CHEMICAL REVIEWS

pubs.acs.org/CR

Review

Metabolic Engineering: Methodologies and Applications

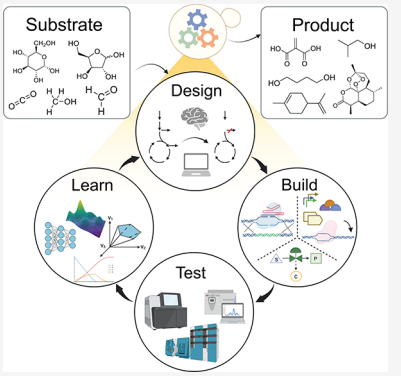
Michael J. Volk,[#] Vinh G. Tran,[#] Shih-I Tan,[#] Shekhar Mishra, Zia Fatma, Aashutosh Boob, Hongxiang Li, Pu Xue, Teresa A. Martin, and Huimin Zhao*

Cite This: *Chem. Rev.* 2023, 123, 5521−5570

Read Online

ACCESS | Metrics & More | Article Recommendations

ABSTRACT: Metabolic engineering aims to improve the production of economically valuable molecules through the genetic manipulation of microbial metabolism. While the discipline is a little over 30 years old, advancements in metabolic engineering have given way to industrial-level molecule production benefitting multiple industries such as chemical, agriculture, food, pharmaceutical, and energy industries. This review describes the design, build, test, and learn steps necessary for leading a successful metabolic engineering campaign. Moreover, we highlight major applications of metabolic engineering, including synthesizing chemicals and fuels, broadening substrate utilization, and improving host robustness with a focus on specific case studies. Finally, we conclude with a discussion on perspectives and future challenges related to metabolic engineering.



CONTENTS

Special Issue: Bridging the Gaps: Learning from Catalysis across Boundaries

Received: June 13, 2022
Published: December 30, 2022

https://doi.org/10.1021/acs.chemrev.2c00403
Chem. Rev. 2023, 123, 5521−5570

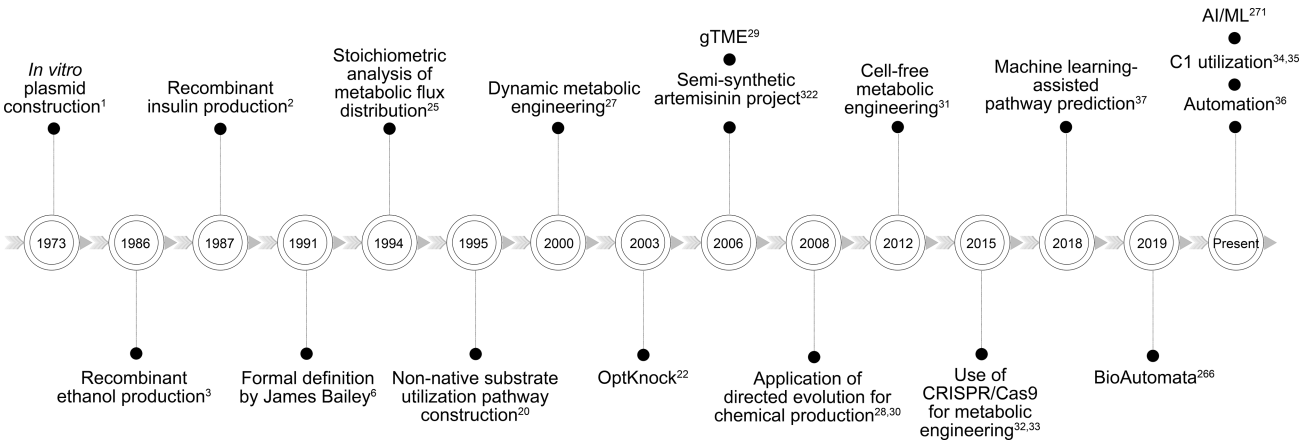


Figure 1. Important milestones in metabolic engineering.

1. INTRODUCTION

In 1973, Cohen and Bailey developed *in vitro* plasmid construction, known as recombinant DNA technology today.[1] This groundbreaking work enabled the biological production of insulin and ethanol by introducing a gene encoding human growth hormone and pyruvate decarboxylase from *Zymomonas mobilis* into *Escherichia coli*, respectively.[2,3] Soon after that, it was realized that the heterologous insertion of a single gene or a biochemical pathway into microorganisms could transform them into tiny chemical factories for the production of molecules. Employing such techniques, a diverse range of chemicals, fuels, pharmaceuticals, and proteins were synthesized by engineered microorganisms in the late 1980s.[4,5] Nevertheless, it was not until 1991 that James Bailey coined the term "metabolic engineering" and defined it as "the improvement of cellular activities by manipulation of enzymatic, transport, and regulatory functions of the cell with the use of recombinant DNA technology."[6] This nascent metabolic engineering field has promptly evolved since then. The early 2000s witnessed the omics era with the development of genomics, transcriptomics, proteomics, and metabolomics, enabling the elucidation of the interactions between these components and uncovering how the host cellular environment changes in response to heterologous pathways or genetic modifications such as gene deletion.[7] Advances in analytical and computational tools, synthetic biology, genome editing tools, and high-throughput technologies developed in the late 2000s further accelerated sophisticated metabolic engineering research.[8−13]

The primary objective of metabolic engineering is to produce functional molecules (e.g., chemicals, fuels, materials, and proteins) using microbial cell factories, while other applications can be found in biological research (e.g., bioremediation and signal transduction) and medical research (e.g., gene therapy and drug discovery).[14−17] In general, metabolic engineering starts with selection of a target product that has high demand and promising applications and selection of a host strain that can be engineered to produce the desired product at high titer, rate, and yield (TRY). Then, a stepwise process of design, build, test, and learn (DBTL) is employed to design and build a native or non-native biochemical pathway that can convert a substrate to a target product, followed by optimization of the metabolic fluxes toward the desired product using tools in systems biology and synthetic biology.

Figure 1 summarizes several significant landmarks in metabolic engineering. Classical metabolic engineering relies on the iterative process of knowledge-guided DBTL in which increasingly better performing strains are constructed based on prior knowledge and intuition.[18−21] Although this approach is widely used to improve the production of chemicals and to extend the substrate range, it is often time-consuming and labor-intensive. To accelerate strain design, several important computational tools, such as metabolic flux analysis, genome-scale metabolic models, and related algorithms like OptKnock, have been established to predict genetic modifications that can lead to higher chemical production.[22−26] Development of synthetic biology tools, such as protein engineering and clustered regularly interspaced short palindromic repeats and CRISPR-associated protein 9 (CRISPR/Cas9), and novel concepts, such as dynamic control and cell-free metabolic engineering, has further facilitated metabolic engineering endeavors.[27−33] More recently, important milestones included metabolic engineering efforts in nonmodel organisms, utilization of C1 compounds, and incorporation of machine learning (ML) techniques.[34−37]

In this review, we first describe the strategies and tools used in metabolic engineering with a focus on the DBTL cycle in the model organisms *E. coli* and *Saccharomyces cerevisiae* because these organisms have been studied and engineered most extensively. Then, we highlight some representative applications of metabolic engineering for production of bulk and fine chemicals and fuels, bioconversion of C1 compounds, and enhancing robustness in model organisms. In addition, we highlight notable examples of metabolic engineering of non-model organisms for production of itaconic acid and 2-propanol/acetone among others. Finally, we briefly discuss the future directions of metabolic engineering, including engineering of non-model organisms, biofoundry development, development and application of ML, utilization of nonsugar substrates, engineering of microbial consortia, and scale-up fermentations.

2. TOOLS

In this section, we provide a brief review of the DBTL cycle used for metabolic engineering of strains for production of chemicals and fuels in the past three decades.

2.1. Design

The first step involved in engineering a strain to produce a molecule of interest is the identification of a biochemical

https://doi.org/10.1021/acs.chemrev.2c00403
Chem. Rev. 2023, 123, 5521−5570

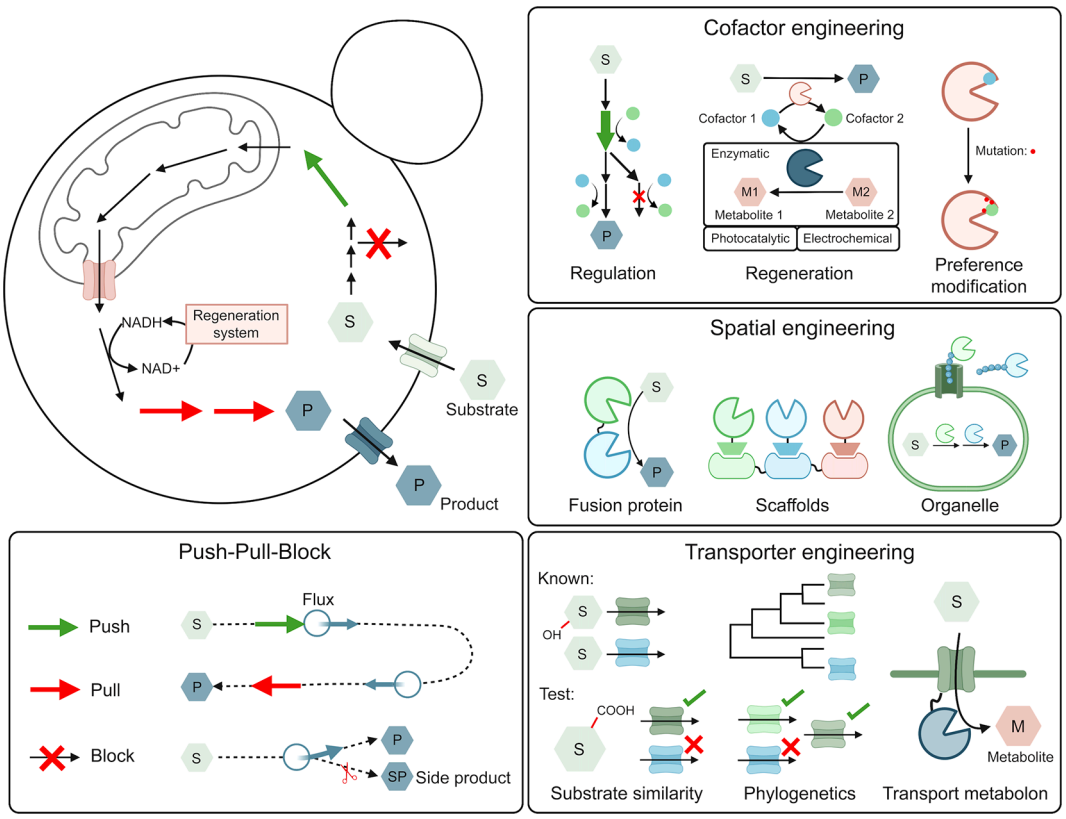


Figure 2. Overview of knowledge-guided approaches for strain design. These approaches can be divided into four categories: push−pull−block (i.e., overexpress or knockout genes to push the carbon flux toward the molecule of interest), cofactor engineering (i.e., perform strain modifications to ensure abundant cofactor supply and maintain cellular redox balance), spatial engineering (i.e., tune the organization of the pathway by mimicking multienzyme complexes or exploit cellular compartments), and transporter engineering (i.e., identify and characterize transporters to facilitate efficient flow across different membrane barriers).

pathway. If a native producer is known, it can be further engineered to improve the titers. However, given the advancement of genetic engineering tools and thorough characterization of metabolism and physiology in model organisms, a more common approach is to construct the pathway in organisms such as *E. coli* or *S. cerevisiae*. The pathway, expressed through plasmid or chromosomal integration, converts a precursor molecule in the host strain to the target molecule. Once the complete biosynthetic pathway is implemented in the host strain, iterative rounds of genetic manipulations are carried out to direct the metabolic flux toward the desired product. In the past, most initial metabolic engineering efforts relied on random mutagenesis and overexpression of a single biosynthetic gene for improving production. The advancement in synthetic biology tools allowed metabolic engineering to continuously tweak enzymatic, transport, and regulatory functions of a cell to create a platform strain for the production of target molecules. Learning from the success stories of the past three decades, various methodologies have been developed based on the knowledge of rewiring cellular metabolism. Moreover, computational approaches such as genome-scale modeling and ML algorithms have been developed to improve metabolic engineering outcomes. Overall, the metabolic engineering strategies for strain design can be primarily divided into three main categories: (1) knowledge-guided design, (2) evolutionary and combinatorial design, and (3) retrobiosynthesis. We summarize these strategies in section 2.1.

2.1.1. Knowledge-Guided Design. Over the last 30 years, the production of numerous molecules has laid the foundation of metabolic engineering as a distinct field. Based on these efforts, several rules have emerged for understanding and making the necessary modifications to the host organisms to improve the TRYs of the desired products. In this section, we provide a brief overview of the strategies implemented based on rational design (Figure 2).

2.1.1.1. Push−Pull−Block. Microorganisms in nature have evolved to improve fitness and maintain metabolic homeostasis in their respective environments. Therefore, the cellular network is highly regulated, and trade-offs are observed when a biochemical pathway is introduced to produce the desired molecule. To improve the TRY, cellular rewiring must be performed. The most straightforward, or the "bread and butter," strategy of researchers to relax some of the regulations and ensure the flow of carbon flux toward the product of interest is to implement the push−pull−block strategy. As the name suggests, the strategy consists of three components: push, driving an increase of flux toward pathway precursors via upregulation of enzymes; pull, drawing higher flux along the biochemical pathway by overexpression of terminal enzymes; and block, removing competing pathways consuming pathway intermediates. Sometimes, overexpression of enzymes in the push strategy is not sufficient as there can be feedback inhibition from the downstream intermediates or the product itself. Therefore, push also includes introducing a heterologous feedback-resistant enzyme to ensure a sufficient precursor pool.[38] Similarly, in the case of a single enzyme catalyzing

https://doi.org/10.1021/acs.chemrev.2c00403
Chem. Rev. 2023, 123, 5521−5570

sequential reactions in a pathway, the block strategy is implemented by mutating the enzyme[39] to prevent leaking of intermediate molecules.

Any metabolic engineering work for efficient production of a molecule generally involves the utilization of all three strategies. However, depending on the target molecule, different versions of push−pull−block are implemented. This includes a two-layer strategy where one of the layers focuses on changing the final product composition while the other layer improves the supply of precursors.[40] Another version is the "push−pull−package−protect" strategy,[41] where package involves modification of vesicles for efficient, degradation-free storage of triacylglycerols (TAG)[42] and protect implies knocking out enzymes to avoid degradation of the desired product.

The requirement of carbon flux for production of the target molecule by the host strain has made the push−pull−block strategy widely applicable irrespective of the target product and the organism in which the metabolic pathway is constructed. In most cases, identifying the gene targets to implement the push−pull−block strategy is relatively straightforward. The advancements of computational learn tools such as genome-scale models, kinetic models, and ML algorithms have further aided in identifying nonintuitive gene targets to assist in implementing the push−pull−block strategy more effectively and are covered in more detail in section 2.4.

2.1.1.2. Cofactor Engineering. Cofactors are known as the helper molecules of enzymes. They can either be organic like NAD^+ or inorganic like iron. They bind proteins and help them carry out their catalytic functions. When a biosynthetic pathway is constructed, some of the heterologous enzymes might not function properly due to the absence of special cofactors or redox imbalance. Examples include the use of multiple cytochrome P450s for production of noscapine, which requires large quantities of NADPH for electron transfer, resulting in possible disruption of $NADPH/NADP^+$ balance.[43] Therefore, multiple strategies have been developed to ensure an adequate supply of cofactors. This includes constructing a cofactor regeneration system, enhancing cofactor levels by regulating the native machinery, and even modifying the enzyme's cofactor preference.

The absence of special cofactors is a major problem in the biosynthesis of plant specialized metabolites such as opioids, alkaloids, and terpenes. Another factor that plays a role is the subcellular location of the cofactor. For example, the lipoylation machinery is only present in yeast mitochondria. Therefore, a functional pyruvate dehydrogenase cannot be expressed in the cytosol. This can be resolved by reconstituting the necessary machinery in the necessary compartment for biosynthesis and regeneration of the desired cofactor.[43,44] Similarly, cofactor supply can also be enhanced by regulating the endogenous system.[45−47]

Another cofactor engineering strategy to maintain redox balance is to change the cofactor preference. Liu and co-workers modified the cofactor specificity of the redox reaction-related enzymes, glyceraldehyde-3-phosphate dehydrogenase and malate dehydrogenase, in the *N*-acetylglucosamine biosynthetic pathway.[48] Protein structure-guided comparison helped in identifying amino acid residues to convert the cofactor preference from NAD^+ to $NADP^+$. Computational tools such as CSR-SALAD[49] are also available to screen for mutation sites *in silico* for changing the cofactor specificity. Further incorporation of ML-guided protein-structure prediction models will benefit the field of cofactor engineering. More details on cofactor engineering can be found in these reviews.[50,51]

2.1.1.3. Spatial Engineering. An emerging strategy to improve the yield and titers of biosynthetic pathways is the spatial organization of enzymes inside cells. Organizing the pathway can boost local enzyme concentrations, remove diffusion limitations, limit accumulation of toxic intermediates, as well as bypass competing enzymes and inhibitory regulatory networks. A variety of strategies have been implemented using synthetic biology tools to exploit these advantages. We divide these strategies into three categories: (1) fusion proteins, (2) synthetic scaffolds, and (3) compartmentalization.

The initial efforts in spatial engineering were inspired by multienzyme complexes which facilitate efficient substrate channeling. To replicate these complexes, a simple approach is to create a synthetic fusion protein by linking two or more genes with a linker sequence. Examples include the construction of a farnesyl pyrophosphate (FPP) synthase−farnesene synthase fusion protein,[52] which reduces the loss of the FPP pool to competing metabolic reactions. Similarly, a 6-fold improvement in the production of raspberry ketone was achieved by utilizing a synthetic fusion protein of coumarate−CoA ligase and benzalacetone synthase.[53] Fusion proteins primarily aid in removing diffusion limitations and might also benefit from protein−protein interactions. However, strenuous characterization of linker sequences with varying length, composition, and enzyme orientation has limited the applications of fusion proteins. Recent advancements in structure prediction have further aided in resolving these issues. For example, a rationally designed biocatalytic cascade of NADH-dependent alcohol and aldehyde dehydrogenase was constructed using Rosetta software to determine the optimal linker placement.[54] Adding the linker close to the active site of the enzymes resulted in electrostatic channeling of acetaldehyde, helping achieve a 500-fold improvement of catalyst turnover frequency compared to the unbound enzymes. To incorporate more enzymes in the metabolic complexes, a better approach is attaching pathway enzymes to a scaffold. Synthetic scaffolds have been created from RNA binding domains,[55] protein−protein interaction domains,[56] and even DNA molecules that are recognized by zinc-finger fused enzymes.[57] Compared to fusion proteins, synthetic scaffolds provide modular control over enzyme stoichiometry and, therefore, the metabolic flux.

The third strategy is compartmentalization. In recent years, unorthodox strategies to spatially localize enzymes in an organelle or *in vitro* compartments have effectively boosted the yield and titer of biochemicals. Apart from the advantages of spatial engineering, organelles provide a unique physicochemical environment that can offer favorable conditions for different metabolic products. This includes varying pH, redox potential, and an adequate supply of essential cofactors and precursors. To realize the benefits of subcellular engineering, well-characterized protein localization tags have been used to localize enzymes in organelles to improve the production of various molecules in diverse organisms.[58−61] We refer readers to a recent in-depth review covering organelle engineering in yeasts.[62] Besides utilizing resources of a single organelle, comprehensive engineering of multiple subcellular compartments is gaining interest. Higher titers have been achieved for isoprene,[63] squalene,[64] astaxanthin,[65] and fatty acid methyl esters[66] by taking advantage of precursors' presence in multiple

https://doi.org/10.1021/acs.chemrev.2c00403
Chem. Rev. 2023, 123, 5521−5570

locations. Apart from only improving production, organelle localization has been implemented to provide spatial separation for improving pathway specificity.[67] Similarly, bacterial microcompartments have been employed for propanediol utilization.[68] Synthetic biology toolkits have also been used to perform *in vitro* compartmentalization[69] and light-based synthetic organelle assembly[70] to cluster enzymes together. Currently, most of the work in the field primarily focuses on targeting pathways to compartments. In a recent work, the overexpression of *INO2*, a regulatory gene in yeast phospholipid biosynthesis, enhanced the activity of the endoplasmic reticulum (ER) and resulted in improved squalene and protopanaxadiol production.[71] Therefore, improving shape, size, and organelle activity by alleviating the compartment-specific constraints should also be prioritized in the future. Overall, the field of spatial engineering has demonstrated the ability to substantially improve the production of various molecules.

2.1.1.4. Transporter Engineering. A rather overlooked but important aspect of strain design is the transportation of molecules. The process of transportation is involved in every aspect of the metabolic pathway starting from the substrate uptake, efficient intermediate transportation, to the export of products. Here, we summarize a range of experimental techniques and computational models developed to engineer transporters of interest.

