparent public support may conceal widespread misunderstanding of the science, leading to misplaced confidence in technology adoption. Continued, accessible, and transparent engagement is essential to support meaningful dialogue.

7.5 Future Workforce Training

Workforce training must evolve alongside emerging technologies to meet growing demand for expertise across biological, computational, and engineering domains relevant to microbial design. Strengthening the domestic biotechnology workforce and sustaining a pipeline of talent nationally and internationally are critical to advancing and securing the biotechnology sector (NSCEB 2025). Key workforce training targets can support this goal.

Foundational and Interdisciplinary Training

As engineering biology and synthetic biology tools become more advanced and expensive, stakeholders must explore how to modernize university curricula and improve access to shared infrastructure to train the next generation of scientists and engineers. Addressing current and future challenges in microbial and enzyme biocatalyst design depends on:

- Developing interdisciplinary student curricula that integrate biology, mathematics, and computing.
- Expanding foundational education to include high-throughput experimentation, automation, and ML.
- Recommending fermentation science training to bridge the divide between laboratory and industrial cultivation scales.
- Offering industrial experience that teaches the feasibility of translating technologies from experimental to applied platforms. Such training is well-established in engineering programs but less common in biology fields.
- Bridging the persistent divide between scientists and engineers by providing interdisciplinary training. This approach can sometimes produce a broadly skilled workforce that lacks deep specialization, but such individuals play a valuable role in facilitating cross-discipline communication.

Specialized Technical Skills

Targeted and platform-specific training enables researchers to operate in nontraditional biological environments and accelerate iterative design cycles. For example, cell-free systems benefit from hands-on instruction in miniaturized experimental platforms, DNA design and preparation, and laboratory automation.

Workforce training must evolve alongside emerging technologies to meet growing demand for expertise across biological, computational, and engineering domains relevant to microbial design.

Recruitment and Science Communication

Building a strong talent pipeline requires intentional recruitment and greater emphasis on science communication, especially public engagement. Despite its importance, many scientists report feeling unqualified or unsupported in conducting outreach activities. Greater institutional and federal investment in science

communication training would strengthen efforts to build trust around novel technologies.

Bridging Knowledge Gaps

Cross-disciplinary training should extend beyond students to include principal investigators and senior researchers. Important gaps persist: many biologists lack AI training, and many computational scientists lack the domain expertise needed to pose meaningful biological questions.

Fellowships like the DOE-funded Computational Science Graduate Fellowship are key to successfully training individuals to navigate both biological and computational domains. Bridging this divide will also require shared platforms, standardized datasets, and clearly defined biological objectives that guide development.

Appendix A

Workshop Agenda

All times U.S. Eastern

Day 1: Tuesday, January 28, 2025

12:00 p.m. **Welcome**

Speakers: Dorothy Koch, Associate Director, U.S. Department of Energy Biological and Environmental Research (BER) program; and Todd Anderson, Director, BER Biological Systems Science Division

12:10 p.m. **Opening Remarks and BER Workshops Overview**

Speaker: Dawn Adin, Program Manager, Genomic Science Program

12:20 p.m. **Introduction to the Workshop and Logistics**

Speakers: James Carothers (Workshop Co-Chair), University of Washington; and Davinia Salvachúa Rodríguez (Workshop Co-Chair), National Laboratory of the Rockies

12:30 p.m. **Overview of Responses to Pre-Workshop Survey Questions**

Key Challenge Question: Focus areas and gaps in microbial design for realizing the bioeconomy

- State-of-the-art strategies for molecule(s) target selection and production
- Fundamental gaps in microbial design when scaling up cultivations
- Successes and barriers when integrating microbial design and artificial intelligence (AI)
- Biotechnology successes and their underlying reasons
- Biocatalysts, bioproducts, and feedstocks

12:45 p.m. **Plenary 1: Connecting Microbial Design to Product Molecules**

Speaker: Brent Shanks, Iowa State University

1:15 p.m. **Break**

1:35 p.m. **Introduction to Breakout Sessions: Goals, Expectations, Logistics, and Ethics**

1:40 p.m. **Breakout 1: Research Opportunities To Fill Key Knowledge Gaps on Bioproduct Synthesis**

The goal of this session is to pinpoint critical gaps, strategies, and opportunities for designing efficient pathways that produce desirable products. The objectives include identifying strategies to manage the complexity of gene regulatory and metabolic networks, mitigate control problems created by new metabolic pathways, and enhance the efficiency of microbial metabolism for the synthesis of desirable products.