The first step in the production of target molecules is substrate utilization. Most of the initial work relied on the use of glucose as a carbon source, but there has been growing interest in converting lignocellulose-type renewable feedstocks, plant extracts, and pulp waste to desired molecules. However, in model organisms like *S. cerevisiae*, the transportation of nonglucose substrates across the membrane is often rate-limiting because of very slow uptake or the absence of necessary import machinery. To overcome this challenge, the identification and characterization of putative heterologous transporters is necessary. For example, the transporters for D-xylose and L-arabinose were identified using various bioinformatics techniques such as genome mining via sequence similarity,[72] transcriptomic analysis, and phylogenetics.[73] Similarly, transporters were characterized for other sugar forms obtained from different feedstocks.[74,75] Because most of the biomass hydrolysates are a mixture of multiple sugars, Because biomass hydrolysates consist of a mixture of multiple sugars, strategies have been developed for substrate co-utilization while overcoming glucose repression[76] and further improved by engineering mutant transporters with directed evolution.[77] Furthermore, inspiration from membrane transport metabolons has led to the design of an artificial enzyme-transporter complex to further enhance substrate utilization.[78]

Early work on the transport of pathway intermediates primarily focused on reuptake[79,80] of the metabolites or avoiding leakage through transporters.[81,82] Emerging methodologies for organelle engineering have given transporter design a new dimension. Smolke and Srinivasan expressed 26 genes across six subcellular locations to synthesize medicinal tropane alkaloids in yeast.[83] Imitating the extensive intra- and intercellular compartmentalization inherent to the plant biosynthetic pathways required addressing the transport limitations across organelle membranes by characterization of necessary plant transporters. A follow-up study[84] implemented an artificial neural network to prioritize transport candidates to screen for tropane alkaloid exporters across the vacuole and improved the production of hyoscyamine and scopolamine by 2- and 1.5-fold, respectively. Apart from increasing titers of a molecule, product preference can be switched by reprogramming interorganelle transporters.[85] For example, systematic identification and replacement of mitochondrial metabolite carriers in *Yarrowia lipolytica* resulted in a complete switch from production of citrate to isocitrate.

Transporters also play an important role in exporting the final product. The advantages include simplified downstream processing, decrease in inhibitory interactions and potential cellular toxicity due to product accumulation, and higher production resulting from a shift in the chemical equilibria. Mukhopadhyay and co-workers employed a bioinformatics screen to identify novel transporters for biofuel tolerance.[86] A library of 43 efflux pumps was screened using a competitive growth assay, and the tolerance was improved for 5 chemicals. Similarly, Dong and co-workers screened transporters for medium-chain fatty acid (MCFA) secretion.[87] A combination of multiple transporters showed a synergistic effect and resulted in more than 2-fold improvement in MCFA production. In another work, a substrate similarity search approach developed by Wang and co-workers exploited the promiscuity of export systems of similar chemicals for malate export.[88]

Transporter engineering is covered for interested readers in these extensive reviews.[89,90] However, in addition to rigorous experimental characterization, more studies need to focus on developing gene mining tools and ML models for *in silico* discovery of transporters.

2.1.2. Evolutionary and Combinatorial Design. Evolutionary design aims to evolve a wild-type or platform strain to a strain that can produce target molecules at improved TRY. Variations of evolutionary design are the most tried and reliable methods for iterative strain improvement. Evolutionary design predates rational and computational design strategies as no causal or mechanistic understanding is necessary for finding improved phenotypes. Evolutionary design relies on the principles of evolution, namely diversification and selection, which in the case of metabolic engineering is typically diversification of DNA and selection of improved mutant strains. These methods are often categorized as adaptive laboratory evolution (ALE), directed evolution (DE), or continuous evolution (CE).

While ALE can be considered a continuous evolution method, it deserves to be commented on separately as it was developed first and is still the most used evolutionary technique. In ALE, random variation in genomic DNA acts as the source of diversification, and growth is used for selection. The defining feature of ALE is that it relies on growth-dependent phenotypes. Common selection criteria include growth rate, tolerance to a range of different chemical or physical stressors, substrate utilization, and growth coupled molecule production. A distinct advantage of ALE associated growth phenotypes is that multistress resistance can be engineered simultaneously by applying selection pressures in series or in parallel.[91] One major disadvantage of ALE is that the mutagenic rates are often low and mostly limited to point mutations. This can lead to long experiments,[92] in some cases years long.[93] We refer readers to a recent in-depth review on ALE.[94]

Historically, DE has been associated with the engineering of individual genes or proteins instead of polygenic phenotypes, but more recently, DE methods have been expanded to

https://doi.org/10.1021/acs.chemrev.2c00403
Chem. Rev. 2023, 123, 5521−5570

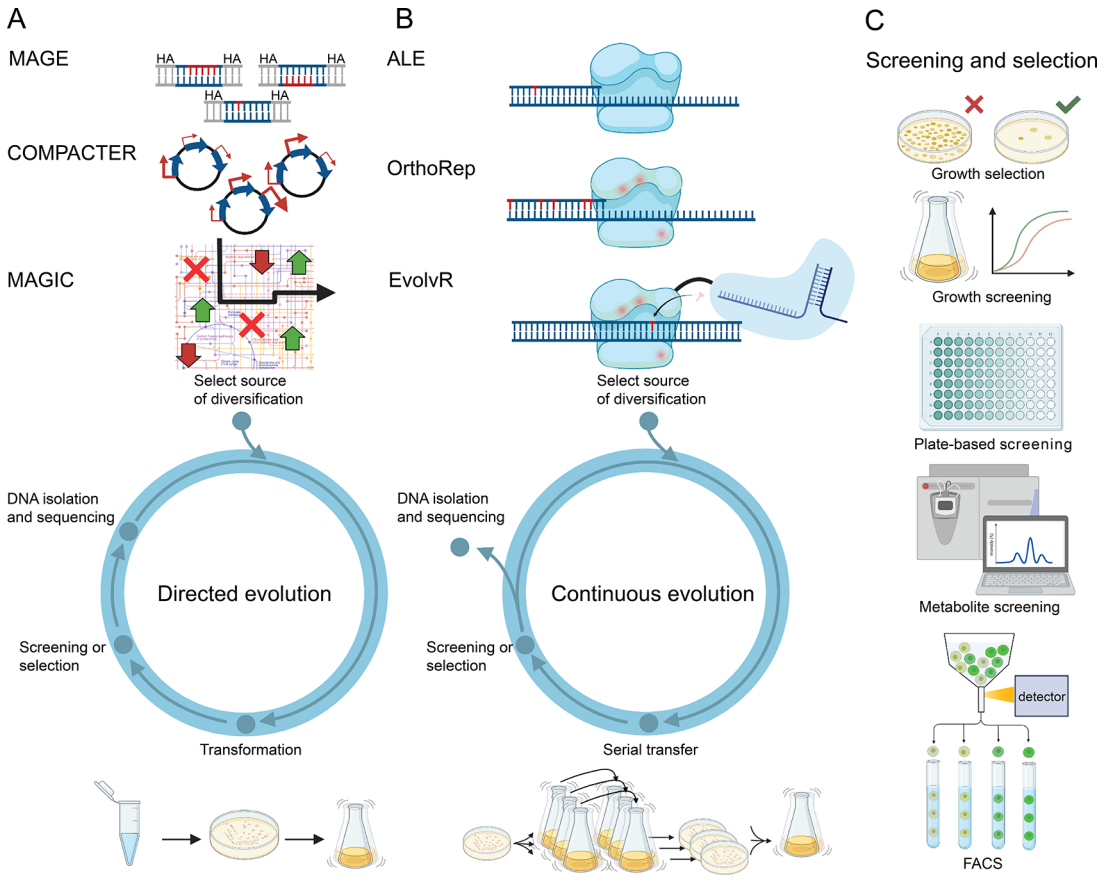


Figure 3. Evolutionary and combinatorial design. (A) The directed evolution (DE) workflow where the source of diversification is selected to target a specific level of the genome. Example sources of diversification include MAGE, COMPACTER, and MAGIC. DNA is transformed and then mutants go through screening/selection. DNA must be isolated during each round so positive hits can be recorded or so permanent modifications can be made on a new parent strain. (B) Continuous evolution (CE) begins with the source of diversification where the simplest version of continuous evolution is ALE in which the source of diversification is limited by the mutagenic rate of wild-type polymerase. Newer methods of CE increase the mutagenic rate of mutation with error prone replication shown with OrthoRep and EvolvR. In the case of EvolvR, a fused dead CRISPR/Cas system allows for more targeted mutations. Mutants are propagated though serial transfer and then go through screening/selection. DNA is only isolated and sequenced according to chosen intervals or at the terminating cycle, which greatly reduces labor and helps increase mutant diversity. (C) Screening and selection methods for evolutionary design. The top two methods shown are dependent on growth and for this reason are often associated with CE methods, specifically ALE. Plate-based screening and metabolite screening are examples of low-throughput methods that are often associated with DE. FACS, while high throughput is expensive and time-consuming and is therefore often used in DE because it only needs to be run a few times. Nothing in principle precludes these methods from being used in CE.

engineer pathways and entire genomes.[95] DE induces diversification with a range of different *in vivo* and *in vitro* methods and these designs can be further reduced in size with computational tools to remove combinatorial elements that are unlikely to be successful. Because DE starts with a more diverse set of candidates than ALE, it can arrive at a suitable design through fewer cell generations and is typically associated with more complex, lower throughput screening/selection strategies where selection methods select growth-linked phenotypes and screening methods evaluate each variant and choose only high performers. DE-associated low throughput screening techniques include microtiter plates, chromatography, mass spectrometry (MS), agar plating, and droplet microfluidics. DE is also amenable to high-throughput selection methods used by ALE and additional high-throughput screening methods like fluorescence-activated cell sorting (FACS) (Figure 3C). One major limitation of DE methods is that they are limited by the transformation efficiency of the host strain which creates an upper limit for library size. We refer readers to a recent in-depth review on DE.[95]

As a mix of DE and ALE, CE aims to generate gene diversification and perform screening/selection without any human intervention. CE methods belong to a broader classification of *in vivo* diversification methods, which includes ALE, but more recent methods have sought to increase mutation rates over basal mutation rates, as in ALE, by using an error-prone polymerase, a deaminase, or by other means. CE methods operate most similarly to ALE, but they require less labor than DE and can generate even greater diversity than DE because they are not limited by transformation efficiency. A few representative CE systems for increasing mutational diversity include OrthoRep,[96] EvolvR,[97] and Muta T7 (Figure 3B).[98] Like DE, CE is amenable to using a more varied list of selection or screening methods in comparison to ALE. Key limitations of CE systems include that they must be properly built and validated in the host of interest if there is not already an established CE system and that each CE system is limited to

https://doi.org/10.1021/acs.chemrev.2c00403
Chem. Rev. 2023, 123, 5521−5570

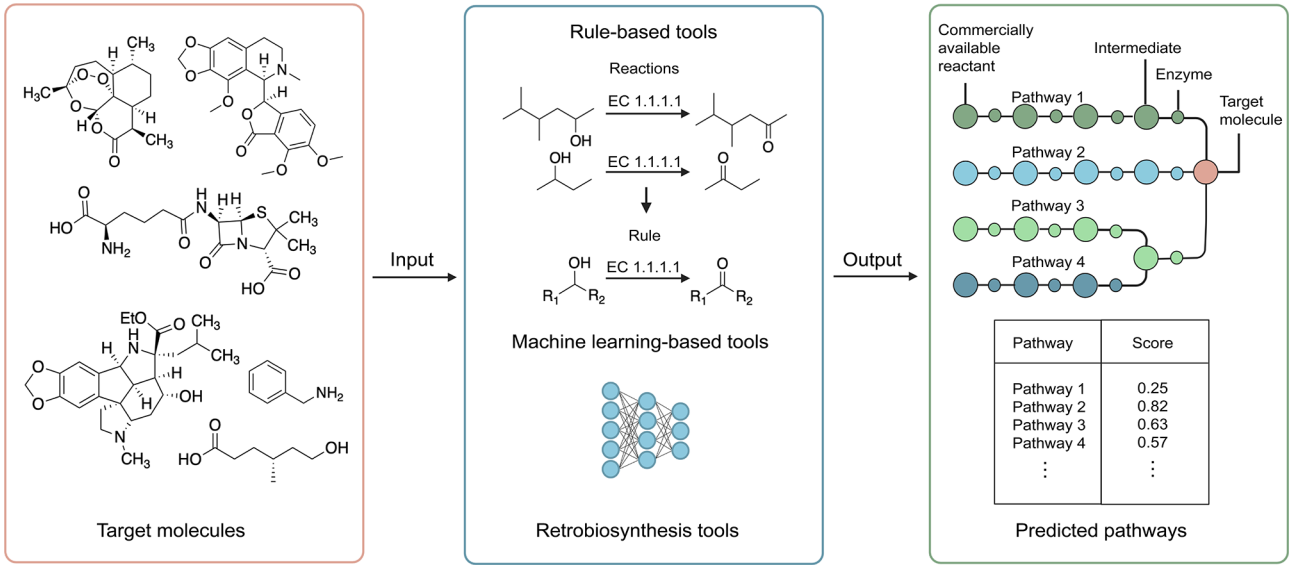


Figure 4. Workflow of retrobiosynthesis. Retrobiosynthesis tools can be divided into two categories, namely rule-based tools and ML-based tools. The rule-based tools can extract reaction rules from a database and then predict pathways based on those rules, while the ML-based tools will use the reactions in a database to directly guide the construction of new pathways. Given the target molecules as an input, both tools can output pathways including several intermediates and enzymes, along with scores indicating the confidence in those pathways.

specific types of mutational patterns. We refer readers to a recent in-depth review on CE.[99]

Combinatorial design approaches include methods for generating diversification with varied sequence or structural patterns. In the case of DNA mutagenesis,[95] *in vitro* diversification is the most common method due to its ease of control and high efficiency, but there have been recent advances in *in vivo* diversification, with opportunities for making the entire process continuous with CE. Here we give a few *in vitro* examples focusing on different types of cellular manipulation. Multiplexed genome engineering (MAGE) allowed for the editing of multiple targets with DNA mismatches, deletions, and small insertions on the genome for the generation of large combinatorial libraries and was used for increasing *E. coli* lycopene production.[100,101] Customized optimization of metabolic pathways by combinatorial transcriptional engineering (COMPACTER) allowed for the recombination of pathway genes on a plasmid with varied promoter strengths, resulting in a library of pathways with varied expression.[102] More recently, CRISPR/Cas methods have been used for precise genome-scale multiplexing. One such example is the multifunctional genome-wide CRISPR (MAGIC) system, which creates diversity by constructing a genome-scale plasmid library for gene activation, interference, and deletion (Figure 3A).[103] We refer readers to further details on evolutionary and combinatorial design with a recently published book.[104]

2.1.3. Retrobiosynthesis. Retrobiosynthesis tools are computational tools that predict metabolic pathways for target molecules. The fact that they can enumerate reliable pathways based on databases makes them popular auxiliary tools for metabolic engineers. There are two types of retrobiosynthesis tools: rule-based tools and ML-based tools. The rule-based tools can transform compounds in databases into readable strings and then extract reaction rules from them, including reaction centers for both reactants and products with their changes. The range of the reaction centers will determine the accuracy of those rules. When a target molecule is input, the search algorithm will find optimal pathways within the extracted rules according to molecular similarity. However, if the input does not match with existing rules, the rule-based tools cannot predict properly. In contrast, the ML-based tools will not explicitly extract rules, they instead use the reactions in the database to directly guide the construction of new pathways. Therefore, the ML-based tools are not limited by rigid rules but may not be as accurate as rule-based tools for known reactions due to the lack of templates. For both complementary tools, the predicted pathways will start from the target molecule, trace back through several intermediates, and finally reach a starting metabolite, suggesting enzymes for each reaction. If many pathways are available for a certain molecule, the evaluation metrics such as yield, toxicity, and thermostability are used to assign a score to each pathway for ranking. SCScore[105] is a widely used score system that captures the synthetic complexity of each molecule (Figure 4).

In the most recent retrobiosynthesis tools, some additional functions are added to enhance their performance. A specialized rule extracting tool called rePrime[106] was integrated with novoStoic[106] to complete all the tasks from extracting rules to predicting pathways. In another example, Monte Carlo tree search reinforcement learning was combined into the previously developed ranking system in RetroPathRL[107] in order to predict longer pathways. In RetroBioCat[108] and novoPathFinder,[109] web servers were adopted to make the tools more user-friendly. Furthermore, Laino and co-workers modified the molecular transformer to create the first ML-based retrobiosynthesis tool, suggesting there will be future applications of ML algorithms in this field.[110]

In order to construct metabolic pathways, retrobiosynthesis tools are often used together with enzyme identification and directed evolution tools.[111] While a series of optimal pathways are suggested by the retrobiosynthesis tools, other tools will predict the probability that a certain enzyme or mutant can be used in each pathway, contributing to a complete pathway with high yield and selectivity. In metabolic engineering, the retrosynthesis tools have been used to guide experiments and

https://doi.org/10.1021/acs.chemrev.2c00403
Chem. Rev. 2023, 123, 5521−5570

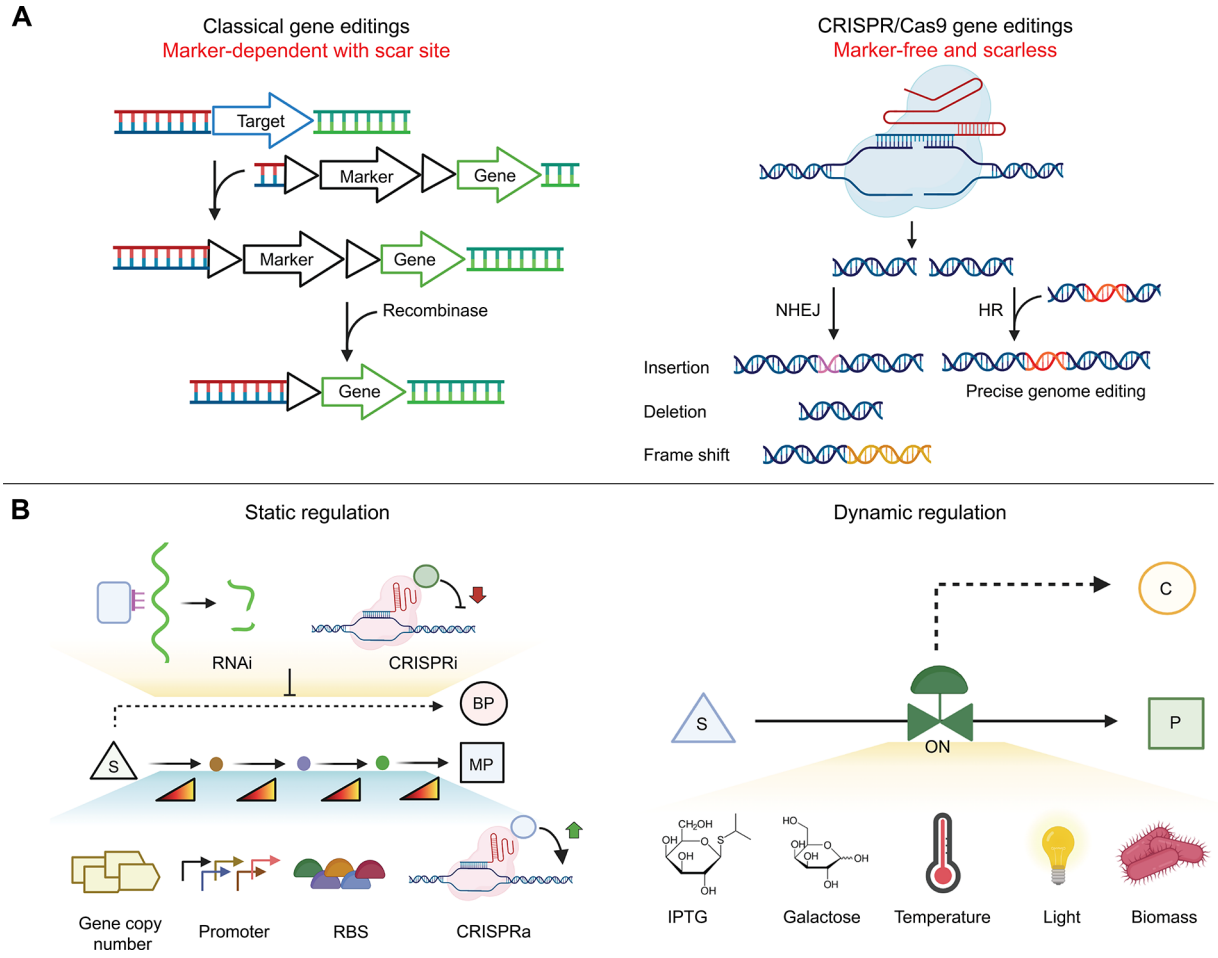


Figure 5. Tools for synthesis in metabolic engineering. (A) Genome editing tools including marker-dependent classical genome editing tools and marker-free CRISPR/Cas9-based genome editing tools. (B) Tools for gene regulation in static and dynamic manners.

the predicted pathways have been tested by *in vitro* and *in vivo* experiments to achieve synthesis of desired products from glucose.[112]

2.2. Build

Once the biochemical pathway to produce the target molecule has been identified using the design strategies described above, the pathway needs to be inserted into a microorganism. Due to the increase in genomic stability, integration of the biochemical pathway into a host chromosome is preferred and can be achieved using genome editing tools. Furthermore, high TRY may not be achieved just by simply expressing the biochemical pathway. To obtain maximal titers, gene regulation tools are utilized to optimize heterologous pathways and host strains at different cellular levels to avoid imbalanced gene expression, which can lead to over- or under-expression of enzymes in the pathway and accumulation of toxic metabolic intermediates. In this section, we describe some traditional genome modification tools and modern genome editing tools based on the CRISPR/Cas system. We also discuss some synthetic biology tools to regulate gene expression in static and dynamic manners.

2.2.1. Genome Editing Tools. Because episomal plasmids with varying copy numbers enable fast, easy, reliable, and well-regulated gene expression, they are the preferred method for gene expression in proof-of-concept metabolic engineering.[113] Nevertheless, replicating plasmids suffer from structural, segregational, and allele distribution instability in selective media. Furthermore, antibiotics are necessary to maintain episomal plasmids, increasing the expense of overall bioprocess cost. Therefore, chromosomal integration of the biochemical pathway is more favorable than plasmid-borne overexpression for maintaining genetic stability and consistent expression throughout a cultured population and for decreasing the metabolic burden and cell-to-cell variability in the absence of selection pressures. Genome editing tools can be used for chromosomal integration as well as for the removal of competing pathways to minimize byproducts and divert carbon and cofactor resources to target products. Here we describe some genome editing tools enabling pathway integration and gene deletion, including classical genome editing tools and modern CRISPR/Cas9-based genome editing tools (Figure 5).

2.2.1.1. Classical Genome Editing Tools. Traditionally, gene knockouts were achieved in bacteria and yeast by homology-directed integration, which relies on homologous recombination (HR) of a donor double-stranded DNA, which can be in the form of a disruption cassette or a suicide plasmid.[114−116] The donor DNA, containing a dominant marker and DNA sequences homologous to the upstream and downstream regions of the gene targeted for deletion, is transformed into a microorganism of interest. HR then leads to the deletion of the target genes, and correct mutants can be selected via the positive selectable markers present on the donor DNA. Unlike *S. cerevisiae*, which possesses a very strong

https://doi.org/10.1021/acs.chemrev.2c00403
Chem. Rev. 2023, 123, 5521−5570

HR machinery, bacteria such as *E. coli* typically have poor HR efficiency, and the red system from the λ phage is expressed to enhance the HR activity.[117] Then, an additional step for elimination of the selection markers from the chromosome can be performed using site-specific recombination to enable subsequent genetic modifications of the deletion strains. Flippase (Flp) and Cre are the commonly used site-specific recombinases that can perform recombination events at two identical FRT and loxP sites, respectively. The dominant marker gene is flanked by the recombination sites and removed after chromosomal integration by expressing the corresponding recombinase. This approach of coupling HR and site-specific recombination has been applied for gene deletion in the metabolic engineering of bacteria and yeasts for the production of various chemicals for decades.[118−124] Nevertheless, the main disadvantage of this method is that various loxP or FRT scars are left behind after several rounds of gene deletion. These multiple identical copies of the scar can serve as future sites for recombinase and thus potentially induce genetic instability. Furthermore, this system is time-consuming, inefficient, and low-throughput.

Genome editing tools can also be used for chromosomal integration of biochemical pathways, which is the preferable method for gene expression compared to plasmid-based expression. In addition to gene deletion, the conventional λ-Red recombineering system can enable chromosomal integration of DNA modules with a size of up to 2,500 bp in bacteria.[125] Furthermore, several λ-Red-based techniques for chromosomal integration of large biochemical pathways were developed, such as I-SceI cleavage-facilitated recombination, knock-in/knockout vector-mediated integration, and pSB1K3-(FRTK) vector aided insertion.[125] Notably, integration of an 8.9 kb synthetic Entner−Doudoroff pathway was one of the large size integrations that could be achieved using λ-Red machinery.[126] Integration of pathways larger than 10 kb in a single step using this system still remains a challenge. To address this limitation, the large pathway can be divided into smaller segments and inserted into the genome through iterative integration. For example, a 15 kb DNA encoding sucrose and lactose consumption pathways were previously divided into four fragments around 3 kb each and then iteratively integrated into the *E. coli* chromosome.[1] In addition to the λ-Red recombineering system, integrases can be used to permanently introduce biochemical pathways to the genome.[127] Integrases perform site-specific recombination of small sequences of DNA called attachment (att) sites to precisely rearrange DNA. The recombination event catalyzed by serine integrases is highly directional and can only be reversed using an accessory protein called a recombination directionality factor. Previously, an intermolecular site-specific TG1 integration system was developed to enable efficient recombination between the attB site on plasmid DNA and the corresponding att site on the genome of *E. coli*.[128] Furthermore, the two compatible integrase systems from phages ΦC31 and ΦBT1 enabled successive three- and two-copy integration of pristinamycin II biosynthetic gene cluster in *Streptomyces pristinaspiralis*.[129] A phage serine integrase-mediated site-specific genome engineering technique also enabled the integration of a butyric acid production pathway from *Clostridium acetobutylicum* into the *Clostridium ljungdahlii* genome.[130]

Several classical tools have been developed for the stable integration of biochemical pathways via HR in *S. cerevisiae*. The series of yeast pRS integration vectors enable single-locus integration and can be used to introduce a heterologous gene into *S. cerevisiae* by homologous integrative recombination events at auxotrophic loci.[131] Nevertheless, superfluous sequences, including the bacterial selection marker and origin of replication of *E. coli*, are also integrated, which might lower the integration efficiency. DNA assembler is another single-locus integration technique and was developed for chromosomal integration of multiple expression constructs, such as xylose utilization and zeaxanthin biosynthetic pathways, with high efficiencies.[132] Nevertheless, this method discourages the use of repetitive elements, such as promoters and terminators, which can result in direct repeat recombination. To address this limitation, a set of *URA3*-based plasmids was designed to enable stable integration of multigene pathways using identical promoters or terminators or multiple copies of the same gene.[133] The integration sites were separated by genetic elements necessary for growth to prevent strain propagation if genes were lost through direct repeat recombination. Strategies for efficient multilocus chromosomal integration in *S. cerevisiae* were also developed to facilitate strain engineering. The Directed-Pop-Out Plasmid system allowed a simple two-step scarless integration method and contained a new set of counterselectable markers, which enabled serial integration followed by a transformation-free marker rescue event.[134] In another study, a new set of integrative vectors, Easy-CloneMulti, which combined consensus sequences targeting Ty sequences and a quickly degrading selection marker, was constructed to enable simultaneous integration of multiple genes in *S. cerevisiae*.[135]

2.2.1.2. CRISPR/Cas-Based Genome Editing Tools. While the classical methods for genome editing have been widely applied in metabolic engineering, a dominant marker must be integrated into the chromosome for selection purposes. Marker removal necessitates the use of recombinase, which renders the genome editing process time-consuming and laborious. The recently developed CRISPR/Cas system has revolutionized metabolic engineering and allowed elegant yet powerful marker-less genome editing.[136] A single guide RNA (sgRNA), which contains a spacer sequence complementary to the targeted DNA sequence, forms a complex with a DNA endonuclease enzyme Cas9. The complex then binds to the genomic target dictated by the spacer sequence. Upon binding, Cas9 creates a DNA double-strand break (DSB), which is then repaired by either HR with a homologous repair donor or nonhomologous end joining (NHEJ). NHEJ is intrinsically mutagenic and creates random indels at the cleavage site. Indel mutations occurring in coding exons may introduce premature stop codons or frame-shift mutations, leading to the deletion of the corresponding proteins. HR is the preferred method for DNA repair in genome editing because it allows precise and accurate editing. Furthermore, because the DSB created by the CRISPR system is lethal, HR serves as a positive selection, making marker-free integration possible. In microorganisms in which NHEJ is predominant over HR, such as *Y. lipolytica*, disruption of *KU70* and *KU80* genes, which are responsible for DSB repair in the NHEJ pathway, can cause the cells to switch to mainly HR-based DNA repair.[137] Nevertheless, deletion of NHEJ-related genes can potentially decrease the overall DNA transformation efficiency.[138] Overall, while off-target modifications is a possible occurrence, CRISPR/Cas9 is an indispensable tool in metabolic engineering and has been implemented with high editing efficiencies in various

https://doi.org/10.1021/acs.chemrev.2c00403
Chem. Rev. 2023, 123, 5521−5570

organisms. Here we discuss the application of CRISPR/Cas9 for gene deletion, integration of biochemical pathways, and substitution and diversification of biological parts. We refer readers to some recent in-depth reviews covering CRISPR/Cas-based systems for genome editing tools.[136,139]

CRISPR-Mediated Gene Deletion. To achieve gene knockout using the CRISPR/Cas9 system, a sgRNA can be used to introduce a DSB inside the target gene followed by repair using an HR donor carrying homology arms flanking the targeted region. The first CRISPR/Cas9-mediated genome editing system for *E. coli* employed ssDNA as an editing template and could achieve gene deletion with 65% efficiency.[33] The CRISPR/Cas9 system for *E. coli* was later improved and could perform gene deletions with near 100% editing efficiency using both ssDNA and dsDNA as editing templates.[33] It was observed that the efficiency with ssDNA as an editing template decreased dramatically with the increased length of the deleted sequence, while the efficiency with dsDNA was more than 90% with a deleted sequence as long as 12 kb. For *S. cerevisiae*, the first CRISPR/Cas9 system employed a 90-bp double-stranded oligonucleotide as the HR template.[140] The donor was centered around the protospacer adjacent motif (PAM) sequence and contained 2 bp changes to mutate the PAM sequence and incorporate a premature TAG stop codon. This system was able to achieve 100% knockout efficiency of the *CAN1* gene. Because desired genome editing could only be attained when all components (i.e., Cas9, sgRNA, and HR donor) are transformed into the same cell, multiplex genome editing can suffer low efficiency when multiple DNA molecules need to be cotransformed. To counter this issue, the homology integrated CRISPR/Cas (HI-CRISPR) system was developed, in which an HR donor was harbored on the CRISPR/Cas9 plasmid to ensure cotransformation of both the CRISPR/Cas9 system and HR donor. HI-CRISPR could perform multiplex gene knockout of *ATF2*, *GCY*1, and *YPR1* with an efficiency of 100%.[141] A CRISPR/Cas9 system was also developed for gene deletion in microorganisms with strong NHEJ such as *Rhodosporidium toruloides*.[142] An HR repair template was not necessary because the NHEJ pathway could generate indels near the cut site, resulting in frameshift mutations. Nevertheless, precise and controlled site-specific gene deletions can be challenging in these organisms.

CRISPR-Mediated Integration of Biochemical Pathways. Marker-free integration of biochemical pathways can be achieved using the CRISPR/Cas9 system. As an early application in metabolic engineering, the CRISPR/Cas9 system was used to integrate the β-carotene synthetic pathway and several genomic modifications into the genome of *E. coli*.[33] Furthermore, a multistep CRISPR/Cas9-based chromosomal integration strategy was developed by dividing a 9.7 kb DNA fragment, including eight genes in the pyrimidine operon of *Bacillus subtilis* F126 into multiple small fragments with the appropriate size followed by successive integration of those fragments into *E. coli*.[143] In another study, CRISPR/Cas9 was used to integrate into *S. cerevisiae* a lactose transporter from *Kluyveromyces lactis* and a heterologous 2′-fucosyllactose (2′-FL) biosynthetic pathway consisting of enzymes Gmd, WcaG, and WbgL from *E. coli*.[144] Multiplex integration can also be achieved using CRISPR/Cas9, in which sgRNAs targeting multiple sites are coexpressed or a sgRNA can be used to target repetitive sequences in the genome. In one study, a sgRNA was expressed targeting multiple repeat sequences called delta sites in the *S. cerevisiae* genome, and up to 18 copies of a set of genes for xylose utilization and (*R*,*R*)-2,3-butanediol (BDO) production pathway were integrated at the delta sites in a single step.[145]

CRISPR-Mediated Substitution and Diversification of Biological Parts. CRISPR/Cas9 systems can also be used to substitute native promoters with stronger constitutive promoters or introduce mutations to existing promoters. This same technique can be applied to other biological parts as well. For example, an *E. coli* strain was engineered to produce isoprenoid by genomic integration of the *cscAKB* operon and a heterologous mevalonate (MVA) pathway. CRISPR/Cas9 was used to substitute the promoters driving the MVA pathway with the stronger T7 promoter.[146] In another study, the expression of the first enzyme in the citric acid cycle, citrate synthase, was reduced by modifying the promoter region.[147] Four different 5′-untranslated region sequences with different strengths were designed and directly edited on the genome using the CRISPR/Cas9 system.

2.2.2. Gene Regulation Tools. Optimizing the flux through a biochemical pathway to maximize production of the target molecule demands tightly regulated and consistent expression of the genes encoding the enzymes in all cells in a cultured population.[148] Gene regulation systems can be broadly classified into static regulation and dynamic regulation.[149] Static regulation systems are traditionally used in metabolic engineering and involve the constitutive expression of biochemical pathways and tuning of these pathways through strategies such as promoter engineering, ribosome binding site (RBS) engineering, and gene copy number engineering. Engineered metabolic pathways can compete with the host endogenous machinery for cellular resources, such as RNA polymerases, ribosomes, ATP, and cofactors.[150] Thus, these engineering pathways can place a metabolic burden on the host by draining resources from essential metabolism or causing improper cofactor balance. In some cases, the expression of heterologous pathways can lead to growth defects and can reduce the production of the target product, which is further exacerbated if the final product or pathway intermediates are toxic. On the other hand, the development of dynamic control systems is motivated by natural metabolic control mechanisms, which enable microorganisms to adjust metabolic flux and maintain homeostasis under new environments.[151] Dynamic control uses a variety of signals to activate or repress enzyme expression and flux toward the desired products. Despite the advantages of dynamic control over static regulation, it can be challenging to design the optimal control topology and tune the control parameters.[151] Here we describe several static and dynamic regulation systems that are frequently applied in metabolic engineering for production of target molecules.

2.2.2.1. Static Regulation. Static manipulation of metabolic pathways refers to the use of genetically encoded components that are expressed constitutively, i.e. independent of changes in cellular and fermentation conditions. Static regulation can be achieved using promoter engineering, RBS engineering, gene copy number engineering, RNA interference, and CRISPR activation and repression.

Promoter Engineering. Promoters are the foundational and ubiquitous genetic component that drives gene expression. Promoter engineering is an efficient method to regulate and construct promoters with diverse strengths and functions and has been recognized as a useful tool to precisely regulate gene expression in synthetic biology and metabolic engineering.[152]

https://doi.org/10.1021/acs.chemrev.2c00403
Chem. Rev. 2023, 123, 5521−5570

Eukaryotic promoters typically contain two modular elements, a core promoter and an upstream activating sequence (UAS), while prokaryotic promoters consist of two short DNA sequences at −10 and −35 positions upstream from the transition start site. Existing promoter databases, such as Anderson promoter collection for *E. coli* and SCPD promoter database of *S. cerevisiae*,[153,154] have enabled selection of promoters with appropriate strength ranging from strong, medium, to weak. Nevertheless, endogenous promoters may not enable efficient optimization of metabolic flux or coexpression of multiple genes in metabolic networks because of poor dynamic range or lack of orthogonality to native regulations.[155] Thus, it can be beneficial to engineer synthetic promoters with better activity. One strategy for promoter engineering is hybrid promoter engineering, in which different promoter elements are combined to create novel promoters. In one study, a library of hybrid promoters with different strengths for gene expression in *Y. lipolytica* was constructed using combinations of promoter elements from *S. cerevisiae* and *Y. lipolytica*.[156] In another study, several UASs were inserted in front of the core *GAL1* promoter from *S. cerevisiae*, creating a library of synthetic galactose-inducible promoters with an expanded dynamic range, in which the best promoter had 2-fold higher activity than the wild-type *GAL1* promoter under a variety of different carbon sources.[157] Furthermore, the activity of promoter Pylb from *B. subtilis* was optimized by substituting −35, −10 core region and upstream sequence with consensus sequences, resulting in an engineered promoter that exhibited 195-fold increase in superfolded green fluorescent protein (GFP) expression compared to the wild-type promoter.[158] In addition to hybrid promoters, promoter engineering can be used to develop minimal promoters, which can reduce the DNA burden of the plasmid construction process. To obtain minimal promoters, Alper and co-workers used random oligonucleotides with different sizes to determine the minimal sizes of core promoter and UAS sequences, and the strongest minimal promoter could achieve 70% of the strength of the stronger constitutive *TDH3* promoter.[159]

In addition to the rational hybrid promoter approach, promoter engineering can also be achieved through random mutagenesis, which has been applied to the construction of promoters with different strengths, leading to the identification of critical regions in promoters.[160] In one study, error-prone PCR was used to mutate the bacteriophage PL-λ promoter.[161] The PL-λ promoter library was used to drive GFP expression and screened based on the fluorescence of *E. coli* colonies, and twenty-two promoter variants were found to span a 325-fold range of mRNA expression. In another study, to optimize a cellobiose consumption pathway in an industrial *S. cerevisiae* strain, random mutagenesis was performed on the *ENO2* and *PDC1* promoters, which were used to drive the expression of a cellobiose transporter and β-glucosidase, respectively.[162] For further reading, we recommend some reviews on promoter engineering.[155,163]

RBS Engineering. The ribosome binding site (RBS) is a genetic element present in prokaryotic promoters, and RBS engineering is another efficient approach to tune gene expression levels in prokaryotic systems. In one study, the Base Editor-Targeted and Template-free Expression Regulation (BETTER) method was developed to diversify multigene expression.[164] Endonuclease deficient Cas9 (dCas9) editor was used to target up to 10 tailored RBSs upstream of genes required for xylose utilization or lycopene production in *Corynebacterium glutamicum*[165] and synthetic RBSs with different translation rates designed with the RBS calculator were used to modulate GPP synthase expression to optimize limonene biosynthesis in cyanobacterium *Synechococcus elongatus*.[166] In another study, 82 operons in *E. coli* were designed and characterized to determine the sequence and structural factors that controlled mRNA stability; redesigning the ribosome binding site was found to lower the translation initiation rate of the first coding sequence by 35-fold and result in 11.8-fold decrease in the mRNA level.[167]

Gene Copy Numbers. In addition to the promoter and RBS, gene copy number is another genetic element that can be optimized to improve the production of the target product. Higher gene copy numbers can be obtained by using high-copy plasmids or integrating several copies of the gene into the chromosome. In one study, the flavonoid (2S)-naringenin biosynthetic pathway was integrated into repetitive rDNA sites in *S. cerevisiae*, creating multiple strains with different copy numbers of the pathway.[168] In another study, the genes of interest and an antibiotic resistance marker were placed between 1-kb homology regions, and RecA-mediated recombination of the homology regions led to daughter *E. coli* cells with higher tolerance to antibiotic concentrations or higher copy number of the integrated genes.[169] A strategy for chromosomal integration of genes with several copies was also developed by leveraging FLP/FRT site-specific recombination, enabling integration of 15 copies of a single gene or 18 total copies of three genes into the chromosome of *E. coli* in one step.[113] Another approach to increase plasmid copy number is to truncate promoters driving the expression of selection marker genes. Because truncated promoters have weakened strength, cells need to maintain higher copy numbers of the plasmid to survive in a selective medium. Zhao and co-workers truncated promoters driving the expression of antibiotics markers, KanMX and HygB, leading to a series of plasmids with stepwise increased copy numbers as high as 100 copies per cell.[170]

RNA Interference. Gene downregulation can be beneficial when the target gene is essential, and thus complete deletion is lethal. Gene repression in *S. cerevisiae* can be achieved using RNA interference (RNAi), which is a post-transcriptional, gene-silencing mechanism present in eukaryotic organisms.[171] RNAi employs an RNA-induced silencing complex to degrade mRNA transcripts. In particular, Dicer (Dcr) cleaves a double-stranded RNA to form smaller guide RNAs that target the gene of interest.[172] Argonaute (Ago) then uses the guide RNAs to degrade the corresponding mRNA, lowering gene expression levels. In one study, the libraries for expression cDNAs targeting the whole genome of *S. cerevisiae* were constructed and used to identify target genes beneficial in improving xylose consumption.[173] In another study, an RNAi library was also combined with high-throughput microfluidic single-cell screening to determine that the downregulation in genes related to cellular metabolism, protein modification, and cell cycle is related to increased amylase production in *S. cerevisiae*.[174] In addition to RNAi, gene knockdown can be achieved using small regulatory RNAs (sRNAs). Previously, a standard platform to develop sRNA for down regulation of biochemical pathways was developed.[175] Screening of 101 *E. coli* sRNAs led to the identification of micC as the optimal scaffold to engineer sRNAs due to its high repression capability. Using a library containing 130 synthetic sRNAs, gene targets that enabled increase in cadaverine production were identified. The sRNA

https://doi.org/10.1021/acs.chemrev.2c00403
Chem. Rev. 2023, 123, 5521−5570

expression platform was then expanded to enable rapid, multiplexed, and genome-scale knockdowns by utilizing compatible sets of origins of replication and antibiotic markers.[176]

CRISPR Interference and CRISPR Activation. In addition to RNAi, CRISPRi is another technique to downregulate gene expression levels. CRISPRi reduces gene expression by targeting a region near the transcription start site with a dCas9, creating a steric hindrance that prevents access of RNA polymerase to initiate transcription.[177] The dCas9 contains mutations in the HNH nuclease and RuvC-like domains and can still bind to the target region without introducing a DSB. The level of repression can be further improved by fusing the dCas9 with a repressor domain, such as the Kruppel-associated box domain.[178] A dCas12a-based CRISPRi system was constructed to repress multiple target genes in a single crRNA array in *S. elongatus* UTEX 2973 with efficiencies ranging from 53% to 94%.[179] In addition to using strong constitutive promoters to enhance gene expression levels, CRISPR activation (CRISPRa), in which a transcriptional activation domain is fused to dCas9, can be used to upregulate gene expression levels. CRISPRa machinery developed in *E. coli* was ported to *Pseudomonas putida*, and the criteria for efficient CRISPRa target sites in *P. putida*, such as distance to the transcription start site and modified sgRNA targeting sequences, were found to be similar to those in *E. coli*.[180] Furthermore, simultaneous activation and interference were achieved by expressing orthogonal Cas proteins, in which one protein was fused to an activation domain while the other protein was fused to a repressor domain, with their cognate sgRNAs targeting the selected genes.[181] In another study, sgRNA-expressing plasmid libraries from pools of array-synthesized oligos for overexpression and downregulation of all the genes in the genome of *S. cerevisiae* were constructed, enabling genome-scale activation and interference.[182] While CRISPRi and CRISPRa are straightforward to design and can be used to achieve multiplex gene regulation, they can show toxicity and off-target effects, and the large size of the CRISPR systems can increase the metabolic burden on cells.