2:50 p.m. **Break**

3:15 p.m. **Plenary 2: Microbial Alchemy—Engineering Life To Fuel the Future**

Speaker: Jay Keasling, University of California–Berkeley

3:45 p.m. **Breakout 2: Research Opportunities To Fill Key Knowledge Gaps on Microbial Genome Engineering**

The goal of this session is to identify routes to advance genome engineering and microbial design. This could include integrating cutting-edge molecular engineering tools with AI-driven predictive models to enhance strain performance and streamline complex pathway engineering. We need to identify approaches to address current limitations and leverage sequencing data to drive innovation and efficiency in developing novel, industrially relevant microorganisms.

4:55 p.m. **Closing Remarks**

Day 2: Wednesday, January 29, 2025

12:00 p.m. **Welcome and Highlights from Day 1**

This session will present a synthesis of the high-level responses to the key questions proposed on Day 1. Through an interactive voting process, we will collectively identify and highlight "5 Key Insights and Learnings" that emerged during the workshop. The purpose of this activity is to reflect on the diversity of ideas shared rather than to establish rankings or priorities. We invite all participants to engage in this collaborative exercise.

12:30 p.m. **Plenary 3: Discovery of Novel Pathways and Molecules Using AI and Machine Learning (ML)**

Speaker: Linda Broadbelt, Northwestern University

1:00 p.m. **Break**

1:20 p.m. **Breakout 3: Research Strategies for Accelerating the Design-Build-Test-Learn Cycle in Microbial Design**

This session will focus on describing the gaps that remain in the ability to efficiently perform high-quality and high-throughput assembly and phenotyping of biocatalysts integrated with AI/ML. Additionally, it will outline future proposed technologies for accelerating and improving the understanding and design of biocatalysts.

2:30 p.m. **Break**

3:00 p.m. **Rapid-Fire Presentations**

Five minutes each on: (1) cell-free bioproduction, (2) cell-free gene expressions and synthetic cells, and (3) orthogonal processes to describe the state-of-the-art and highest-priority knowledge gaps and opportunities for realizing the full potential of these biological strategies for product generation.

3:15 p.m. **Breakout 4: Research Opportunities To Leverage Alternative Biocatalysts for Product Generation and Understanding Microbial Design**

Compared to microbial systems, nonmicrobial biocatalysts—such as cell-free systems, synthetic cells, and orthogonal processes—offer an alternative to expand the reaction space, streamline reaction control, and bypass cellular regulation, among other features. This session will focus on proposing strategies for leveraging alternative biocatalysts to generate products and gain fundamental understanding of microbial design.

4:25 p.m. **Closing Remarks**

Day 3: Thursday, January 30, 2025

12:00 p.m. **Welcome and Highlights from Day 2**

This session will present a synthesis of the high-level responses to the key questions proposed on Day 2. Through an interactive voting process, we will collectively identify and highlight "5 Key Insights and Learnings" that emerged during the workshop. The purpose of this activity is to reflect on the diversity of ideas shared rather than to establish rankings or priorities. We invite all participants to engage in this collaborative exercise.

12:30 p.m. **Plenary 4: Designing Biology for Remediation and Sustainability**

Speaker: Pamela Silver, Wyss Institute at Harvard University

1:00 p.m. **Plenary 5: Transferring Lab Innovation into Large-Scale Application by Thorough Understanding of Strains and Processes**

Speaker: Ralf Takors, University of Stuttgart, Germany

1:30 p.m. **Break**

1:50 p.m. **Breakout 5: Research Opportunities To Fill Key Knowledge Gaps To Improve and Predict Scalability of Biocatalysts**

There is often a disconnect between research teams developing proof-of-concept systems in the laboratory and those driving the innovations required for the scale-up of microbial cultivations. Bioprocess development and engineering efforts involve optimizing bioreactor design; refining cultivation techniques; and developing scalable, real-time measurement technologies to improve production (titers, rates, yields). Additionally, successful bioprocesses may also include the integration of microbial conversion with upstream substrate preprocessing and downstream separations. This session will focus on describing the fundamental knowledge and innovations required to design biocatalysts that can predict and ensure successful performance at large cultivation scales, either independently or in conjunction with upstream or downstream processes.

3:00 p.m. **Break**

3:20 p.m.	**Breakout 6: Shared Tools and Resources To Accelerate Biocatalyst Design for a Developing Bioeconomy** This session will focus on describing the resources necessary to enable or accelerate biocatalyst design as well as the implications of expanding a microbial-based bioeconomy. In each breakout room we will create a microbial bioeconomy around a model microbe, a nonmodel microbe, a cell-free system, and a synthetic cell.
4:30 p.m.	**Breakouts 5 and 6 Report-Out**
4:50 p.m.	**Next Steps and Closing Remarks**