2.2.2.2. Dynamic Regulation. In contrast to static regulation, dynamic regulation seeks to genetically modify an organism to achieve a balance between growth and production or to shift its metabolism from growth to production in response to an external environmental signal or internal metabolic state. Dynamic regulation can be classified into two different modes: manual induction and autonomous induction.[183] In manual induction, cells are first grown to a target density and then an exogenous signal, including chemical inducers, carbon sources, or environmental signals, is introduced to minimize the cell growth while activating the production pathways. On the other hand, autonomous induction relies on key intermediates or quorum signals to dynamically shift the metabolic flux toward the desired product without human intervention. We refer readers to some in-depth reviews covering dynamic control.[183,184]

Manual Induction Using Exogenous Signals. Chemical inducers are typically used in bacteria to drive the expression of biochemical pathways and can be classified into nonsugar inducers and carbon sources. Anhydrotetracycline (aTc) and isopropyl β-D-1-thiogalactopyranoside (IPTG) are common nonsugar inducers and have been used for the production of various chemicals in *E. coli*.[185−187] Inducible promoter systems using aTc and IPTG have also been developed for bacterial species with limited genetic tools, such as *C. acetobutylicum* and *B. subtilis*.[188,189] In addition to nonsugar inducers, carbon sources have been used to induce gene expression. In *E. coli*, arabinose is a commonly utilized sugar inducer, while inducible systems in *S. cerevisiae* mainly rely on galactose. The endogenous yeast galactose-inducible (GAL) expression system has been widely used to induce biochemical pathways by switching to galactose-containing media following growth phases in glucose.[190,191] Recently, the glucose-inducible system of *P. putida* was ported into an *E. coli* strain with an activated Entner−Doudoroff pathway, enabling tight control of gene expression over a wide range of glucose.[192]

Manual Induction Using Environmental Signals. While dynamic regulation using exogenous signals is effective, reversing the induction by exchanging the medium is laborious and impractical. On the other hand, environmental signals, such as temperature, oxygen, and light, are low cost, generally nontoxic, and can be applied and removed from the medium several times during growth and production phases. Temperature-inducible systems can be controlled externally and are easily timed. The λ phage major rightward (pR) and leftward (pL) promoters, which are induced at 42 °C and inactivated at 33 °C, were utilized to develop a heat-inducible system.[193,194] The system was used to regulate the expression level of D-lactate dehydrogenase and L-alanine dehydrogenase in *E. coli*.[195,196] Cells were grown aerobically at 33 °C followed by thermo-induction at 42 °C under oxygen-limited conditions. In another study, the heat-inducible system was used to downregulate the expression of the isocitrate dehydrogenase gene (*IDH*) in an *E. coli* strain.[197] In addition to temperature, the dissolved oxygen level is another environmental signal that can be induced easily and reversibly. The *nar* promoter in *E. coli* is induced under anaerobic conditions and was used to dynamically drive the expressions of genes for production of D-lactate, 2,3-BDO, and 1,3-propanediol (1,3-PDO).[198] Fermentations were initially performed under aerobic conditions at 250 rpm for optimal cell growth, followed by a reduction in shaking speed to 100 rpm for production under anaerobic conditions. Furthermore, dynamic control using light is highly tunable and reversible. Optogenetic switches are typically derived from the photoreceptors that undergo a conformational change in response to light.[199] Previously, a chromatic acclimation sensor/regulator (CcaSR) optogenetic system was developed to control the expression of glucose-6-phosphate isomerase, a glycolytic enzyme that dictates flux distribution between Embden−Meyerhof−Parnas (EMP) and pentose phosphate pathways.[200] Using such a system, the EMP flux could be controlled between 0.5% to 50% of the total glycolytic flux. In another study, by combining the blue light-activated *EL222* gene expression system with the GAL regulators and a photosensitive degron domain, two optogenetic gene expression systems (OptoEXP and OptoINVRT) were developed, enabling a switch from a light-induced growth phase to a darkness-induced production phase.[201]

Autonomous Induction Using Metabolite-Based Biosensors. Unlike several dynamic regulation systems that rely on the manual introduction of exogenous inducers or environmental signals at a predetermined time, autonomous dynamic control does not require human intervention and can be engineered by using metabolite-responsive biosensors. Liao and co-workers developed the first autonomous dynamic control system, in which an acetyl phosphate-responsive promoter was used to control the expression of phosphoe-

https://doi.org/10.1021/acs.chemrev.2c00403
Chem. Rev. 2023, 123, 5521−5570

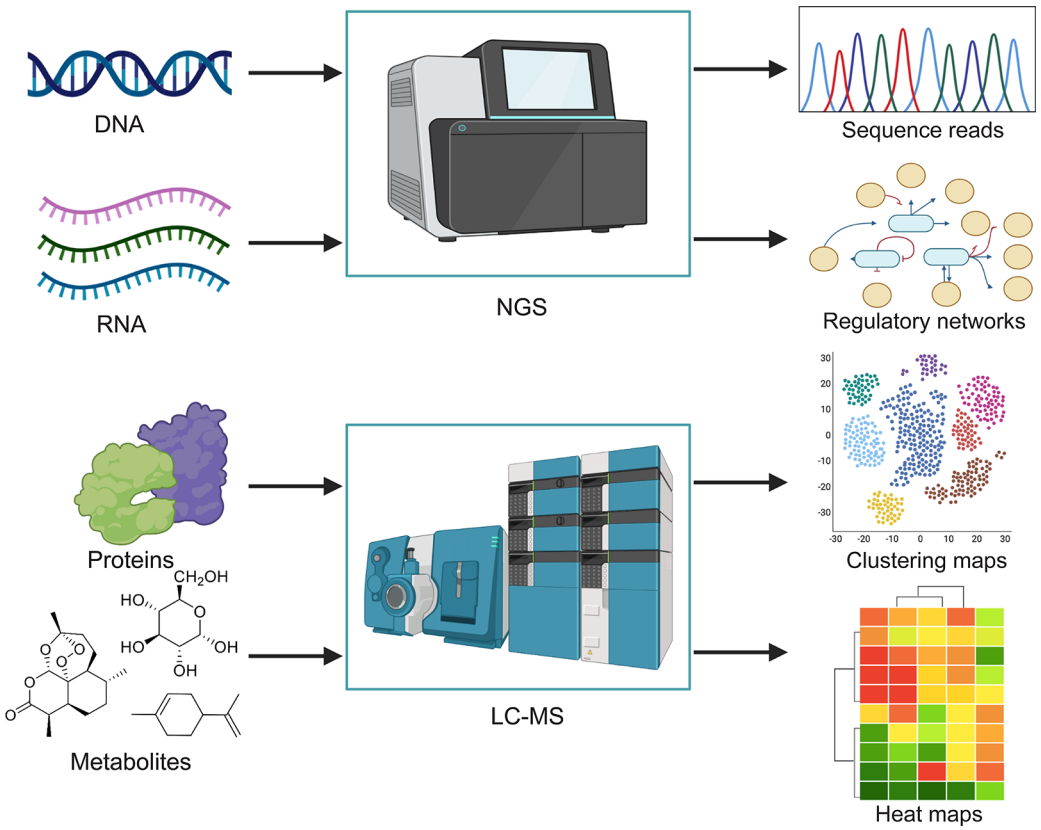


Figure 6. Omics tools for analysis in metabolic engineering. The most applied omics tools for characterization of host and mutant strains in metabolic engineering roughly fall under two major categories: next-generation sequencing (NGS) and mass spectrometry (MS)-based techniques. NGS is used for both DNA-Seq and RNA-Seq analysis, while MS-based methods, typically liquid chromatography-MS (LC-MS), are employed for proteomic and metabolomic analysis.

nolpyruvate synthase and isopentenyl diphosphate isomerase.[27] The system used acetyl phosphate to sense excess glycolytic flux and could redirect carbon flux from acetate production toward lycopene. In another study, an acyl-CoA-responsive promoter was used to control the expression of ethanol and fatty acid ethyl ester (FAEE) producing modules.[202] As acyl-CoA accumulated, the two modules were upregulated, enabling tight control over FAEE production. Furthermore, metabolite-responsive biosensors could be coupled with CRISPRi and RNAi to achieve autonomous gene downregulation. For example, muconic acid and glucosamine-6-phosphate served as the inducing molecules for dynamic repression of competing pathways in muconic acid and *N*-acetylglucosamine production, respectively.[203,204]

Autonomous Induction Using Quorum Sensing. In addition to dynamic control using metabolite-responsive biosensors, quorum sensing-based dynamic regulation is inducer-free and triggered when a certain cell density is achieved, which is an important parameter in metabolic engineering. Previously, Prather and co-workers designed a quorum sensing circuit to degrade the phosphofructokinase-1 gene (*pfk-1*).[205] The *pfk-1* was expressed under the P_{esaS} promoter, which was activated by transcriptional regulator EsaRI70 V in the absence of *N*-acyl homoserine lactone (AHL). As cells grew, accumulation of AHL reduced activity of EsaRI70 V and eventually turned off the expression of *pfk-1*. The same group further developed a layered dynamic regulation tool by combining the Lux and Esa quorum sensing systems, enabling independent, simultaneous, and dynamic activation and downregulation of two sets of genes.[206] Recently, a synthetic quorum sensing circuit was developed in *S. cerevisiae* by combining the hormone cytokinin system from *Arabidopsis thaliana* with the endogenous Ypd1-Skn7 signal transduction pathway.[205] The quorum sensing circuit was further coupled with an auxin-inducible protein degradation system to achieve autonomous degradation of the growth-related Erg9p.

2.3. Test

The advancements in measurement techniques, specifically -omics technologies, have led to improvements in the reliability and sensitivity of quantitative biology methods. Reduction in price and easier operation has resulted in a widespread adoption of many such 'omics-based tools. Microorganism characterization via tools such as genomics or metabolomics offers a deeper look into the microbial processes around reference and mutant strains, enabling improved design for metabolic engineering (Figure 6).

2.3.1. Genomics and Transcriptomics. With the advent of whole genome sequencing (WGS) using next-generation sequencing (NGS) techniques, it has become considerably easier to match genotypes to phenotypes in microorganisms. This has also benefited efforts in metabolic engineering to reverse engineer favorable phenotypes from microorganisms screened in natural environments.[207] By combining NGS and ALE, researchers now have the capability to easily piece together various mutations found in the genomes of high producing evolved mutants and trace the phenotypes to a set of causal gene mutations.[208] Databases such as ALEdb have

https://doi.org/10.1021/acs.chemrev.2c00403
Chem. Rev. 2023, 123, 5521−5570

further streamlined the process of storing sequenced mutations and recovering them later for hypothesis testing.[209] Apart from whole genome resequencing after ALE experiments, NGS has also been used in metabolic engineering for quantitative trait loci (QTL) mapping to identify loci that grant varying hydrolysate tolerances to different strains of *S. cerevisiae*. The lab strain was then reengineered to contain a subset of the favorable allelic loci to demonstrate a significant increase in robustness.[210]

NGS-based transcriptomics techniques such as RNA-Seq have also been employed to study the transcriptional profiles of high performing mutants evolved from ALE experiments to identify the impact of mutations on the regulatory architecture. More specifically, in applications tailored toward metabolic engineering, combining regulatory data with metabolic models results in improved accuracy of FBA simulations. This was demonstrated by the IDREAM framework by integrating metabolic models of *S. cerevisiae* with the inferred regulatory networks from transcriptomic data.[211] Transcriptomics data sets, on the whole, are relatively underused for metabolic engineering applications compared to other omics data sets.[212]

2.3.2. Proteomics and Metabolomics. Along with the rise of NGS-based genomics and transcriptomics, improvements in MS techniques have also contributed to a rise in the use of proteomics and metabolomics analysis within metabolic engineering research. An often-encountered limitation of transcriptomics is the lack of correlation between mRNA transcript levels and enzyme activities, leading to an incorrect understanding of the genotype−phenotype relationship. One way to overcome this limitation is to directly measure the protein copy numbers in a cell population using proteomic techniques which can serve as a more accurate proxy for enzyme activity levels. Proteomics-guided metabolic engineering, however, is still a nascent field with scarce literature. Recently, a protocol for proteomic analysis of *E. coli* tailored specifically toward downstream metabolic engineering applications was reported.[213] Alonso-Gutierrez and co-workers developed a statistical analytical workflow for proteomics data sets, which performed principal components analysis (PCA) on such data.[214] Known as PCA of proteomics (PCAP), the method applied PCA to proteomics and target molecule levels to highlight specific enzymes that require modulation to guide metabolic engineering efforts. Analysis via this method was used to overexpress the downstream enzymes of the mevalonate pathway in *E. coli* and maintain a balanced expression of the remaining enzymes, resulting in a 40% increase in the production of the terpenes, limonene, and bisabolene. Within metabolic engineering applications, proteomic data sets have also been used to engineer higher tolerance to inhibitors. Proteomic analysis of the stress response of an *E. coli* strain grown in different inhibitor conditions identified a list of differentially expressed proteins that showed a clear tolerance response to stress. Overexpressing the most differentially upregulated protein from this list, YcfR, resulted in an engineered *E. coli* that showed higher growth fitness when exposed to stressors such as acetic acid, furfural, and phenol.[215]

Although proteomic analysis generates a useful snapshot of cellular resource allocation, it does not provide the full picture of carbon distribution within metabolism. Thus, a better understanding of accumulations and bottlenecks is often obtained by performing metabolomic analysis of host microorganisms instead. The implementation of mass spectrometric analyses of microbial strains for metabolomic analysis of metabolite species has been extensively studied and reported in literature.[216,217] Gold and co-workers applied a protocol for targeted metabolomics to overproduce L-tyrosine in *S. cerevisiae*.[218] The study utilized targeted metabolomics of certain metabolites to assess the efficacy of different strain engineering strategies, including expression of tyrosine feedback resistant genes and global engineering of central carbon metabolism. Metabolomic data also has the potential for integration into FBA frameworks via methods like MetDFBA, which allow for more accurate flux profile predictions, enabling better designed mutant strains in the future.[219] Published data sets of metabolomic analysis of mutant libraries, such as transcriptional factor knockouts, can expand the accessibility of systems-level metabolic engineering efforts to the larger scientific community.[220]

2.3.3. Fluxomics and ^{13}C Metabolic Flux Analysis. Fluxomics refers to the variety of techniques used to estimate metabolic fluxes. Reaction rate fluxes are a better representation of the dynamic metabolic trends compared to metabolome measurements, which only provides a static snapshot of metabolism. Thus, flux measurements are considered more informative for strain engineering efforts.[221]

However, measurement of fluxes *in vivo* is not straightforward because flux is not an observed but a calculated phenotype. Calculation of every reaction flux would require time series measurements of every participating metabolite, which is not possible. Thus, computational methods are employed where stoichiometric models of metabolism are solved at metabolic steady state to obtain reaction flux values. One such method, known as flux balance analysis (FBA), employs metabolic stoichiometric models to predict flux values by solving an optimization problem involving a steady-state constraint.[222] To constrain the space of possible flux solutions, experimentally measured rates of extracellular transport are employed as constraints. To further constrain the solution space, thermodynamic constraints are also enforced to ensure thermodynamic feasibility of the flux solutions.[223,224]

The inclusion of isotopic tracers, mathematical models, and high-resolution mass spectrometric analysis has enabled the inclusion of better constraints for flux calculations, thus increasing the accuracy of flux predictions. ^{13}C metabolic flux analysis (^{13}C-MFA) is used to trace the flow of carbon through metabolic pathways by introducing ^{13}C-labeled substrates into a microorganism's growth medium and then by analyzing their incorporation into intracellular metabolite pools via MS.[225] In recent years, rigorous work has been done to develop reliable methods for isotopic flux analysis, including work in experimental design, MS analysis, as well as computational techniques to calculate flux values.[226,227]

Isotopic flux analysis has been employed to identify bottlenecks that can help lead to overproduction of target molecules.[228] ^{13}C labeling data is also shown to be useful in constraining the prediction space of GEMs, serving as an additional data set for integration in mathematical models.[229] Various review articles in recent years have addressed the benefits and limitations of employing isotopic flux analysis in metabolic engineering.[221,225,230]

2.3.4. *In Vivo* Biosensors. Evolutionary or combinatorial approaches have been increasingly used for metabolic engineering. However, despite the availability of various MS techniques, screening and evaluating the phenotypes of modified microorganisms remain a challenge. Besides, real-time monitoring of metabolite concentrations is critical for the

https://doi.org/10.1021/acs.chemrev.2c00403
Chem. Rev. 2023, 123, 5521−5570

dynamic regulation of metabolic networks, which are rarely captured by techniques mentioned previously. Alternatively, biosensors have emerged as an enabling high-throughput technology for faster screening, selection, and study of metabolically engineered strains.[231−234] Biosensors are typically genetically encoded components, either proteins or nucleic acids, but they can also be derived from the naturally occurring biosensors that reflect changes in the intracellular metabolite concentration by transducing the metabolite interaction into a detectable output signal like gene expression.[235,236] These biosensors are often linked with other techniques like flow cytometry, MS, or colorimetric assay to measure intracellular metabolites quickly.[237] Biosensor-coupled screening of metabolic regulation has been explored in the prokaryotic and eukaryotic organisms, including plants.[238,239]

Transcription factor (TF)-based biosensors are the most adeptly used protein-based biosensors in metabolic engineering. TFs often possess ligand-binding domains that bind and sense small molecules such as amino acids, lipids, intracellular metabolites, sugars, and other metabolites or environmental stressors. TFs responsive to these natural metabolites have been used to engineer biosensors for metabolic engineering applications.[250] A variety of metabolite-responsive biosensors have been successfully used in designing synthetic gene circuits capable of detecting different amino acids, acids, secondary metabolites, fatty acids, fatty acyl-CoA, malonyl-CoA, and alkanes.[234,240−243] Recently, Borodina and co-workers employed a biosensor for *cis*,*cis*-muconic acid (CCM) coupled with GFP expression for high throughput screening of yeast UV-mutagenic libraries, where high GFP expressing cells, corresponding to high CCM concentrations, were sorted using FACS helping identify a mutant capable of producing 49.7% higher CCM.[237] A challenge associated with natural biosensors is that they are specific to the cognate ligands and have lower responses to effector molecules. Liu and co-workers utilized the continuous evolution system, OrthoRep, to evolve the CCM biosensor, BenM, to achieve a higher dynamic range and higher activity toward desired cognate ligand (CCM) and noncognate ligand (adipic acid). Evolved biosensors outperform the parent BenM-based biosensors by achieving a 180-fold higher dynamic range and a broadened operational range.[244,245]

Another type of protein-based biosensor are Förster resonance energy transfer (FRET)-based biosensors, which have been developed to sense sugars, amino acids, pyruvate, lactate, redox conditions, and environmental stressors.[231] The design of FRET-based biosensors involves metabolite binding protein (MBP) fused in the middle of two fluorescent proteins, including a donor and an acceptor. The binding of a ligand molecule to MBP creates a conformational change and modulates energy transfer from donor to acceptor protein, producing a detectable output. Ahmad and co-workers developed FRET-based biosensors for real-time monitoring of intracellular lysine concentration in bacterial and yeast systems. The lysine binding periplasmic protein (LAO) from the *Salmonella enterica* LT2 strain was linked with the cyan fluorescent protein (CFP) and yellow fluorescent protein (YFP) to develop the FRET-based biosensor system.[246] Further, the dynamic range of lysine detection was expanded by performing mutagenesis in the region critical for lysine binding. Zamboni and co-workers developed a trehalose-6-phosphate (T6P) FRET system to visualize the intracellular T6P accumulation under *in vivo* microscopy in different osmotic stress conditions.[247] Recently, Pohl and co-workers developed a FRET-based glucose biosensor applied to glucose monitoring in microbioreactor cultivations of *E. coli* and *C. glutamicum*.[248] Overall, monitoring metabolites based on FRET biosensors shows a wide variety of applications in metabolic engineering, biomedicine,[249] pharmacology, toxicology, and food sciences.[250] However, challenges still lie in identifying specific MBPs to develop new FRET-based biosensors.

Lastly, riboswitches are transcription-based RNA biosensors that are usually built upon aptamer domains to either facilitate or disrupt the formation of a terminator, which creates transcriptional repression or activation by preventing the synthesis of long mRNAs. The main advantage of riboswitches is their improved response time compared to protein-based biosensors.[234,251,252] Such RNA biosensors are usually only specific to limited metabolites, such as folinic acid and theophylline, due to limited availability of aptamers.[241] Weinberg, Hartig, and co-workers recently identified a new class of guanidine riboswitch called the guanidine-IV, where motifs associated with guanidine exporter associated sugE were identified via a bioinformatics approach. The efficacy of the new riboswitch was confirmed by addition of 5 mM guanidine, which resulted in an 80-fold increase in GFP response in *Staphylococcus aureus*.[253] These findings have opened the door to the identification of new metabolite-binding RNAs, which are needed for designing improved regulatory circuits in metabolic engineering.

2.4. Learn

Learn is primarily composed of computational tools applied to metabolic hosts. They have been indispensable for their ability to help identify meaningful biological patterns and discover engineering targets. The goal of these computational tools is to design strains from first-principles, gain further insight on metabolism, help influence knowledge-guided design, unify multiple data modalities, and achieve TRYs that might not be achievable through evolutionary engineering. There are an abundant number of different approaches to modeling in metabolic engineering, which can be grouped into three categories: (1) genome-scale models (GEMs), (2) kinetic models, and (3) ML models. We discuss the core principles of these methods and reference a few key examples.

2.4.1. Genome-scale Models. GEMs are network representations of metabolic pathways that include information about the topology and stoichiometry of the reactants and the products involved in each reaction.[254] The stoichiometric coefficients are represented compactly in the form of a single stoichiometric matrix. GEMs can be solved at the metabolic steady state to obtain the flux vector consisting of reaction rates at steady state. As the linear equations being solved are underdetermined, additional constraints are often imposed to reduce the feasible solution space. Constraints are often derived from prior biological knowledge such as lower and upper bounds on flux values. Extracellular measurements of metabolites can also be used to obtain rates of cellular uptake or secretion, and they can be included as additional constraints. While the stoichiometric, inequality, and measurement constraints narrow the solution space considerably, the linear equations that remain undetermined allow for multiple flux vector solutions. Thus, an optimization problem is solved for a phenotypic objective, most often growth, and the flux vector

https://doi.org/10.1021/acs.chemrev.2c00403
Chem. Rev. 2023, 123, 5521−5570

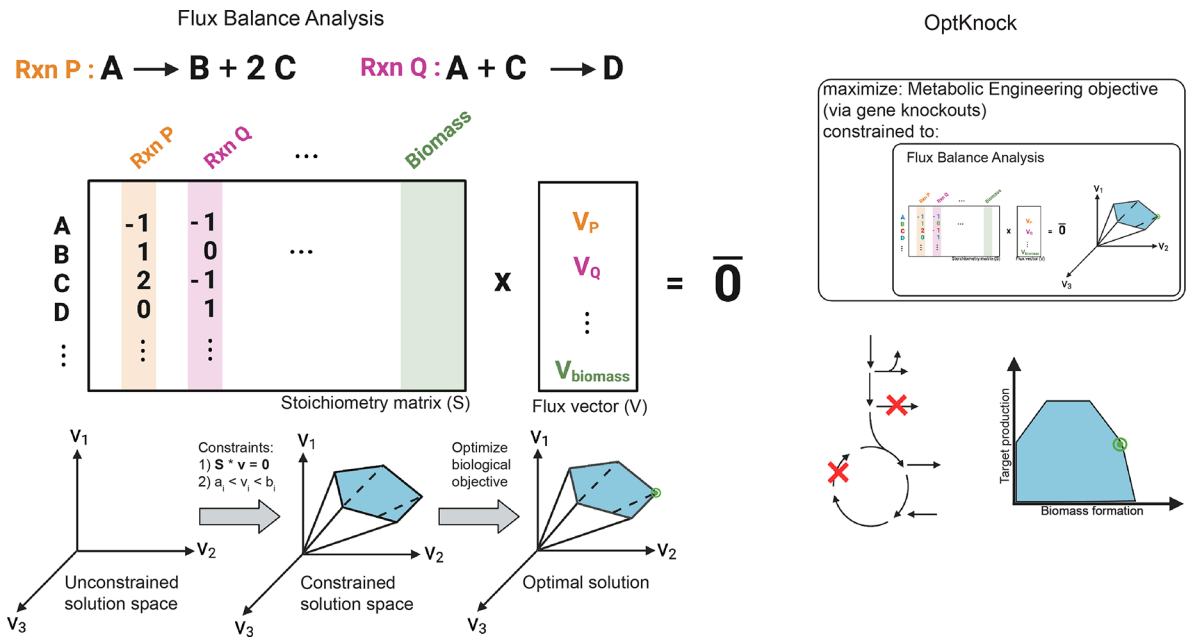


Figure 7. Genome-scale models (GEMs) in metabolic engineering. GEMs are mostly used with flux balance analysis (FBA) to generate flux profile predictions at metabolic steady state. To construct a GEM, extensive knowledge of the metabolic reactions at the genome-scale is required, which are transcribed as a stoichiometric matrix (S). The linear algebra equation is solved at steady state and used to constrain the solution space for an acceptable flux vector. OptKnock is a commonly used algorithm that optimizes an objective function (related to a metabolic engineering objective) and solves an FBA problem underneath to obtain flux vectors corresponding to the engineering objective.

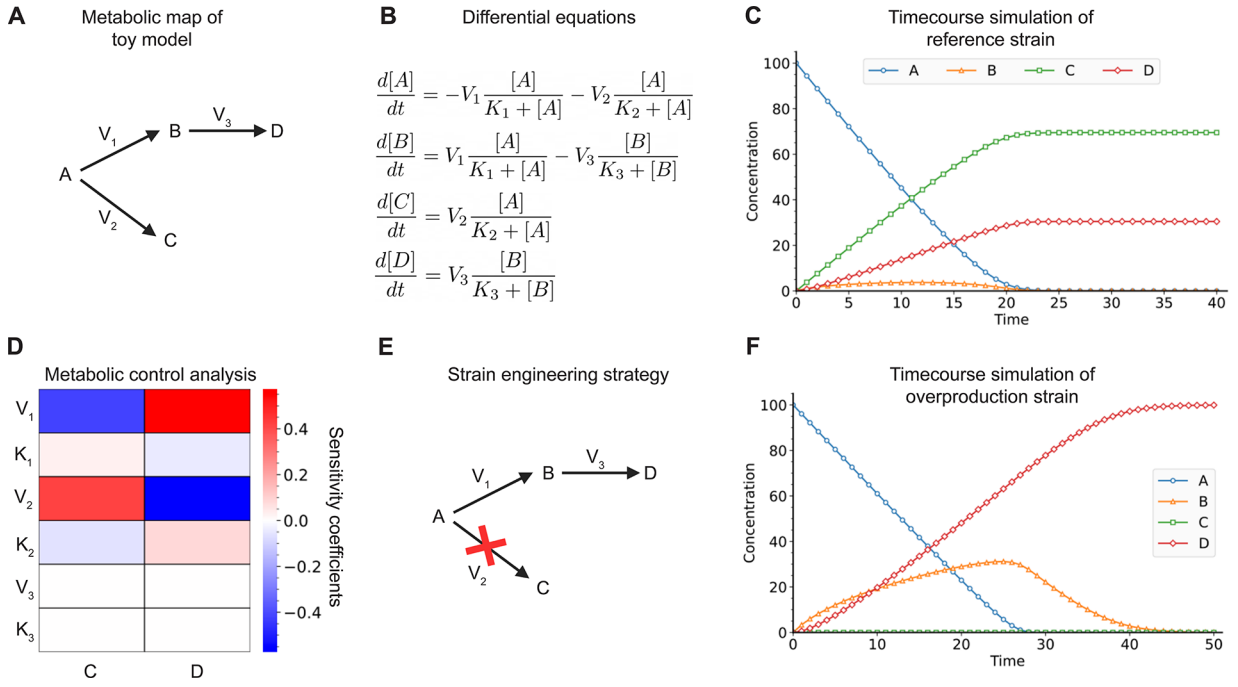


Figure 8. Kinetic models in metabolic engineering. (A) The construction of kinetic models of a metabolic system requires knowledge of the reaction pathways and network topology. Here, D is the target chemical and C is a byproduct. (B) Ordinary differential equations (ODEs) are written to model the rate of change of each metabolite within the model, with Michaelis−Menten rate expressions commonly used. (C) Integration of the ODEs is used to predict the dynamical trajectories and steady-state concentrations of the metabolites. (D) Metabolic control analysis (MCA) is performed to shed light on the sensitivity of each parameter within the model on the concentrations of metabolites. Here, the rows represent model parameters and columns represent the metabolites C and D. Higher positive values imply that increasing the parameter increases the corresponding metabolite concentration, and higher negative values imply vice versa. (E) From the MCA heatmap, it can be inferred that down-regulating or knocking out reaction V_2 can increase the concentrations of the target metabolite D. Thus, the model is modified to knockout reaction V_2. (F) In silico prediction of the engineering strategy shows that such an intervention can help improve concentration of D, thereby helping in the design of an overproducing mutant strain.

https://doi.org/10.1021/acs.chemrev.2c00403
Chem. Rev. 2023, 123, 5521−5570

satisfying these criteria is obtained. This methodology is known as flux balance analysis (FBA).[222]

GEMs and their analysis using FBA are one of the most common mathematical frameworks to model metabolic phenotypes for the purpose of metabolic engineering. Iterative updates to GEMs and the publication of consensus models have led to well-characterized GEMs for numerous model organisms such as *E. coli*, *S. cerevisiae*, cyanobacteria, and Chinese hamster ovary (CHO) cells. A recently published consensus-based GEM of *S. cerevisiae*, named Yeast8, offers the incorporation of enzyme constraints to obtain accurate flux solutions in a better-constrained flux space.[255] This feature was employed to improve the phenotype of protein secretion experimentally in *S. cerevisiae*.[256]

While GEMs offer insight into the metabolic behavior around a desired phenotype, they do not directly suggest genetic interventions required to achieve such a phenotype. Several algorithms have been implemented that impose an optimization condition along with the FBA framework to achieve a metabolic engineering target. The algorithm known as OptKnock[22] solves optimization problems at two levels: at the outer level, it identifies gene knockouts that couple cellular growth to the production of a target molecule, and at the inner level, it optimizes a cellular objective for each of the tested perturbations from the outer level, as described in the FBA method above. The algorithm was later evolved to include kinetic constraints, known as k-OptForce (Figure 7).[257]

2.4.2. Kinetic Models. Kinetic models are another type of mathematical framework that allows the prediction of metabolic states in time. The models use reaction rate expressions, such as mass action, Michaelis−Menten, or Hill kinetics to model the rate of reactions of processes such as synthesis, degradation, enzymatic conversion, or metabolite transport. The rate expressions can also accommodate regulation in the form of allosteric or transcriptional effects. These expressions are then linked to each other in the form of ordinary differential equations, where each equation is expressed as the mass balance of a given metabolite within the model. The complete set of differential equations containing the rates of formation and consumption of each metabolite within the model constitutes a kinetic model of metabolism. The capability of kinetic models to predict metabolite levels and reaction rates (metabolic fluxes) over time within a modeled biosystem makes them a valuable tool for designing metabolic engineering strategies aimed at the overproduction of target molecules.

Kinetic models are commonly utilized in conjunction with metabolic control analysis (MCA),[258] a method of analysis that computes the sensitivity of any model output with respect to a model parameter. Several commonly studied forms of sensitivity are available in strain design for metabolic engineering. The first is known as flux control coefficient, which indicates the expected change in a metabolic flux to variations in a parameter, typically enzyme activity or concentration. A second commonly studied metric is the concentration control coefficient, which indicates the expected change in the concentration of a metabolite to variations in enzyme activity. Therefore, MCA can be employed to guide insightful decisions while constructing microbial strains by utilizing the above control coefficients to predict bottlenecks and suggest solutions (Figure 8).[259]

Kinetic models that emulate large-scale biosystems with numerous biological components often contain many parameters. The estimation of these parameters relies on the availability of high-quality biological data, most commonly metabolite and flux measurements, which are often accompanied by proteomic and transcriptomic measurements. Such measurements are challenged by experimental throughput, sensitivity, and signal-to-noise ratio. Thus, the use of kinetic models to predict phenotypic behavior of large-scale biosystems is restricted by uncertainty in parameter estimation, which often leads to phenotypic predictions by kinetic models being accompanied by large confidence intervals. To address this limitation, several studies have turned toward ensemble modeling of metabolism.[260] Ensemble modeling employs the same underlying structure as kinetic models described above. However, instead of constructing a single kinetic model with a single set of estimated parameters, ensemble modeling constructs a population of models, each with their own set of parameters. The values for these parameters are estimated via sampling schemes such that the models satisfy constraints of thermodynamics, like reaction reversibilities. Employing an ensemble of metabolic models to predict phenotypic behavior allows for the prediction of a distribution of values instead of a single value, thus offering a precise confidence interval of such a prediction. The ensemble modeling framework was employed in combination with a genetic algorithm to construct a genome-scale kinetic model of *E. coli* metabolism.[261] In short, an ensemble of models was first constructed on the basis of steady-state flux data of the reference strain. Then a genetic algorithm was used to cross different parameter sets from the ensemble while minimizing the mismatch between experimental data and model predictions to identify the best combination of parameters that resulted in the lowest mismatch. This final parameter set resulted in a kinetic model that could accurately predict the product yields of 24 metabolites with a Spearman correlation coefficient of 0.84.

Another framework within kinetic models that employ MCA to study metabolic networks is the Optimization and Risk Analysis of Complex Living Entities (ORACLE).[262] Within this framework, flux and/or concentration values are uniformly sampled, and samples are rejected or accepted based on thermodynamic constraints. The ORACLE method was used to identify metabolic engineering targets for the production of 1,4-butanediol in *E. coli*.[263]

2.4.3. Machine Learning. Broadly, ML is a set of methods that uses algorithms that learn from experience without explicit instructions. In direct contrast to mechanistic models including GEMs and kinetic models, ML models are not often readily interpretable for biology. Lack of intrepretibility affords ML models to have a great deal of flexibility in combining multiomics data and expressivity in modeling complex nonlinear relationships, but it positions them one step further from causal explanations. Currently, ML is used at every stage of the systems metabolic engineering pipeline from feedstock to target product recovery.[111,264] The majority of the studies published related to both metabolic engineering, and ML can be roughly categorized into either design or characterization. ML for design aims to learn from data to suggest improved future designs, while ML for characterization aims to broaden knowledge over a specific domain for improved rational engineering or support of computational design pipelines. The most common classes of ML algorithms applied to metabolic engineering fall into supervised ML, where a model is trained to predict on a labeled data set and unsupervised ML where a model helps to identify patterns in an unlabeled data set. ML

https://doi.org/10.1021/acs.chemrev.2c00403
Chem. Rev. 2023, 123, 5521−5570

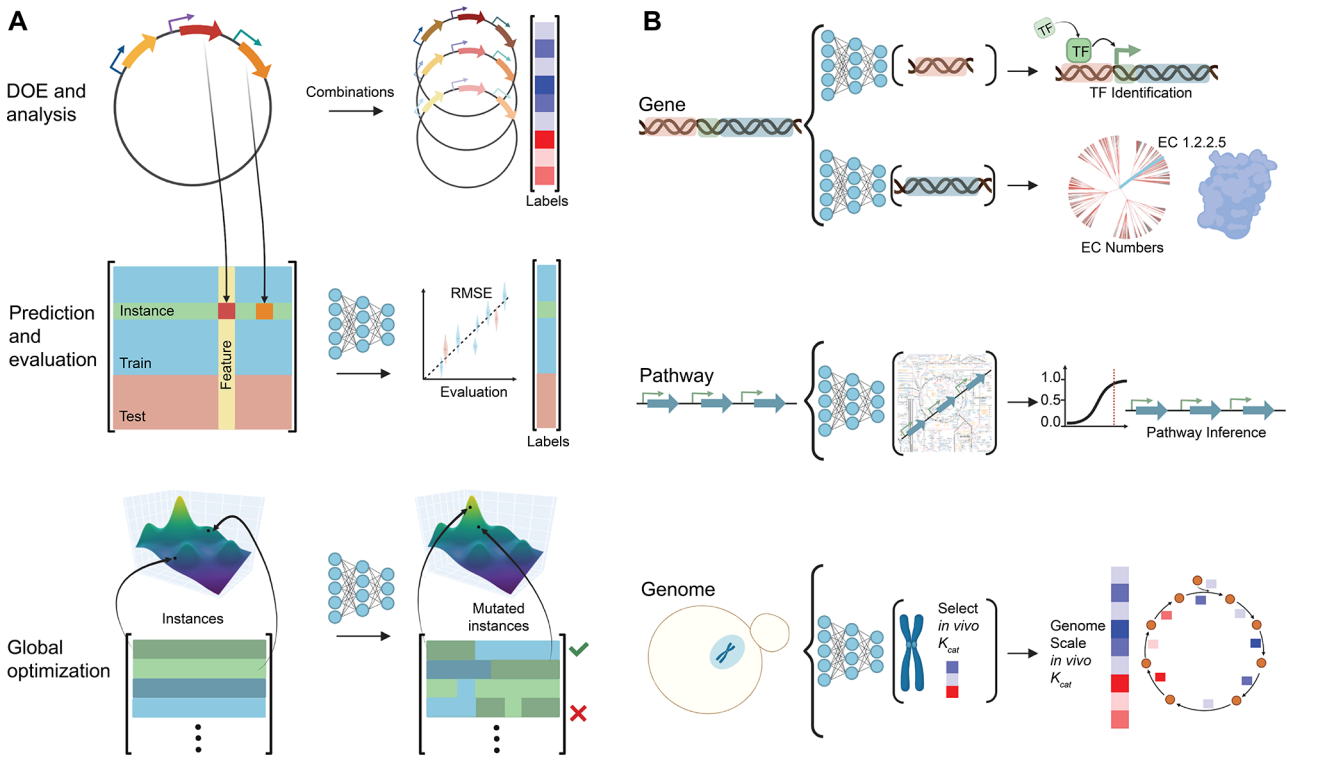


Figure 9. ML in metabolic engineering. (A) Standard workflow that integrates ML with metabolic engineering. First, a library of variants is constructed and analyzed by assigning labels to each variant. In this example, we can assume labels are titers associated with a pathway on one of the plasmid variants. Then, the data about the pathway is converted to a data matrix, where a ML model is trained to predict on the reserved test data. New variants that are predicted to perform well are then recommended for future design by a global optimization method. (B) ML for characterization. ML has been used to make inferences at different scales relevant to metabolic engineering. Example tasks include transcription factor binding and Enzyme Commission (EC) number prediction at the sequence level, pathway inference at the pathway level, and GEM parametrization at the genome scale.

consists of a vast set of methods, so to limit scope, we will focus on design and characterization that helps engineer organisms to overproduce a target molecule.