Appendix B

Workshop Participants

Chairs

James Carothers, *University of Washington*

Davinia Salvachúa Rodríguez, *National Laboratory of the Rockies*

Participants

Caroline Ajo-Franklin, *Rice University*

Ludmilla Aristilde, *Northwestern University*

Adam Arkin, *University of California–Berkeley*

José Avalos, *Princeton University*

Ahmed Badran, *Scripps Research Institute*

Jay Bardhan, *Pacific Northwest National Laboratory*

Alexander Beliaev, *Pacific Northwest National Laboratory*

Andrew Borchert, *Crysalis Biosciences*

Linda Broadbelt, *Northwestern University*

Virginia Cornish, *Columbia University*

Taraka Dale, *Los Alamos National Laboratory*

Adam Feist, *University of California–San Diego*

Hector García Martín, *Lawrence Berkeley National Laboratory*

Adam Guss, *Oak Ridge National Laboratory*

Farren Isaacs, *Yale University*

Laura Jarboe, *Iowa State University*

Neel Joshi, *Northeastern University*

Jay Keasling, *University of California–Berkeley*

Anna Kuchina, *Institute for Systems Biology*

Han Li, *University of California–Irvine*

Costas Maranas, *The Pennsylvania State University*

John McGeehan, *National Laboratory of the Rockies*

Aindrila Mukhopadhyay, *Lawrence Berkeley National Laboratory*

Vincent Noireaux, *University of Minnesota–Twin Cities*

Pamela Peralta-Yahya, *Georgia Institute of Technology*

Brian Pfleger, *University of Wisconsin–Madison*

Ritimukta Sarangi, *Stanford Synchrotron Radiation Lightsource*

Brent Shanks, *Iowa State University*

Pamela Silver, *Harvard University*

Nádia Skorupa Parachin, *Cemvita*

Ralf Takors, *University of Stuttgart*

Meltem Urgun-Demirtas, *Argonne National Laboratory*

Yasuo Yoshikuni, *Lawrence Berkeley National Laboratory*

Karsten Zengler, *University of California–San Diego*

Huimin Zhao, *University of Illinois Urbana-Champaign*

Jeremy Zucker, *Pacific Northwest National Laboratory*

Appendix C

Breakout Questions

Breakout 1

Research Opportunities To Fill Key Knowledge Gaps in Bioproduct Synthesis

1. Despite several decades of successes in metabolic engineering that have increased production of amino acids, biofuels, and other valuable chemicals, why are metabolic pathways still hard to engineer?
2. In what ways can pathway complexity be reduced while improving yields? Examples?
3. What major challenges arise when designing new-to-nature pathways to synthesize molecules? What are the most promising ways to overcome these? What are the successes so far?
4. Bioproduct types: How are the challenges different when the microbe is the product (i.e., living materials) versus when the microbe produces the product (e.g., macromolecules versus small molecules versus inorganics)? What are the challenges between small and macromolecular products? What are the challenges with molecules with different physicochemical properties?
5. When and where should we use nonbiological, transformation approaches (e.g., photocatalysis, electrocatalysis, thermocatalysis, polymerization), and how do we integrate those directly with biological transformations? How do we optimally deliver abiotic inputs to biological systems?
6. Many chemical transformations we aim to replace are conducted in nonaqueous media today. How do you break free from biological processes that depend on water (e.g., integrate biology with water-stable catalysts or integrate biocatalysts with organic solvents)?
7. How do we advance control and predictability of metabolic pathway engineering? For instance, how do we incorporate experimental, mathematical, and computational (e.g., kinetic, AI, and machine learning) analyses to design new pathways and bring metabolic network engineering to the next level?
8. How can we effectively use published information along with chemical and enzymatic catalysis databases and AI to facilitate and accelerate target molecule selection?

Breakout 2

Research Opportunities To Fill Key Knowledge Gaps in Microbial Genome Engineering

1. In the last decade, which tools have successfully meaningfully improved strain performance, and how can we leverage that knowledge for the next generation of tool development?
2. What are the rate-limiting steps in genome engineering? What is the largest number of genome modifications to date (for chemical bioproduction purposes)? What can't we do today?
3. Which tools should be prioritized to accelerate and simplify microbial and viral design?
4. How do we more rapidly and reliably discover new nonintuitive genetic targets to improve microbial genome engineering efficacy?
5. How does the interplay between application needs and tool development shape innovation in this field? Does application drive tool development, or is it the other way around?
6. How could information gained from model microbes be translated to other, more

bioenergy- and/or industrially relevant (novel) organisms?

7. What are knowledge gaps and opportunities in re-coded, minimal, and synthetic genomes (new microbes) to develop state-of-the-art biomolecular and pathway engineering in novel platform organisms?
8. How can we leverage the vast amounts of DNA sequences generated each year to improve genome engineering?