ML for design consists of two main components, the first is to learn a model that can predict the outcome from varied designs and the second is to choose the next set of designs to synthesize for improved production. To start the iterative learning process, design of experiments (DOE) is used for careful selection from the initial design space.[265] To predict the outcome of new designs, a model learns from training instances of different designs with their corresponding features and labels and attempts to predict the label for hidden test instances from test features. Common predictive models in metabolic engineering include Gaussian process,[266] linear models,[267] support vectors machine,[268] random forests,[269] neural networks,[270] and ensemble models that combine many of the previously mentioned individual models.[271] Following predictive modeling, a new design combination is suggested by a global optimization scheme like Bayesian optimization[272] or genetic algorithm (Figure 9A).[22]

ML studies related to characterization in metabolic engineering are more varied in approach, and they typically consider a much larger data domain such as the space of all genetic components, all proteins, or all pathways. Their aim is to further define the components of the cell, which in turn allows for more plausible engineering targets. Each level of the organism including gene, pathway, and genome has been the subject of ML studies that advance the metabolic engineering project. At the gene level, models have been constructed to predict promoter strength,[273] protein function,[274] promiscuity,[275] and transcription factor binding domains.[276] At the pathway level, there are models that allow for native pathway inference[277] and prediction of pathway dynamics.[37] At the genome scale, there are models that can help parametrize GEMs,[278] predict cellular phenotypes,[279] predict interaction networks,[280] and characterize metabolic networks.[281] Because these models typically operate on a larger data domain under more general conditions, they learn more generic patterns that are less useful for context-specific engineering design (Figure 9B). Well-curated databases have allowed for the application of ML algorithms for characterization, but there is still a need for more metabolic engineering-specific data sets. Another route to circumvent this is through literature mining, which can be used to predict phenotypes over a range of different environmental and genetic perturbations for different organisms.[282,283]

3. APPLICATIONS

In this section, we focus on case studies that implement the DBTL cycle to convert organisms into microbial cell factories for synthesis of chemicals, biofuels, utilization of different substrates, and improvement of strain robustness. In each section, compounds that were produced via comprehensive strategies were selected to discuss how metabolic engineering advances microbial production. Figure 10A shows 11 icons that each represents a particular strategy for metabolic engineering of a strain.

https://doi.org/10.1021/acs.chemrev.2c00403
Chem. Rev. 2023, 123, 5521−5570

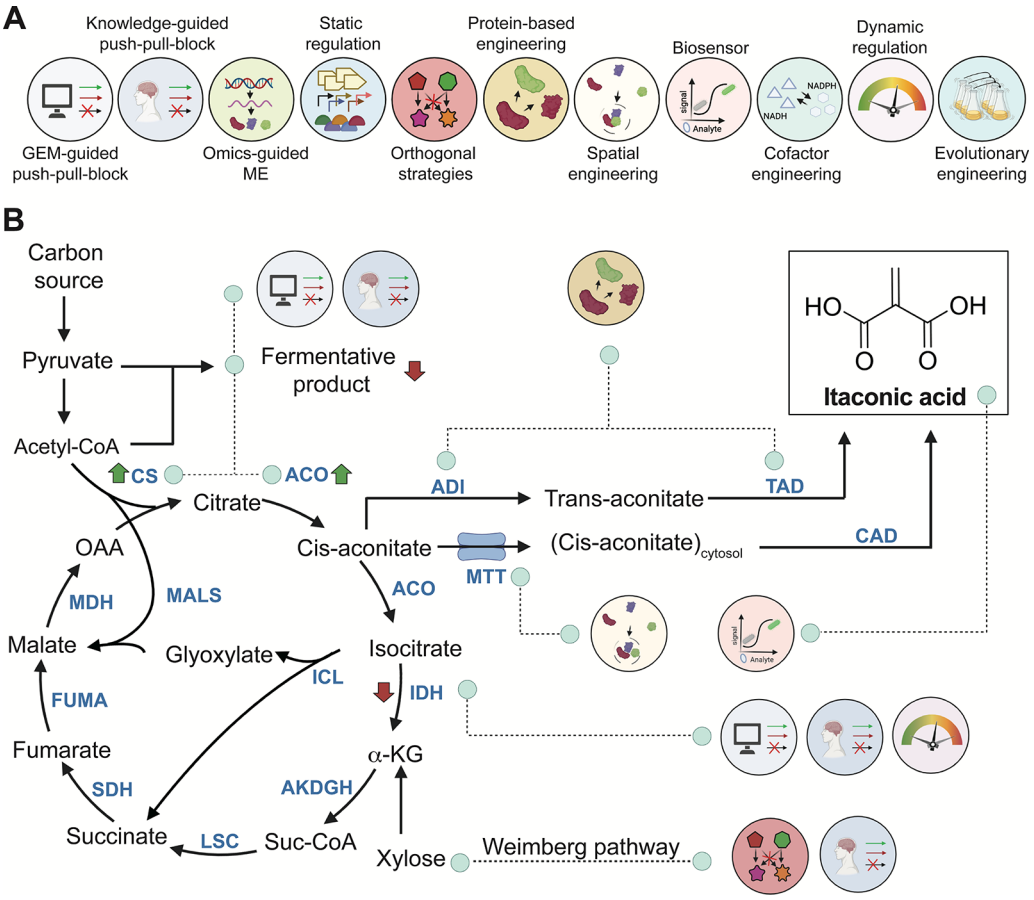


Figure 10. (A) Icons of metabolic engineering methods used to annotate case studies. (B) Pathway and metabolic engineering strategies for production of itaconic acid. Enzymes: CS, citrate synthase; ACO, aconitase; ADI, aconitate isomerase; TAD; *trans*-aconitate decarboxylase; CAD, *cis*-aconitate decarboxylase; IDH, isocitrate dehydrogenase; AKGDH, α-ketoglutarate dehydrogenase; LSC, succinyl-CoA ligase; SDH, succinate dehydrogenase; FUMA, fumarase; MDH, malate dehydrogenase; MALS, malate synthase. Compounds: Suc-CoA, succinyl-CoA; α-KG, α-ketoglutarate; OAA, oxaloacetic acid.

3.1. Synthesis of Chemicals

In the past few decades, most metabolic engineering studies have been focused on the production of chemicals. Below we will discuss the design and engineering of microbial cell factories for synthesis of bulk chemicals and fine chemicals.

3.1.1. Bulk Chemicals. Bulk chemicals are commodity chemicals used at a considerably large scale. Thus, their production via metabolically engineered organisms has gained much attention, and extensive effort has been taken in developing the cell-based production of bulk chemicals, such as alcohols, amines, and organic acids. Table 1 lists selected examples of important bulk chemicals with the currently highest reported titers and the metabolic engineering strategies associated with engineering production strains. Here, itaconic acid, 1,4-BDO, and 2-propanol/acetone are selected as case studies to discuss how metabolic engineering strategies have been used to enhance the production of bulk chemicals.

3.1.1.1. Itaconic Acid. Itaconic acid is a C5 organic acid and has been recognized as one of the top 12 building block chemicals by the U.S. Department of Energy.[284] Itaconic acid could be applied in the polymer industry as a comonomer as well as in agricultural or medical fields as a bioactive compound. Microbial production of itaconic acid is catalyzed by *cis*-aconitate decarboxylase (CAD), which converts *cis*-aconitate, one of the intermediates in the TCA cycle, into itaconic acid (Figure 10B). Klamt and co-workers constructed a GEM for optimizing the production of itaconic acid in *E. coli*.[285] In the final strain, both the acetate formation and pyruvate formation pathways were knocked out to enhance the carbon flux toward the TCA cycle. The push strategy was used to overexpress citrate synthetase (CS), *cis*-aconitase (ACO), and CAD. To increase the availability of *cis*-aconitate, the isocitrate dehydrogenase (IDH) gene, which is an essential gene, was knocked down by replacing its native promoter with a weak one. This GEM strategy enabled *E. coli* to produce 32 g/L itaconic acid. Furthermore, because *cis*-aconitate is unstable and toxic to the cell, the intracellular concentration of *cis*-aconitate is quite low. Thus, Guss and co-workers introduced a noncanonical pathway from *Ustilago maydis* into *P. putida*[286,287] to convert *cis*-aconitate to itaconic acid by aconitate isomerase (ADI) and *trans*-aconitate decarboxylase (TAD) because the *trans*-aconitate is more stable and nontoxic to the cell. The yield increased from 0.23 g/g to 0.34 g/g when changing the cis-pathway to the trans-pathway. The precursor for itaconic acid is in the mitochondria in eukaryotic cells, however, the expression of CAD is in the cytosol which creates an inefficiency in conversion. Liu and co-workers demonstrated that the overexpression of a mitochondria transporter (MTT) could enhance the production of itaconic acid by 10.5-fold in *Y. lipolytica*.[288] Moreover, because the quantification of itaconic acid is time-consuming, Malys and co-workers developed a whole-cell biosensor by identifying the novel

https://doi.org/10.1021/acs.chemrev.2c00403
Chem. Rev. 2023, 123, 5521−5570

Table 1. List of Several Bulk Chemicals Produced through Metabolic Engineering

chemicals	host	titer	metabolic engineering strategies	ref
		Alcohol		
ethylene glycol (EG)	*E. coli* W3110	108 g/L, 0.36 g/g-xylose, 2.25 g/L/h	• screen the best *E. coli* strain • push−pull • GEM • small RNA library	300
1,3-propanediol	*E. coli*	135 g/L, 0.51 g/g-glucose, 3.5 g/L/h	• mine new enzymes • cofactor engineering	301,302
1,2-propanediol	*E. coli* MG1655	17.3 g/L, 0.18 g/g-glucose, 0.72 g/L/h	• novel pathway construction • block the pathway of fermentative product	303,304
1,4-butanediol	*E. coli*	125 g/L, 0.4 g/g-glucose, 3.5 g/L/h	• GEM • cofactor engineering • energy balancing for cell growth • medium optimization	263,295
1,3-butanediol	*E. coli*	2.4 g/L, 56 mg/g-glucose	• block the competing pathway and fermentative product • screen the best enzyme • RBS engineering • gene copy number tuning	305
2,3-butanediol	*S. cerevisiae*	178 g/L, 2.64 g/L/h	• block the competing pathway • redox balancing	306
1,5-pentanediol	*E. coli* BW25113	0.97 g/L	• rewire the amino acid metabolism as a synthetic pathway • push−pull • protein engineering • transcription factor engineering	307
		Amine		
12 primary amine	*E. coli*	• detectable 12 primary amine • 10.67 g/L iso-butylamine from glucose	• retrobiosynthesis design by rewiring the amino acid metabolism	112
1,3-diaminopropane	*E. coli*	13 g/L, 0.1 g/g-glucose, 0.19 g/L/h	• GEM • push−pull • small RNA library	308
putrescine	*E. coli* W3110	42.3 g/L, 0.26 g/g-glucose, 1.26 g/L/h	• GEM • small RNA library	309
1,5-diaminopentane	*C. glutamicum*	103.8 g/L, 0.31 g/g-glucose, 1.47 g/L/h	• screen the best *C. glutamicum* strains • promoter tuning • transporter engineering	310
		Organic Acid		
3-hydroxypropionic acid (3-HP)	*Halomonas bluephagenesis*	154 g/L, 0.93 g/g-1,3-propanediol, 2.4 g/L/h	• transcriptome analysis • screening the best enzyme • promoter and RBS tuning	311
	K. pneumoniae	102.6 g/L, 0.86 g/g-glycerol, 1.07 g/L/h	• promoter library screening • promoter repetition	312
lactic acid	*C. glutamicum*	• L-lactic acid: 212 g/L, 97.9 g/g-glucose • D-lactic acid: 264 g/L, 95.0 g/g-glucose	• push−pull • introduce the ED pathway	313

https://doi.org/10.1021/acs.chemrev.2c00403
Chem. Rev. 2023, 123, 5521−5570

Table 1. continued

chemicals	host	titer	metabolic engineering strategies	ref
		Organic Acid		
			• redox balancing	
glycolic acid	*E. coli* MG1655	65.5 g/L, 0.765 g/g-glucose, 0.85 g/L/h	• block the byproduct formation and product degradation pathway • push−pull	314
malonic acid	*E. coli* MG1655	3.6 g/L, 0.1 g/L/h	• block the byproduct formation • screen the best enzyme	315
succinic acid	*Y. lipolytica*	160 g/L, 0.4 g/g-glycerol, 0.4 g/L/h	• block succinate degradation pathway	316
glutaric acid	*C. glutamicum*	105.3 g/L, 0.54 g/g-glucose, 1.53 g/L/h	• GEM • transcriptome analysis. • block the degradation of product precursor	317
itaconic acid	*E. coli*	43 g/L, 0.6 g/g-glycerol, 1.34 g/L/h	• push−pull−block	318
muconic acid	*C. glutamicum*	54 g/L, 0.197 g/g-glucose, 0.34 g/L/h	• push−pull−block	319
		Amino Acid		
lysine	*C. glutamicum*	120 g/L, 0.55 g/g-glucose, 4 g/L/h	• isotope flux analysis • in silico flux analysis • block the completing pathway	320

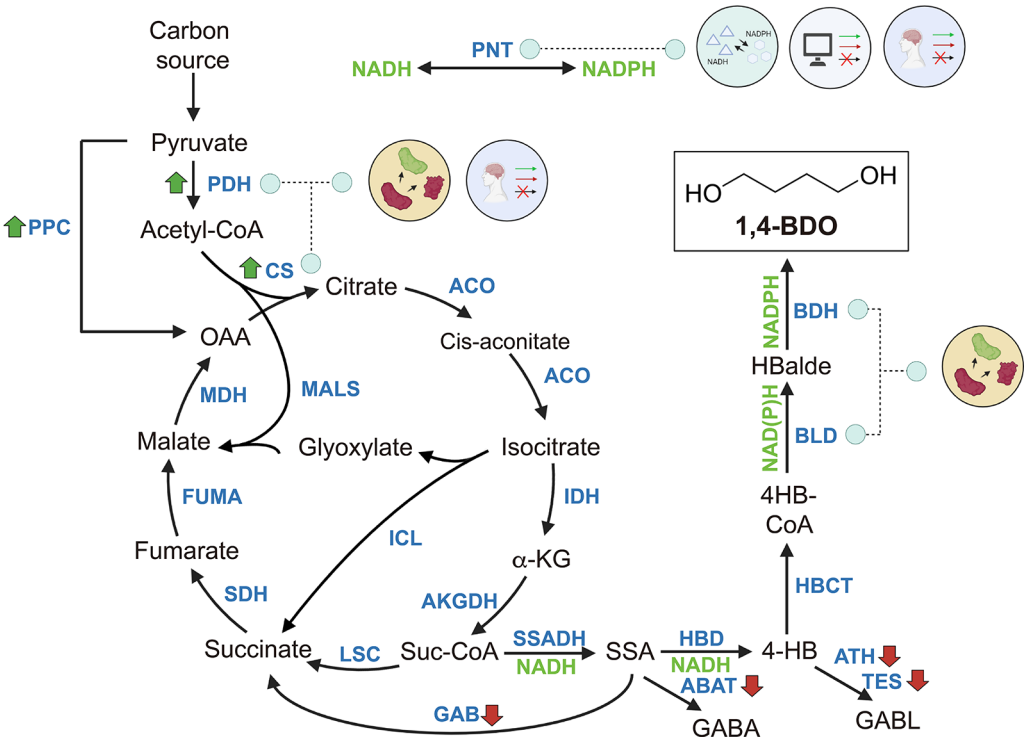


Figure 11. Pathway and metabolic engineering strategies for production of 1,4-BDO. Enzymes: PPC, phosphoenolpyruvate carboxylase; SCT, CoA-dependent succinyl semialdehyde dehydrogenase; HBD, 4-hydroxybutyrate dehydrogenase; HBCT, 4-hydroxybutyryl-CoA transferase; BLD, 4-hydroxybutyryl-CoA reductase; BDH, butyraldehyde dehydrogenase; PNT, pyridine nucleotide transhydrogenase; ABAT, 4-aminobutyrate aminotransferase; GAB, succinate semialdehyde dehydrogenase; ATH, acyl-CoA thioester hydrolase; TES, thioesterase II. Compounds: PEP, phosphoenolpyruvate; SSA, succinate semialdehyde; 4HB, 4-hydroxybutyrate; 4HB-CoA, 4-hydroxybutyryl CoA; Hbald, 4-hydroxybutyraldehyde; GABL, γ-butyrolactone.

itaconic acid-inducible mechanism from *Yersinia pseudotuberculosis*.[289] This novel itaconic acid-inducible system is based on a transcription factor ItcR and *ccl* promoter, in which ItcR can activate the transcription of the *ccl* promoter after binding with itaconic acid. The biosensor was used to optimize the

https://doi.org/10.1021/acs.chemrev.2c00403
Chem. Rev. 2023, 123, 5521−5570

expression of CAD, which led to increased production of itaconic acid.

As mentioned before, the availability of the precursor, *cis*-aconitate, is the main limitation for itaconic acid production. Therefore, knock-down or knockout of the IDH gene is a key metabolic engineering strategy. However, the IDH gene is an essential gene, and deletion of the IDH gene would cause the cell to be auxotrophic to glutamate. Thus, knocking down and dynamic control of the expression of the IDH gene are both prominent metabolic strategies to enhance itaconic acid production. For knocking down the IDH gene expression, its promoter has been replaced with the weak promoter for reduced transcription or the start codon has been changed to GTG or TGT for reduced translation.[285,287] The GTG or TTG replacement for the ATG start codon resulted in the increased yields from 0.09 g/g to 0.21 g/g or 0.24 g/g, respectively. Apart from direct repression of gene expression, Alper and co-workers implemented the expression of AMP deaminase to reduce the AMP level, which is necessary for IDH activity, and this approach increased the production of itaconic acid by 5-fold in *Y. lipolytica*.[290] As for the strategy of dynamic control, Guss and co-workers developed a two-phase fermentation approach, in which cells grew in the first stage, and after the nitrogen source was close to being exhausted, the cell transitioned into a production phase to produce itaconic acid by overexpressing the CAD which was controlled by a nitrogen biosensor.[287] Overall, the cell growth rate increased from 0.35 h^{-1} to 0.58 h^{-1}, and the yield increased from 0.24 g/g to 0.4 g/g using dynamic control. Klamt and co-workers also demonstrated that expression of the IDH gene under temperature-dependent dynamic control could enhance the peak productivity from 0.32 g/L/h to 0.39 g/L/h and the titer from 32 g/L to 47 g/L.[197] Instead of reducing the IDH activity, Shen and co-workers attempted to overcome the glutamate auxotroph by decoupling substrate utilization of glutamate formation and itaconic acid production.[291] In this strategy, glycerol served as the sole carbon source for growth, and the native xylose pathway in *E. coli* was replaced by the Weimberg pathway; thus, xylose was solely converted to α-ketoglutarate for glutamate formation. This strategy led to an accumulation of 20 g/L itaconic acid in *E. coli* without adding exogenous glutamate.

3.1.1.2. 1,4-Butanediol (1,4-BDO). 1,4-Butanediol (1,4-BDO) is an important tetracarbon diol and a platform chemical because it can be converted to industrially important chemicals, including tetrahydrofuran, γ-butyrolactone, and polybutylene succinate. Besides, 1,4-BDO serves as a monomer for biodegradable plastic, such as polybutylene terephthalate (PBT) and polybutylene adipate terephthalate (PBAT). Several pathways have been discovered and utilized for 1,4-BDO production,[292] and the main pathway is the CoA-dependent pathway (Figure 11). The CoA-dependent pathway begins with succinate, which is an intermediate of the TCA cycle and converted to succinyl-CoA by succinyl-CoA synthase. Succinyl-CoA is then converted to 4-hydroxybutyrate (4HB), which is subsequently converted to 1,4-BDO by a CoA-dependent reduction process catalyzed by 4-hydroxybutyryl-CoA transferase (HBCT), 4-hydroxybutyryl-CoA reductase (BLD), and butyraldehyde dehydrogenase (BDH). Pathway redox maintenance requires several gene modifications, which can be identified with computational methods. Hatzimanikatis and co-workers developed a kinetic model which showed that high precursor concentration, energy supply, and reducing equivalents contributed to an increase in the synthesis of 1,4-BDO by 20%. Furthermore, an increase in the activity of phosphofructokinase and adenosine triphosphate (ATP) synthase resulted in a nearly 48.5% increase in 1,4-BDO production.[263] Hu and co-workers applied the CRISPR/Cas system to engineer *E. coli* by performing site-specific mutagenesis of citrate synthase (CS) and gene replacement of the native pyruvate dehydrogenase (PDH) with *Klebsiella pneumonia* PDH to enhance the flux toward the TCA cycle. Furthermore, CRISPR/Cas system was used for whole-pathway integration, and the engineered strain could produce 0.9 g/L 1,4-BDO. Finally, the CRISPRi system was used for blocking the downstream pathway degrading the precursor of 1,4-BDO. By combining the CRISPRi system to simultaneously suppress competing genes that diverted the flux from *gabD*, *ybgC*, and *tesB*, the titer was increased to 1.8 g/L.[293] Meanwhile, slower enzyme kinetics of the downstream enzymes reduced the conversion efficiency of 4HB to 1,4-BDO, so the mutagenesis-based strategies (e.g., error-prone mutagenesis) for butyraldehyde dehydrogenase and butanol dehydrogenase were applied to improve the accumulation of 1,4-BDO, increasing the titer by almost 4-fold.[294] Moreover, the downstream enzymes 4-hydroxybutyryl-CoA transferase and butyraldehyde dehydrogenase were modified to increase the yield by 20%. Furthermore, 1,4-BDO titer of 125 g/L was achieved by integrating strategies including improvement of cofactor regeneration, destruction of byproducts, and over-expression of membrane-bound transhydrogenase.[295]

3.1.1.3. 2-Propanol and Acetone. 2-Propanol is mainly used as a disinfectant and is often called rubbing alcohol. It is also used in pharmaceuticals and personal care products, and it is used as a solvent for herbicides, pesticides, inks, and resins.[296] Acetone is also primarily used as a solvent, but it can also be used in the current biofuel infrastructure by enhancing performance of existing fuels or serving as a precursor to green diesel or jet fuels.[297] 2-Propanol and acetone can be produced together in acetone−butanol−ethanol (ABE) fermentation if the host microorganism harbors a primary−secondary alcohol dehydrogenase (sADH) which converts acetone to 2-propanol.[298] Clostridium species have traditionally been used for ABE fermentation, but the process has been phased out due to limited selectivity of a single product and high cost of C5 and C6 substrates.[297,299] In this case study, Köpke and co-workers chose *Clostridium autoehthanogenum* as their host strain as it is an anaerobic acetogen, so it does not need light like other autotrophs, and it makes use of efficient CO_2 fixing pathways. These selection criteria combined with the fact that there have been significant advancements in engineering tools for acetogens makes *C. autoehthanogenum* a good candidate for scale up gas fermentation.[297]

The pathway for acetone/2-propanol production starts from acetyl-CoA, followed by converting acetyl-CoA into 2-propanol by four cascade enzymes, including thiolase (THL), CoA-transferase (CT), acetoacetate decarboxylase (ADC), and alcohol dehydrogenase (ADH) (Figure 12). Thus, accumulation of intermediate acetone is a key step for high 2-propanol production. Köpke and co-workers first performed enzyme mining of THL, CT, and ADC and assembled a representative group of candidate proteins, including 4 THLs, 16 CTs, and 10 ADCs with strong, medium, and weak promoters, which generated 247 different recombinant *C. autoethanogenum* strains. 247 strains produced a wide range of acetone titer, and the highest titer could reach up to 100 mM, an 11-fold

https://doi.org/10.1021/acs.chemrev.2c00403
Chem. Rev. 2023, 123, 5521−5570

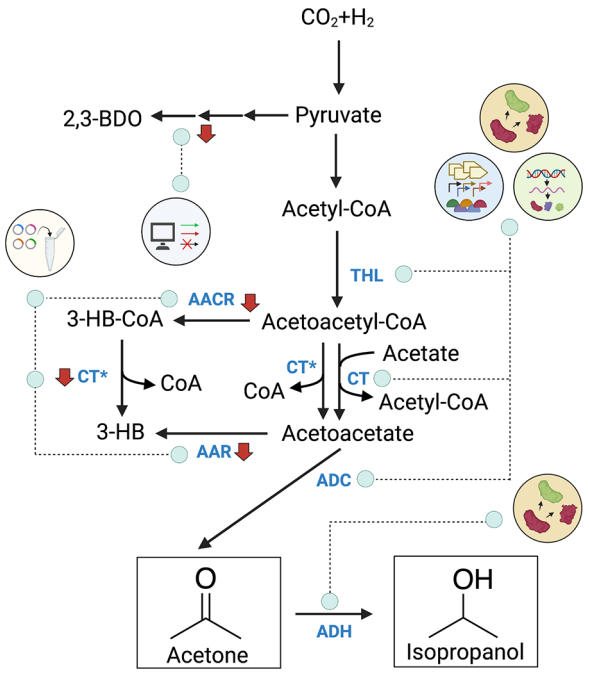


Figure 12. Pathway and metabolic engineering strategies for production of acetone and 2-propanol. Enzymes: THL, thiolase; CT, CoA-transferase; ADC, acetoacetate decarboxylase; ADH, alcohol dehydrogenase; AACR, acetoacetyl-CoA reductase; AAR, acetoacetate reductase. Compounds: 3-hydroxybutyrate, 3-HB; 2,3-butanediol, 2,3-BDO.

enhancement compared to the reference strain. Next, to further increase the flux to acetone, a GEM was used to eliminate the unwanted byproducts including 3-butanediol (2,3-BDO) and 3-hydroxylbutarate (3-HB). One successful gene knockout was identified for reducing 2,3-BDO, but due to lack of information in the current metabolic network, gene targets for reducing 3-HB production were not identified. As an alternative method to find candidate genes, an in silico homology search was carried out on enzymes from literature that perform similar reactions, which led to the identification of 13 gene knockout candidates. However, *in vivo* characterization of the 13 candidates was infeasible due to underdeveloped genetic tools in *C. autoethanogenum*. Instead of *in vivo* characterization, a cell-free strategy called, *in vitro* prototyping and rapid optimization of biosynthetic enzymes (iPROBE), was adopted to determine gene targets. Using iPROBE individual gene targets were cloned and put into cell-free reactions to assess the effect of enriching a single target protein on acetone biosynthesis. Enriched proteins that resulted in low acetone production helped identify 3 targets among 13 candidate genes for reducing 3-HB production. *In vivo* knockout of 3 gene targets identified by iPROBE and one gene target identified by the GEM resulted in a 27-fold increase of acetone production with more than 50 mol% yield. Furthermore, proteomic analysis, kinetic modeling, and reapplying iPROBE, revealed that the CT is the rate limiting step in the acetone-production pathway. To address this bottleneck, an additional copy of CT was overexpressed, and the acetone productivity reached around 2.5 g/L/h and selectivity of acetone/2-propanol increased to 40%. Finally, with expression of mutant ADH and scale-up fermentation to a 120 L fermenter, production rate of ∼3 g/L/h and ∼90% selectivity was achieved. Further life cycle assessment showed that the entire process is carbon negative, which means that it fixes carbon instead of releasing carbon into the atmosphere like traditional production processes.[297]

3.1.2. Fine Chemicals. Besides bulk chemicals, microbial cell factories have been applied to produce fine chemicals. Table 2 lists several examples of oleo-chemical products that are produced by engineered microorganisms. Table 3 lists several natural products that many researchers have worked on extensively including artemisinin, resveratrol, carotenoid, heme, and taxol-derived compounds with the highest reported titers. In this section, artemisinin and omega-3 fatty acids are selected as case studies to discuss how metabolic engineering strategies have been used to enhance the production of fine chemicals.

3.1.2.1. Artemisinin. Artemisinin is naturally produced by a plant, *Artemisia annua*, which has a long history of use in traditional Chinese medicine, and its derivatives were designated as first-line antimalarial drugs by the World Health Organization in 2002.[321] Until now, microbial synthesis of artemisinin is the most successful case for the biosynthesis of natural products and drugs. The effort on metabolic engineer-

Table 2. List of Representative Oleo-chemicals Produced through Metabolic Engineering

chemicals	host	titer	metabolic engineering strategies	ref
		Oleo-chemicals		
oleoylethanolamide	*S. cerevisiae*	8.1 mg/L, 405.8 μg/g-glucose	• host strain engineering to ensure flux toward phospholipids • heterologous enzyme expression to shunt carbon away from native lipids • relaxation of native regulation	329
triacylglycerol	*S. cerevisiae*	1.76 g/L, 0.088 g/g-glucose	• push−pull−block: ACC1**, PAH1, DGA1 • knockout of lipases	81
human *ceramide-NS*	*S. cerevisiae*	N/A	• introduction of human hDES1 • ER localization of hDES1 • push−pull−block of host genes	330
omega-3 fatty acids	*Y. lipolytica*	56.6% w/w of total fatty acids, 30% of DCW	•push−pull−block of host genes • introduction of heterologous genes for chain elongation and desaturation	328

https://doi.org/10.1021/acs.chemrev.2c00403
Chem. Rev. 2023, 123, 5521−5570

Table 3. List of Representative Natural Product Produced through Metabolic Engineering

chemicals	host	titer	metabolic engineering strategies	ref
		Natural Product		
artemisinin	*S. cerevisiae*	25 g/L, 0.156 g/L/h	• mine new enzyme	327
resveratrol	*E. coli* BW27784	2.34 g/L	• screen the best enzyme • promoter replacement • medium optimization	331,332
carotenoid	*S. cerevisiae* *Y. lipolytica*	lycopene: 2.3 g/L *β*-carotene: 6.5 g/L	• spatial engineering • promoter shuffling	333,334 334,335
	E. coli	astaxanthin: 432 mg/L, 9.62 mg/L/h	• protein engineering for soluble expression • spatial engineering • kinetic model analysis	334,336
heme	*C. glutamicum*	309 mg/L. 2.1 mg/g-glucose, 6.4 mg/L/h	• screen the best *C. glutamicum* strain • transcription factor engineering • membrane engineering	337
taxol-derived chemicals	*E. coli*	300 mg/L taxadiene	• promoter tuning • protein engineering	338,339
opioids-derived chemicals	*S. cerevisiae*	6.4 *μ*g/L thebaine 0.3 *μ*g/L hydrocodone	• modular design • protein engineering • mine new enzymes • push−pull−block	13
cannabinoids	*S. cerevisiae*	8.0 mg/L tetrahydrocannabinolic acid 4.2 *μ*g/L cannabidiolic acid	• modular design • mine new enzymes • spatial engineering • push−pull	340
vinblastine	*S. cerevisiae*	23.9 *μ*g/L	• push−pull−block • redox balance • spatial engineering • mine new enzyme • protein engineering • proteomic analysis	341

ing of artemisinin has been comprehensively reviewed.[322] *E. coli* was the first engineered microorganism for artemisinin production and produced 25 g/L of the precursor amorphadiene through enzyme screening and process optimization. However, it is hard for *E. coli* to further convert amorphadiene into artemisinin because it is unsuitable for the expression of eukaryotic P450 enzymes. Here, we only briefly introduce a study based on *S. cerevisiae*. The goal is to produce artemisinic acid by the microorganism, and the final step converting artemisinic acid into artemisinin is accomplished by chemical conversion under mild industrially favorable conditions. The metabolic pathway and strategies are illustrated in Figure 13. In the initial attempt, the mevalonate pathway, geranyl diphosphate synthase (GPPS), and amorpha-4,11-diene synthase (ADS) were overexpressed in *S. cerevisiae* S288C, but only 150 mg/L amorphadiene was produced.[323] Additional expression of the cytochrome P450 enzyme from *A. annua* and its cognate reductase (CPR1) resulted in 100 mg/L artemisinic acid. In order to increase the artemisinic acid production, Regentin and co-workers optimized the culture conditions and inhibited the sterol production pathway to increase the precursor availability and the final strain produced 2.5 g/L artemisinic acid.[324] Instead of using *S. cerevisiae* S288C, Paddon and co-workers attempted to engineer *S. cerevisiae* CEN.PK2 with the similar metabolic engineering strategies developed by Keasling and co-workers while also optimizing the culture process. However, only 1.6 g/L artemisinic acid was produced, while a production of amorphadiene up to 40 g/L was unexpectedly observed, which indicated that expression of cytochrome P450 is not sufficient for converting amorphadiene to artemisinic acid.[325] Further effort was spent on optimizing the conversion of amorphadiene to artemisinic acid by exploring additional cytochrome systems. Because the interaction of cytochrome b5 with cytochrome P450 enzymes can increase the cytochrome P450 reaction rate,[326] cytochrome b5 from *A. annua* was expressed in the CEN.PK2 production strain, which resulted in increased production of artemisinic acid from 0.7 g/L to 2.4 g/L. Furthermore, the production of artemisinic acid was increased to 7.1 g/L by overexpressing the *A. annua* artemisinic aldehyde dehydrogenase (ALDH) and NAD-dependent artemisinic alcohol dehydrogenase (ADH) and by deleting *GAL80* for enhanced protein expression. Finally, 25 g/L

https://doi.org/10.1021/acs.chemrev.2c00403
Chem. Rev. 2023, 123, 5521−5570

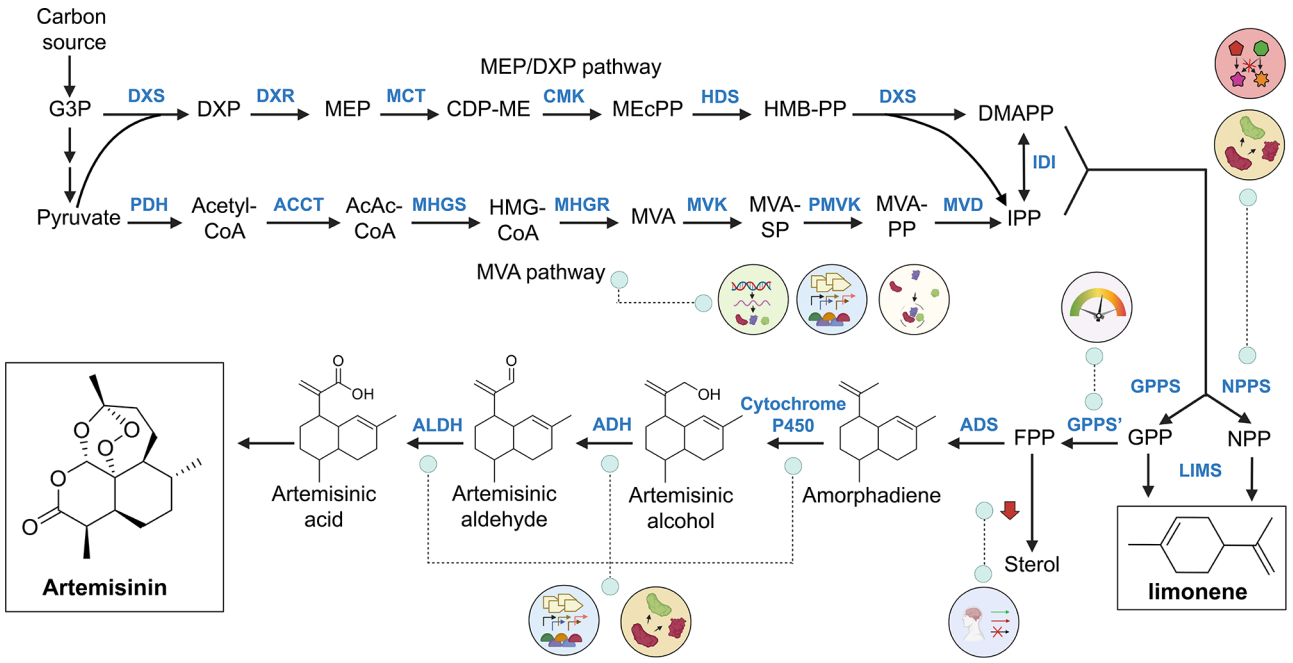


Figure 13. Pathway and metabolic engineering strategies for production of limonene and artemisinin. Enzymes: DXS, 1-deoxy-D-xylulose 5-phosphate synthase; DXR, 1-deoxy-D-xylulose 5-phosphate reductase; MCT, methyl-D-erythritol 4-phosphate cytidylyltransferase; CMK; MDS, 2-*C*-methyl-D-erythritol 2,4-cyclodiphosphate synthase; HDS, 4-hydroxy-3-methylbut-2-enyl diphosphate synthase; HDR, 4-hydroxy-3-methylbut-2-enyl diphosphate reductase; PDH, pyruvate dehydrogenase; ACCT; MHGS; MHGR; MVK; PMVK; MVD; IDI, isopentenyl-diphosphate isomerase; NPPS; GPPS; LIMS, limonene synthase; ADS, amorpha-4,11-diene synthase; ADH, artemisinic alcohol dehydrogenase; ALDH, artemisinic aldehyde dehydrogenase. Compounds: G3P, glyceraldehyde-3-phosphate; DXP, 1-deoxy-D-xylulose-5-phosphate; MEP, 2-*C*-methyl-D-erythritol-4-phosphate; CDP-ME, 4-diphosphocytidyl-2-*C*-methyl-D-erythritol; CDP-MEP, 4-diphosphocytidyl-2-*C*-methyl-D-erythritol 2-phosphate; MEcPP, 2-*C*-methyl-D-erythritol-2,4-cyclodiphosphate; HMB-PP, 4-hydroxy-3-methylbut-2-enyl-diphosphate; AcAc-CoA, acetoaceyl-CoA; HMG-CoA, 3-hydroxy-3-methylglutaryl-CoA; MVA, mevalonate; MVA-5P, MVA-5-phosphate; MVA-PP, 5-bisphosphomevalonate; DMAPP, dimethylallyl pyrophosphate; IPP, isopentenyl pyrophosphate; NPP, neryl diphosphate; GPP, geranyl diphosphate; FPP, arnesyl pyrophosphate.

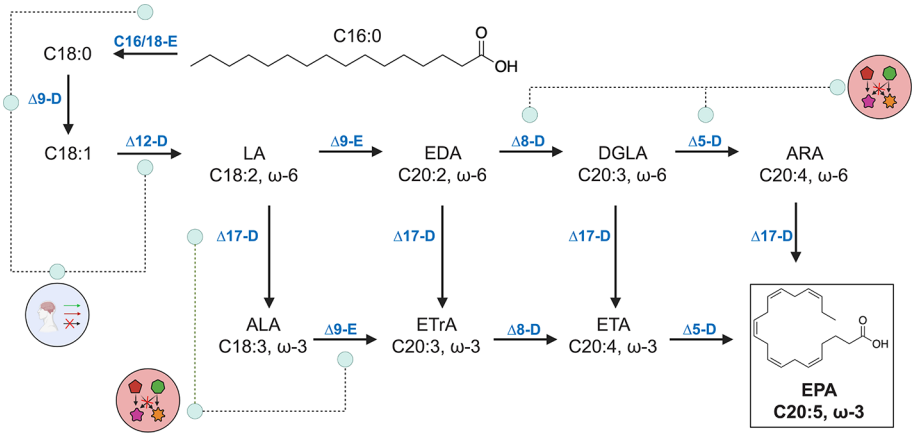


Figure 14. Metabolic engineering strategies for the overproduction of ω-3 fatty acids involves a combination of push−pull−block deriving from knowledge-guided design and orthogonal strategies for expressing various desaturases and elonases in the host. Enzymes: C16/18-E, C16/18 elongase; Δ9-D, Δ-9 desaturase; Δ12-D, Δ-12 desaturase; Δ9-E, Δ-9 elongase; Δ17-D, Δ-17 desaturase; Δ8-D, Δ-8 desaturase; Δ5-D, Δ-5 desaturase. Compounds: C16:0, palmitic acid; C18:0, stearic acid; C18:1, oleic acid; LA, linoleic acid; ALA, α-linolenic acid; EDA, eicosadienoic acid; ETrA, eicosatrienoic acid; DGLA, dihomo-γ-linolenic acid; ETA, eicosatetraenoic acid; ARA, arachidonic acid; EPA, eicosapentaenoic acid.

artemisinic acid was produced after optimizing the fermentation process.[327]

3.1.2.2. Omega-3 Fatty Acids. Omega-3 fatty acids are traditionally produced via marine fisheries, a slow process that struggles to keep up with the growing market demand for nutraceuticals. For example, eicosapentaenoic acid (EPA) and docosahexaenoic acid (DHA) have contributed to increased market demand as their supplemental health benefits have become clear. Microbial synthesis of these molecules offers a promising solution to address the growing demand. Xue and co-workers engineered a *Y. lipolytica* strain that produced EPA at 56.6% of the total fatty acids through microbial fermentation.[328] The researchers integrated a heterologous pathway that elongates the natively produced linoleic acid (C18:2 fatty acid (FA)) to a 20 carbon FA, followed by 3 desaturation steps leading to the synthesis of EPA (w-3 C20:5 FA). The integrations were performed using NHEJ and the sites of random integration were identified using genome

https://doi.org/10.1021/acs.chemrev.2c00403
Chem. Rev. 2023, 123, 5521−5570

Table 4. List of Representative Biofuels Produced through Metabolic Engineering

chemicals	host	titer	metabolic engineering strategies	ref
		Short-Chain Biofuels		
propanol	*E. coli*	10.3 g/L, 0.259 g/g-glucose, 0.083 g/L/h	• protein engineering • knockout the RpoS sigma factor to enhance the TCA	38
2-propanol	*E. coli*	143 g/L, 0.23 g/g-glucose, 0.6 g/L/h	• screen the best enzyme set • medium optimization • product extraction	362,363
	C. autoethanogenum	3g/L/h and 90% selectivity	• mine new enzymes • promoter and enzyme candidate combinatorial library • kinetic model • GEM • cell-free ML assessment by iPROBE • proteomic analysis	
butanol	*Clostridium acetobutylicum*	130 g/L, 0.31 g/g-glucose, 1.32 g/L/h	• metabolic flux and mass balance for selecting a better pathway • push−block	364
isobutanol	*E. coli*	50.9 g/L, 0.37 g/g-glucose, 0.7 g/L/h	• rewire the amino acid metabolism • block byproduct formation pathway • ALE	345,350,365
pentanol	*E. coli*	4.3 g/L	• protein engineering	366
isopentanol	*C. glutamicum*	2.76 g/L, 0.1 g/g, 0.058 g/L/h	• block the competing pathway • screen the best enzyme	367,368
isoprene	*S. cerevisiae*	3.7 g/L, 22.9 mg/g-glucose 39 mg/L/h	• transcription factor engineering • promoter engineering • protein engineering	369
		Medium-Chain Biofuels		
limonene	*E. coli* BL21 (DE3)	3.6 g/L, 0.07 g/g-glycerol, 0.15 g/L/h	• push−pull • medium optimization	359
		Long-Chain Biofuels		
free fatty acids	*S. cerevisiae*	33.4 g/L, 0.1 g/g-glucose	• heterologous pathway for FFA • cofactor engineering • adaptive evolution	360
fatty alcohols	*Y. lipolytica*	5.8 g/L, 36 mg/g-glucose, 39 mg/L/h	• heterologous FAR expression • bioreactor optimization • extractive fermentation	370

walking and genome sequencing. It was discovered that NHEJ-mediated insertions led to the disruption of the genes *PEX10*, *LEU2*, *LIP1*, and *SCP2*. Three of these genes are related to lipid metabolism, and their disruption led to a positive impact on the EPA titer. The authors studied the impact of *PEX10* inactivation, which resulted in a defective β-oxidation pathway. The strains with *PEX10* disruption contained deformed and dysfunctional peroxisomes. This defect led to an accumulation of lipids, resulting in higher titers of EPA as well as other long-chain polyunsaturated fatty acids such as dihomo-γ-linolenic acid and arachidonic acid. In all, the engineered strain contained 30 copies of 9 genes and accumulated EPA at 56.6% of the total fatty acid content, while saturated fatty acids accumulated at less than 5%. The net accumulation of EPA in the engineered strain was 15% of the dry cell weight, making it the highest reported titer of EPA in yeast at the time of publication. This work constitutes an early example of the production of a nutritional supplement in yeast to replace animal-derived products (Figure 14).