Breakout 3

Research Strategies for Accelerating the Design-Build-Test-Learn (DBTL) Cycle in Microbial Design

1. High-Throughput Technologies:
   a. What are opportunities to enhance the effectiveness of the DBTL cycle in a disruptive fashion?
   b. How does DBTL success depend on the number of samples that are tested? What does that number look like?
   c. Which considerations are needed in phenotyping technologies to boost speed, applicability to different organisms, translatability to different scales, and integration with different technologies?
   d. Which breakthroughs are needed to maximize genetic diversity (genes/libraries/strains/enzymes) and minimize phenotyping as much as possible to accelerate the DBTL cycle?
   e. Do we need standardized assays for the key products we would like to make?
   f. How can we accelerate how fast we generate high-throughput assays for different products?
2. Omic Technologies:
   a. How can we combine and leverage the complexity of omic data to accelerate microbial and viral genome engineering or enzyme design?
   b. What is the next emerging omic technology that will be critical in advancing biocatalyst design?
3. Visualization Tools:
   a. What specific cellular features or processes could we focus on to better understand biocatalyst performance and optimization from nano- to macroscale?
4. Integration of AI and Machine Learning (AI/ML):
   a. Assuming high-throughput and data quality requirements are significantly exceeded, what are the key challenges we should anticipate in automating laboratory workflows and integrating data with AI/ML across various platforms?
   b. What novel ways can AI accelerate DBTL optimization and enhance or reduce the need for high throughput/automation?

Breakout 4

Research Opportunities To Leverage Alternative Biocatalysts for Product Generation and Understanding Microbial Design

1. With current or novel tools, how can we accelerate the design of enzymes with desired characteristics (e.g., stability, durability, optimal pH, optimal temperature, kinetics, allostery)?
2. How can we leverage cell-free systems to learn about post-translational pathway control (enzyme allosteric control, phosphorylation cascades) to expand the toolbox for pathway engineering?
3. How can cell-free systems accelerate our understanding of in-pathway metabolite leakage into other pathways/fates?
4. How can nonmicrobial biocatalysts help us gain a rapid understanding of cellular toxicity?
5. How can cell-free systems contribute to accelerating the design, prototyping, and discovery of metabolic pathways or genetic circuits?
6. How can cell-free systems be optimized to more accurately model or predict *in vivo* performance?
7. What breakthroughs are required to effectively integrate orthogonal processes into biological engineering and biosystems design?

8. How can we push the boundaries of current synthetic cell design to create viable platforms for bioproduct generation?

Breakout 5

Research Opportunities To Fill Key Knowledge Gaps To Improve and Predict Scalability of Biocatalysts

1. What does successful microbial performance at scale (>1,000 L) mean from a biocatalyst point of view?
2. Achieving consistent performance at different scales and over time remains a significant challenge, likely due to the varying stresses that biocatalysts encounter (e.g., high product concentrations intracellularly versus extracellularly, microenvironments in bioreactors, high cell densities, evolution). What does the ideal biocatalyst look like for enhancing, controlling, and uncovering novel approaches to maintaining consistency?
3. The integration of biological tools (e.g., omics, biosensors) and computational modeling could be useful for predicting scalability and optimizing microbial processes. What tools and data could we use now to predict microbial performance in large-scale cultivations? What new tools and data do we need to develop to improve those predictions?
4. Integrating the upstream deconstruction of polymeric substrates and direct conversion by microbes into high-yield products remains a challenge. What radical fundamental innovations in biosystem design and integration are required to overcome these obstacles and ensure performance at scale?
5. How could we design biocatalysts to enable facile downstream product separations?
6. How, and at what stages, should economic feasibility be incorporated to effectively support and guide the science of scale-up?
7. What other key scientific questions need to be answered to enable seamless large-scale microbial cultivations?

Breakout 6

Shared Tools and Resources To Accelerate Biocatalyst Design for a Developing Bioeconomy

1. Resources, technologies, facilities, and capabilities.
2. Standardized terminology and data-sharing methods.
3. Benchmarking and sharing new methods (experimental and computational) to benefit the entire community.
4. Mitigating the risks or unintended consequences of expanding microbial bioeconomy applications (i.e., ethical, legal, and societal implications).
5. Education and public engagement to build greater familiarity with microbial systems.
6. Training the workforce of the future.
7. What else are we missing?

Appendix D

References

Abramson, J., et al. 2024. "Addendum: Accurate Structure Prediction of Biomolecular interactions with AlphaFold 3," *Nature* **636**(8042), E4. DOI:10.1038/s41586-024-08416-7.

An, B., et al. 2023. "Engineered Living Materials for Sustainability," *Chemical Reviews* **123**(5), 2349–419. DOI:10.1021/acs.chemrev.2c00512.

Antoniewicz, M. R. 2021. "A Guide to Metabolic Flux Analysis in Metabolic Engineering: Methods, Tools and Applications," *Metabolic Engineering* **63**, 2–12. DOI:10.1016/j.ymben.2020.11.002.

Arkin, A. P., et al. 2018. "KBase: The United States Department of Energy Systems Biology Knowledgebase," *Nature Biotechnology* **36**(7), 566–69. DOI:10.1101/096354.

Bhattarai-Kline, S., et al. 2022. "Recording Gene Expression Order in DNA by CRISPR Addition of Retron Barcodes," *Nature* **608**, 217–25. DOI:10.1038/s41586-022-04994-6.

Biggs, B. W., et al. 2021. "Enabling Commercial Success of Industrial Biotechnology," *Science* **374**(6575), 1563–65. DOI:10.1126/science.abj5040.

Biner, O., et al. 2020. "Bottom-Up Construction of a Minimal System for Cellular Respiration and Energy Regeneration," *ACS Synthetic Biology* **9**(6), 1450–59. DOI:10.1021/acssynbio.0c00110.

Boo, A., et al. 2019. "Host-Aware Synthetic Biology," *Current Opinion in Systems Biology* **14**, 66–72. DOI:10.1016/j.coisb.2019.03.001.

Burley, S. K., et al. 2017. "Protein Data Bank (PDB): The Single Global Macromolecular Structure Archive," In: *Protein Crystallography. Methods in Molecular Biology*, vol 1607. Eds. Wlodawer, A., Z. Dauter, and M. Jaskolski. Humana Press, New York, NY. DOI:10.1007/978-1-4939-7000-1_26.

Cardiff, R. A. L., et al. 2025. "Carbon-Conserving Bioproduction of Malate in an *E. coli*-Based Cell-Free System," *Metabolic Engineering* **91**, 59–76. DOI:10.1016/j.ymben.2025.03.020.

Chowdhury, S., et al. 2024. "Carbon Negative Synthesis of Amino Acids Using a Cell-Free-Based Biocatalyst," *ACS Synthetic Biology* **13**(12), 3961–75. DOI:10.1021/acssynbio.4c00359.

Cong, L., et al. 2013. "Multiplex Genome Engineering Using CRISPR/Cas Systems," *Science* **339**(6121), 819–23. DOI:10.1126/science.1231143.

Cordell, W. T., et al. 2023. "Milligrams to Kilograms: Making Microbes Work at Scale," *Trends in Biotechnology* **41**(11), 1442–57. DOI:10.1016/j.tibtech.2023.05.002.

Crater J. S., and J. C. Lievense. 2018. "Scale-Up of Industrial Microbial Processes," *FEMS Microbiology Letters* **365**(13), fny138. DOI:10.1093/femsle/fny138.

Dahl, R., et al. 2013. "Engineering Dynamic Pathway Regulation Using Stress-Response Promoters," *Nature Biotechnology* **31**, 1039–46. DOI:10.1038/nbt.2689.

Darch, S. E., et al. 2012. "Density-Dependent Fitness Benefits in Quorum-Sensing Bacterial Populations," *Proceedings of the National Academy of Sciences of the USA* **109**(21), 8259–63. DOI:10.1073/pnas.1118131109.

Delvigne, F., et al. 2014. "Metabolic Variability in Bioprocessing: Implications of Microbial Phenotypic Heterogeneity," *Trends in Biotechnology* **32**(12), 608–16. DOI:10.1016/j.tibtech.2014.10.002.

Delvigne, F., et al. 2017. "Bioprocess Scale-Up/Down as Integrative Enabling Technology: From Fluid Mechanics to Systems Biology and Beyond," *Microbial Biotechnology* **10**(5), 1267–74. DOI:10.1111/1751-7915.12803.

Deutsch, E. W., et al. 2023. "The ProteomeXchange Consortium at 10 Years: 2023 Update," *Nucleic Acids Research* **51**(D1), D1539–48. DOI:10.1093/nar/gkac1040.

Dong, C., et al. 2018. "Synthetic CRISPR-Cas Gene Activators for Transcriptional Reprogramming in Bacteria," *Nature Communications* **9**, 2489. DOI:10.1038/s41467-018-04901-6.

Freemont, P. S., et al. 2024. *Engineering Biology Metrics and Technical Standards for the Global Bioeconomy*. Engineering Biology Research Consortium, London, UK. DOI:10.25561/110822.

Gach, P. C., et al. 2017. "Droplet Microfluidics for Synthetic Biology," *Lab on a Chip* **17**, 3388–400. DOI:10.1039/C7LC00576H.

Gebru, K. A., et al. 2021. "Bio-Cement Production Using Microbially Induced Calcite Precipitation (MICP) Method: A Review," *Chemical Engineering Science* **238**(116610). DOI:10.1016/j.ces.2021.116610.

Gibson, D. G., et al. 2010. "Creation of a Bacterial Cell Controlled by a Chemically Synthesized Genome," *Science* **329**(5987), 52–6. DOI:10.1126/science.1190719.

Gu, C., et al. 2019. "Current Status and Applications of Genome-Scale Metabolic Models," *Genome Biology* **20**(121). DOI:10.1186/s13059-019-1730-3.

Hannon, J. R., et al. 2020. "Technoeconomic and Life-Cycle Analysis of Single-Step Catalytic Conversion of Wet Ethanol into Fungible Fuel Blendstocks," *Proceedings of the National Academy of Sciences of the USA* **117**(23), 12576–83. DOI:10.1073/pnas.1821684116.

Haug, K., et al. 2020. "MetaboLights: A Resource Evolving in Response to the Needs of Its Scientific Community," *Nucleic Acids Research* **48**(D1), D440–44. DOI:10.1093/nar/gkz1019.

Heckmann, D., et al. 2020. "Kinetic Profiling of Metabolic Specialists Demonstrates Stability and Consistency of *In Vivo* Enzyme Turnover Numbers," *Proceedings of the National Academy of Sciences of the USA* **117**(37), 23182–90. DOI:10.1073/pnas.2001562117.

Heili, J. M., et al. 2024. "Controlled Exchange of Protein and Nucleic Acid Signals from and Between Synthetic Minimal Cells," *Cell Systems* **15**(1), 49–62.e4. DOI:10.1016/j.cels.2023.12.008.

Hodgson, A., et al. 2022. *The U.S. Bioeconomy: Charting a Course for a Resilient and Competitive Future*, Schmidt Futures, New York, NY. https://www.schmidtfutures.org/wp-content/uploads/2022/04/Bioeconomy-Task-Force-Strategy-4.14.22.pdf

Hren, A., et al. 2025. "High-Density CRISPRi Screens Reveal Diverse Routes to Improved Acclimation in Cyanobacteria," *Proceedings of the National Academy of Sciences of the USA* **122**(12), e2412625122. DOI:10.1073/pnas.2412625122.

Hu, M., et al. 2024. "KETCHUP: Parameterizing of Large-Scale Kinetic Models Using Multiple Datasets with Different Reference States," *Metabolic Engineering* **82**, 123–33. DOI:10.1016/j.ymben.2024.02.002.

Isaacs, F. J., et al. 2011. "Precise Manipulation of Chromosomes *In Vivo* Enables Genome-Wide Codon Replacement," *Science* **333**(6040), 348–53. DOI:10.1126/science.1205822.

Jin, J., et al. 2022. "Robotic Data Acquisition with Deep Learning Enables Cell Image–Based Prediction of Transcriptomic Phenotypes," *Proceedings of the National Academy of Sciences of the USA* **120**(1), e2210283120. DOI:10.1073/pnas.2210283120.

Jinek, M., et al. 2012. "A Programmable Dual-RNA–Guided DNA Endonuclease in Adaptive Bacterial Immunity," *Science* **337**(6096), 816–21. DOI:10.1126/science.1225829.

Johnson, C. W., et al. 2019. "Innovative Chemicals and Materials from Bacterial Aromatic Catabolic Pathways," *Joule* **3**(6), 1523–37. DOI:10.1016/j.joule.2019.05.011.

Johnston, T. G., et al. 2020. "Compartmentalized Microbes and Co-Cultures in Hydrogels for On-Demand Bioproduction and Preservation," *Nature Communications* **11**, 563. DOI:10.1038/s41467-020-14371-4.

Jung, J. K., et al. 2022. "Programming Cell-Free Biosensors with DNA Strand Displacement Circuits," *Nature Chemical Biology* **18**(4), 385–93. DOI:10.1038/s41589-021-00962-9.

Kampers, L. F. C., et al. 2021. "From Innovation to Application: Bridging the Valley of Death in Industrial Biotechnology," *Trends in Biotechnology* **39**(12), 1240–42. DOI:10.1016/j.tibtech.2021.04.010.

Keasling, J., et al. 2021. "Microbial Production of Advanced Biofuels," *Nature Reviews Microbiology* **19**, 701–15. DOI:10.1038/s41579-021-00577-w.

King, Z. A., et al. 2015. "Next-Generation Genome-Scale Models for Metabolic Engineering," *Current Opinion in Biotechnology* **35**, 23–9. DOI:10.1016/j.copbio.2014.12.016.

Ko, Y. S., et al. 2020. "Tools and Strategies of Systems Metabolic Engineering for the Development of Microbial Cell Factories for Chemical Production," *Chemical Society Reviews* **49**(14), 4615–36. DOI:10.1039/D0CS00155D.

Kuchina, A., et al. 2021. "Microbial Single-Cell RNA Sequencing by Split-Pool Barcoding," *Science* **371**(6531), eaba5257. DOI:10.1126/science.aba5257.

Landwehr, G. M., et al. 2025. "Accelerated Enzyme Engineering by Machine-Learning Guided Cell-Free Expression," *Nature Communications* **16**(1), 865. DOI:10.1038/s41467-024-55399-0.

Lara, A. R., et al. 2006. "Living with Heterogeneities in Bioreactors: Understanding the Effects of Environmental Gradients on Cells," *Molecular Biotechnology* **34**(3), 355–81. DOI:10.1385/MB:34:3:355.

Lee, S. Y., et al. 2019. "A Comprehensive Metabolic Map for Production of Bio-Based Chemicals," *Nature Catalysis* **2**, 18–33. DOI:10.1038/s41929-018-0212-4.

Liew, F. E., et al. 2022. "Carbon-Negative Production of Acetone and Isopropanol by Gas Fermentation at Industrial Pilot Scale," *Nature Biotechnology* **40**(3), 335–44. DOI:10.1038/s41587-021-01195-w.

Ling, C., et al. 2022. "Muconic Acid Production from Glucose and Xylose in *Pseudomonas putida* via Evolution and Metabolic Engineering," *Nature Communications* **13**, 4925. DOI:10.1038/s41467-022-32296-y

Liu, N., et al. 2023. "Application of Artificial Scaffold Systems in Microbial Metabolic Engineering," *Frontiers in Bioengineering and Biotechnology* **11**. DOI:10.3389/fbioe.2023.1328141.

Macklin, D. N., et al. 2020. "Simultaneous Cross-Evaluation of Heterogeneous *E. coli* Datasets via Mechanistic Simulation," *Science* **369**, eaav3751. DOI:10.1126/science.aav3751.

Martin, H. G., et al. 2023. "Perspectives for Self-Driving Labs in Synthetic Biology," *Current Opinion in Biotechnology* **79**, 102881. DOI:10.1016/j.copbio.2022.102881.

Martin, J. P., et al. 2023. "A Dynamic Kinetic Model Captures Cell-Free Metabolism for Improved Butanol Production," *Metabolic Engineering* **76**, 133–45. DOI:10.1016/j.ymben.2023.01.009.

McBee, R. M., et al. 2022. "Engineering Living and Regenerative Fungal–Bacterial Biocomposite Structures," *Nature Materials* **21**, 471–78. DOI:10.1038/s41563-021-01123-y.

Moreno-Gámez, S., et al. 2023. "Quorum Sensing as a Mechanism To Harness the Wisdom of the Crowds," *Nature Communications* **14**, 3415. DOI:10.1038/s41467-023-37950-7.

Morrell, W. C., et al. 2017. "The Experiment Data Depot: A Web-Based Software Tool for Biological Experimental Data Storage, Sharing, and Visualization," *ACS Synthetic Biology* **6**(12), 2248–59. DOI:10.1021/acssynbio.7b00204.

National Security Commission on Emerging Biotechnology (NSCEB). 2025. *Charting the Future of Biotechnology: An Action Plan for American Security and Prosperity*, U.S. Senate, Washington, DC. https://www.biotech.senate.gov/wp-content/uploads/2025/04/NSCEB-Full-Report-%E2%80%93-Digital-%E2%80%934.28.pdf

Nielsen, J., and J. D. Keasling. 2016. "Engineering Cellular Metabolism," *Cell* **164**(6), 1185–97. DOI:10.1016/j.cell.2016.02.004.

Ostrov, N., et al. 2016. "Design, Synthesis, and Testing Toward a 57-Codon Genome," *Science* **353**(6301), 819–22. DOI:10.1126/science.aaf3639.

Pandi, A., et al. 2022. "A Versatile Active Learning Workflow for Optimization of Genetic and Metabolic Networks," *Nature Communications* **13**, 3876. DOI:10.1038/s41467-022-31245-z.

Peruzzi, J. A., et al. 2023. "Engineering Transmembrane Signal Transduction in Synthetic Membranes Using Two-Component Systems," *Proceedings of the National Academy of Sciences of the USA* **120**(19), e2218610120. DOI:10.1073/pnas.2218610120.

Richardson, K. N., et al. 2020. "Aldehyde Production in Crude Lysate– and Whole Cell–Based Biotransformation Using a Noncanonical Redox Cofactor System," *ACS Catalysis* **10**(15), 8898–903. DOI:10.1021/acscatal.0c03070.

Richardson, S. M., et al. 2017. "Design of a Synthetic Yeast Genome," *Science* **355**(6329), 1040–4. DOI:10.1126/science.aaf4557.

Rogers, J. K., et al. 2016. "Biosensor-Based Engineering of Biosynthetic Pathways," *Current Opinion in Biotechnology* **42**, 84–91. DOI:10.1016/j.copbio.2016.03.005.

Roquet, N., et al. 2016. "Synthetic Recombinase-Based State Machines in Living Cells," *Science* **353**, aad8559. DOI:10.1126/science.aad8559.

Rugbjerg, P., et al. 2018. "Diverse Genetic Error Modes Constrain Large-Scale Bio-Based Production," *Nature Communications* **9**, 787. DOI:10.1038/s41467-018-03232-w.

Salvachúa, D., et al. 2021. "Process Intensification for the Biological Production of the Fuel Precursor Butyric Acid from Biomass," *Cell Reports Physical Science* **2**(10), 100587. DOI:10.1016/j.xcrp.2021.100587.

Sandberg, T. E., et al. 2019. "The Emergence of Adaptive Laboratory Evolution as an Efficient Tool for Biological Discovery and Industrial Biotechnology," *Metabolic Engineering* **56**, 1–16. DOI:10.1016/j.ymben.2019.08.004.

Shin, W., et al. 2023. "An Automated Model Annotation System (AMAS) for SBML Models," *Bioinformatics* **39**(11), btad658. DOI:10.1093/bioinformatics/btad658.

Shipman, S. L., et al. 2016. "Molecular Recordings by Directed CRISPR Spacer Acquisition," *Science* **353**, aaf1175. DOI:10.1126/science.aaf1175.

Sullivan, K. P., et al. 2022. "Mixed Plastics Waste Valorization Through Tandem Chemical Oxidation and Biological Funneling," *Science* **378**, 207–11. DOI:10.1126/science.abo4626.

Tickman, B. I., et al. 2022. "Multi-Layer CRISPRa/i Circuits for Dynamic Genetic Programs in Cell-Free and Bacterial Systems," *Cell Systems* **13**(3), 215–29.e8. DOI:10.1016/j.cels.2021.10.008.

Vandelook, S., et al. 2021. "Current State and Future Prospects of Pure Mycelium Materials," *Fungal Biology and Biotechnology* **8**(20). DOI:10.1186/s40694-021-00128-1.

Vardon, D. R., et al. 2016. "*cis,cis*-Muconic Acid: Separation and Catalysis to Bio-adipic Acid for Nylon-6,6 Polymerization," *Green Chemistry* **18**, 3397–413. DOI:10.1039/C5GC02844B.

Vögeli, B., et al. 2022. "Cell-Free Prototyping Enables Implementation of Optimized Reverse β-Oxidation Pathways in Heterotrophic and Autotrophic Bacteria," *Nature Communications* **13**, 3058. DOI:10.1038/s41467-022-30571-6.

Xu, Y. F., et al. 2012. "Ultrasensitive Regulation of Anaplerosis via Allosteric Activation of PEP Carboxylase," *Nature Chemical Biology* **8**, 562–68. DOI:10.1038/nchembio.941.

Yang, Y., et al. 2024. "Improved Enzyme Functional Annotation Prediction Using Contrastive Learning with Structural Inference," *Communications Biology* **7**(1), 1690. DOI:10.1038/s42003-024-07359-z.

Yu, T., et al. 2023. "Enzyme Function Prediction Using Contrastive Learning," *Science* **379**(6639), 1358–63. DOI:10.1126/science.adf246.

Zhang, J., et al. 2020. "Combining Mechanistic and Machine Learning Models for Predictive Engineering and Optimization of Tryptophan Metabolism," *Nature Communications* **11**, 4880. DOI:10.1038/s41467-020-17910-1.

Zhao, E., et al. 2018. "Optogenetic Regulation of Engineered Cellular Metabolism for Microbial Chemical Production," *Nature* **555**, 683–87. DOI:10.1038/nature26141.

Zournas, A., et al. 2025. "Machine Learning-Led Semi-Automated Medium Optimization Reveals Salt as Key for Flaviolin Production in *Pseudomonas putida*," *Communications Biology* **8**, 630. DOI:10.1038/s42003-025-08039-2.

Appendix E

Acronyms and Abbreviations

AI	artificial intelligence
ALE	adaptive laboratory evolution
ATAC-seq	assay for transposase-accessible chromatin using sequencing
ATP	adenosine triphosphate
BER	DOE Biological and Environmental Research program
BSSD	DOE Biological Systems Science Division
CFD	computational fluid dynamics
ChIP-seq	chromatin immunoprecipitation sequencing
CRISPR	clustered regularly interspaced short palindromic repeats
CRISPRa	CRISPR activation
CRISPRi	CRISPR interference
DAP-seq	DNA affinity purification sequencing
DBTL	design-build-test-learn
DOE	U.S. Department of Energy
EBRC	Engineering Biology Research Consortium
EMSL	Environmental Molecular Sciences Laboratory
GEM	genome-scale metabolic modeling
GMO	genetically modified organism
JGI	DOE Joint Genome Institute
KBase	DOE Systems Biology Knowledgebase
LLM	large language model
ML	machine learning
mRNA	messenger ribonucleic acid
P450	cytochrome P450
PTM	post-translational modification
TEA	techno-economic analysis
Tn-seq	transposon insertion sequencing