3.2. Synthesis of Fuels

Recent awareness of sustainability and eco-friendliness has drawn attention toward biofuels for their ability to replace petroleum-based fuels.[342] It was estimated that biofuels could fill 27% of the global demand for transportation fuels by 2050.[343] Thus, engineering microbial cell factories for biofuels production is highly desirable. Table 4 lists several biofuels that can be produced via metabolic engineering. Here, isobutanol, limonene, and free fatty acids are selected as case studies of short-, medium-, and long-chain biofuel to showcase how metabolic engineering strategies have been used to enhance biofuel production.

3.2.1. Isobutanol. Isobutanol is an attractive option for biofuel because its energy density is similar to that of 1-butanol

https://doi.org/10.1021/acs.chemrev.2c00403
Chem. Rev. 2023, 123, 5521−5570

but possesses a higher octane number, which is preferable for blending into gasoline to reduce engine knocking.[344] The pathway of isobutanol production involves the valine synthesis pathway, but precursors will be directed to isobutanol instead of valine (Figure 15).[345] In the valine synthesis pathway,

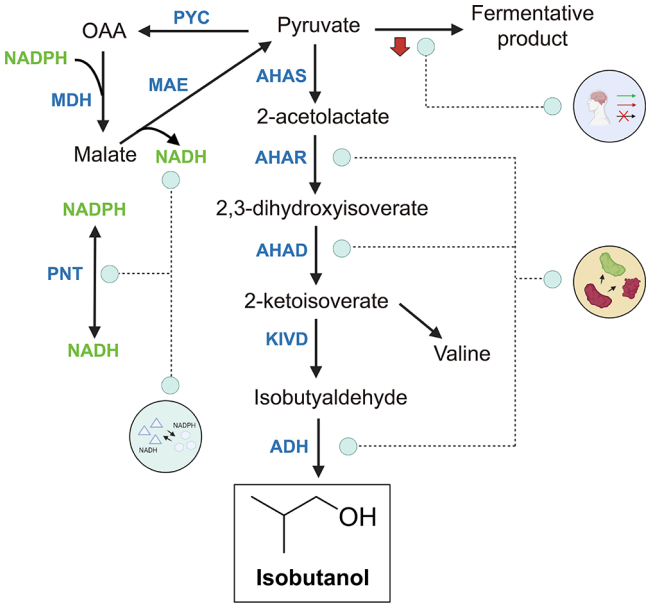


Figure 15. Pathway and metabolic engineering strategies for production of isobutanol. Enzymes: AHAS, acetohydroxyacidsynthase; AHAR, acetohydroxyacid reductase; DHAD, 2,3-dihydroxyisoverate dehydrogenase; KIVD, 2-ketoisoverate dehydrogenase; ADH, aldehyde dehydrogenase; MAE, malic enzyme.

pyruvate is converted to 2-ketoisovalerate by three enzymes, including the acetolactate synthase (AHAS), keto acid reductoisomerase (AHAR), and dihydroxy-acid dehydratase (DHAD). 2-Ketoisovalerate is the key precursor for valine and isobutanol production, thus, to rewire the flux to the isobutanol production, two enzymes, including 2-ketoacid decarboxylases (KIVD) and alcohol dehydrogenase (ADH), must be expressed. In this pathway, two molecules of pyruvate will be converted to one molecule of isobutanol, and two molecules of reducing equivalent are required in the reactions catalyzed by AHAR and ADH. Thus, redox balancing is important for isobutanol production because *E. coli* does not produce NADPH efficiently during anaerobiosis. To overcome this limitation, one strategy is to select enzymes utilizing NADH as a cofactor instead of NADPH. Liao and co-workers attempted to overexpress the ADH from different organisms and found that expression of ADH from *Lactococcus lactis* led to 8 g/L isobutanol production as compared to 6 g/L using the ADH from *E. coli*, which is attributed to the fact that the ADH from *L. lactis* utilizes NADH as a cofactor instead of NADPH.[346] While selecting enzymes from different organisms is a popular strategy, protein engineering is also an eminent strategy to switch cofactor specificity. Arnold and co-workers performed site-specific saturation mutagenesis on AHAR, and one mutant showed a higher specific activity with NADH than with NADPH (i.e., 0.65 vs 0.07 U/mg).[347] In addition to the engineering of AHAR, the same study also performed the random mutagenesis of ADH from *L. lactis* and successfully enhanced the catalytic efficiency by 29-fold. By utilizing the mutant ADH, the isobutanol production increased from 1.5 g/L to 13 g/L.[347] A third approach is to increase the NADPH supply via converting the NADH to NADPH by pyridine nucleotide transhydrogenase. The same study also demonstrated that 8.2 g/L of isobutanol could be produced by overexpression of a pyridine nucleotide transhydrogenase.[347] In yeast, pyridine nucleotide transhydrogenase cannot be successfully expressed, thus Kondo and co-workers regenerated the NADPH by creating a transhydrogenase-like shunt in which pyruvate is sequentially converted to oxaloacetate, malate, and pyruvate by pyruvate carboxylase (PYC), malate dehydrogenase (MDH), and malic enzyme (MAE). With this shunt, one molecule of NADH is converted to one molecule of NADPH, and the production of isobutanol could be enhanced from 44 mg/L to 83 mg/L.[348] In the same study, 12 individual gene deletions were performed to increase the pyruvate availability, and the strain with LPD1 deletion showed a 7.5-fold improvement of isobutanol production.[348] Substrate channeling has also been pursued to increase product formation. For isobutanol production with enhanced substrate channeling, Stephanopoulos and co-workers established the isobutanol pathway in mitochondria and showed a 2.6-fold improvement (i.e., 635 mg/L).[61] In addition, ALE provides an alternative for enhancing isobutanol production. Smith and Liao performed random mutagenesis of *E. coli* to screen for a strain that could resist a toxic valine analogue (i.e., norvaline). This resulted in a pull of more flux toward the valine pathway and enhanced the precursor pool for isobutanol production. By ALE, a mutant *E. coli* strain was obtained that improved isobutanol production from 3.1 g/L to 6.1 g/L.[349] With additional metabolic engineering and process optimization, the highest isobutanol production could reach 50 g/L with a yield of 0.37 g/g glucose and a productivity of 0.7 g/L/h.[349,350]

3.2.2. Limonene. Limonene is regarded as a potential next-generation jet biofuel because its hydrogenated form has a low freezing point and is immiscible with water,[351] thereby enhancing cold-weather performance. Limonene is also notable for its pleasant orange-scented fragrance and its designation as a Generally Recognized As Safe (GRAS) molecule has driven demand for the inclusion of limonene in eco-friendly cleaning products. Limonene and its derivatives are synthesized by condensing two basic building blocks, isopentenyl diphosphate (IPP) and dimethylallyl diphosphate (DMAPP). For the synthesis of the building blocks, there are two main pathways, the mevalonate (MVA) pathway, and the methylerythritol phosphate (MEP) pathway (Figure 13).

The MVA pathway is typically found in animals, yeasts, and some Gram-positive bacteria. On the other hand, the MEP pathway can be discovered in most Gram-negative bacteria, green algae, and cyanobacteria.[64] In the MVA pathway, MVA is the key precursor for the biosynthesis of IPP and DMAPP, thus balancing the overexpression of the acetoacetyl-CoA transferase (ACCT), 3-hydroxy-3-methylglutaryl-CoA synthase (MHGS), and 3-hydroxy-3-methylglutaryl-CoA reductase (MHGR) is important for limonene production. It has been reported that accumulation of the pathway intermediate HMG-CoA caused the observed growth inhibition in *E. coli*,[352] while it has also been observed that reduced expression of HMGS and MHGR enhanced MVA production 7-fold.[353] Zhao and co-workers tried to improve the limonene production in *E. coli* by optimizing the RBS of ACCT and MHGS expression, and 1.29 g/L limonene was produced.[354] Although RBS engineering is an eminent way of optimizing the MVA pathway, the exact expression level is still not clear. Lee

https://doi.org/10.1021/acs.chemrev.2c00403
Chem. Rev. 2023, 123, 5521−5570

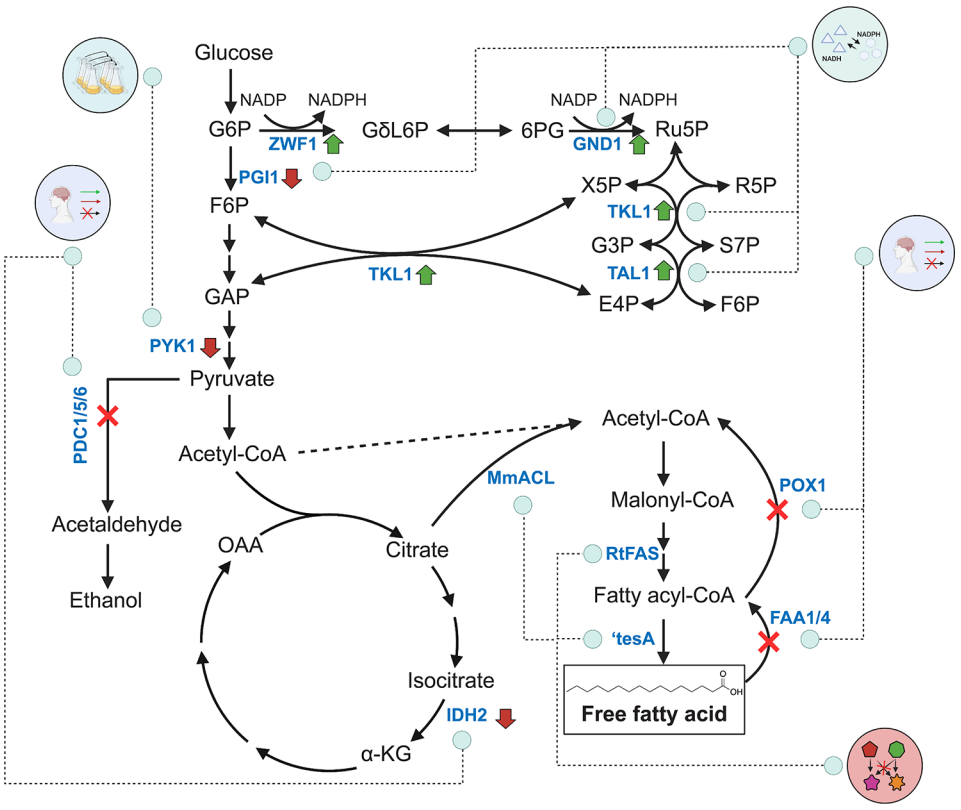


Figure 16. Metabolic engineering strategies for the overproduction of free fatty acids involve a combination of knowledge-guided design, cofactor engineering, orthogonal strategies, and adaptive laboratory evolution. Enzymes: ZWF1, cytoplasmic glucose-6-phosphate dehydrogenase; PGI1, phosphoglucose isomerase; GND1, isoform 1 of phosphogluconate dehydrogenase; TKL1, transketolase 1; TAL1, transaldolase 1; PYK1, pyruvate kinase; PDC1/5/6, pyruvate carboxy 1/5/6; IDH2, subunit 2 of mitochondrial NAD^+-dependent isocitrate dehydrogenase; MmACL, ATP:citrate lyase from *Mus musculus*; RtFAS, fatty acid synthase from *R. toruloides*; 'tesA, truncated *E. coli* thioesterase; POX1, fatty acyl-CoA oxidase; FAA1/4, fatty acyl-CoA synthetase 1/4. Compounds: G6P, glucose 6-phosphate; GδL6P, 6-phosphogluconolactone; 6PG, 6-phosphogluconate; Ru5P, D-ribulose 5-phosphate; X5P, xylulose 5-phosphate; R5P, ribose 5-phosphoric acid; GAP, glyceraldehyde 3-phosphate; S7P, sedoheptulose 7-phosphate; E4P, erythrose 4-phosphate; F6P, fructose 6-phosphate; α-KG, α-ketoglutarate; OAA, oxaloacetic acid.

and co-workers performed a proteomic analysis and used the computational tool PCAP to guide metabolic engineering, which resulted in over a 40% improvement in the production of limonene with a titer of 0.605 mg/L in *E. coli*.[214]

The MVA pathway starts with acetyl-CoA, followed by a series of enzymatic reactions to produce the precursor MVA. In eukaryotic cells, acetyl-CoA is synthesized in several compartments, such as mitochondria and peroxisomes. Therefore, repurposing the compartment for chemical production has been an effective metabolic engineering strategy for isoprenoid production.[64,355] This strategy is also useful for minimizing the toxic effect of isoprenoids produced by the MVA pathway.[62] On the other hand, the dynamic control also plays an important role in limonene production in yeast because the precursor GPP, which pulls flux away from limonene, will be converted to downstream essential cellular components, such as sterols, ubiquinone, and FPP for protein prenylation. Therefore, dynamic control could help cells accumulate GPP without cell growth impairment for limonene production. Vickers and co-workers came up with the idea to dynamically degrade FPPS by N-degron and coupled it with the ergosterol responsive promoter (i.e., ERG1 promoter) to express the N-degron-fused GPPS, which resulted in 76 mg/L limonene production in *S. cerevisiae*.[356] Several groups also coupled the dynamic control with orthogonal biosynthesis, in which limonene was produced by the precursor of neryl diphosphate (NPP) instead of GPP. Kampranis and co-workers first engineered the limonene synthase to be NPP-specific and then expressed the mutant limonene synthase with the NPPS under the control of the ERG1 promoter, which allowed *S. cerevisiae* to produce 130 mg/L limonene.[357] Another strategy is to express the GPPS under the glucose-sensing promoter, thereby creating a two-phase fermentation system where glucose is first directed to produce the essential component GPP and then the cells enter the production stage by expressing the NPPS and limonene synthase. With this approach, the final strain could produce 917 mg/L limonene.[358] Currently, the highest production of limonene was achieved in *E. coli* with a titer of 3.6 g/L, a yield of 0.07 g/g glycerol, and a productivity of 0.15 g/L/h, which is accomplished by overexpressing the MVA pathway from different designs of the plasmid library, using proteomic analysis to determine the amounts of pathway enzymes, optimizing the IPTG concentration (i.e., tuning the gene expression), using glycerol as a substrate, and using an extracted two liquid-phase fed-batch strategy.[359]

3.2.3. Free Fatty Acids. Numerous studies on lipid metabolism have enabled considerable metabolic engineering research to produce lipid-related chemicals in yeast *S. cerevisiae*. However, *S. cerevisiae* is restricted in lipid production compared to other oleaginous yeasts due to its evolutionary trajectory toward ethanol production. To address this

https://doi.org/10.1021/acs.chemrev.2c00403
Chem. Rev. 2023, 123, 5521−5570

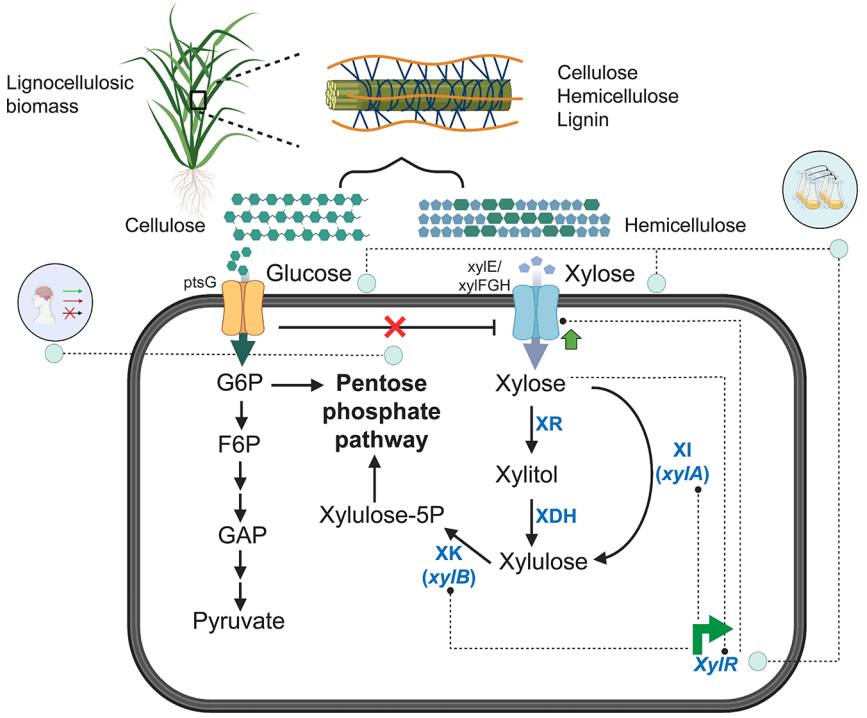


Figure 17. Co-utilization of glucose and xylose sugars for efficient conversion of lignocellulosic substrates into target molecules. The pretreatment of lignocellulosic substrates mostly yields glucose and xylose, each of which are transported into the cell via their respective transporters. Rational engineering strategies to improve co-utilization efficiency involve the inactivation of the CCR, including transporters, to allow the activation of xylose metabolism. The XylR transcriptional activator serves as a central regulator for the entire xylose metabolism, increasing expression of relevant genes upon sensing xylose in the cell. Thus, a common strategy involves generating mutations in XylR that show improved binding and activation of its targets via directed evolution. On the whole-cell level, ALE methods have also been applied to evolve better coutilization phenotypes. Enzymes: XR, xylose reductase; XDH, xylitol dehydrogenase; XI, xylose isomerase; XK, xylulose kinase; XylR, transcriptional regulator of xylose metabolism. Compounds: G6P, glucose 6-phosphate; GAP, glyceraldehyde 3-phosphate; F6P, fructose 6-phosphate.

limitation, Yu and co-workers re-engineered the native metabolism of *S. cerevisiae* into that of an oleaginous yeast.[360] First, a platform strain was chosen from a different study conducted in the same lab that employed a heterologous pathway for the synthesis of free fatty acids (FFA) in *S. cerevisiae*.[361] The construction of the platform strain followed traditional push-pull-block strategies to increase the supply of the precursor, cytosolic acetyl-CoA (push), increase the flux toward fatty acyl-CoA and FFAs (pull), and disrupt the consumption of FFAs elsewhere (block). The resulting platform strain could produce FFAs at a titer of 7.0 g/L in fed-batch fermentation. Second, several strategies were borrowed from the metabolism of oleaginous yeasts to ensure higher FFA production, such as increasing NADPH supply by increasing flux through the pentose phosphate pathway, downregulating flux in the citric acid cycle, thereby pushing more citrate from mitochondria to cytosol, and decoupling lipid accumulation to growth by cultivating under nitrogen-limited conditions. Finally, the authors abolished ethanol formation by deleting PDC1 from the FFA overproducing strain. To recover growth on glucose after this deletion, ALE was performed that resulted in a "synthetic oil yeast" capable of a high titer of FFA without producing any ethanol. The highest titer from this study was 33.4 g/L of FFA in fed-batch cultures, representing the highest reported titer of FFA production in yeast at the time of its publication (Figure 16).

3.3. Substrate Utilization

Whenever a new biosynthetic pathway is constructed in microorganisms, glucose serves as the most common choice of carbon source. Other alternative hexose sugars (e.g., galactose, mannose, fructose), pentose sugars (e.g., xylose and arabinose), 3 carbon substrates like glycerol, 2 carbon substrates like acetate, and 1 carbon substrates (e.g., CO_2, CO, methanol) have been explored as alternative substrates and have showed additional sustainability benefits, and in some cases better conversion to products. Here we discuss the utilization and coutilization of different carbon sources.

3.3.1. Co-utilization of Glucose and Xylose. To reduce both the dependency on glucose as a carbon source and the production cost, researchers have shifted focus to the utilization of alternative, cheap, renewable feedstocks to produce various biochemicals, and these include lignocellulosic biomass,[371−373] crude glycerol source,[316,374−376] and other carbon sources. Glucose, xylose, and arabinose are the main carbon sources obtained after the pretreatment of biomass. Co-utilization of these carbon sources by the microbial system is a promising approach to reduce cost and at the same time improve carbon yield. However, co-utilization of mixed carbon sources (e.g., glucose-xylose/glucose-glycerol/glucose-acetate) is challenging and requires an intensive rewiring of the microbial metabolic network.[372,377]

Numerous model and nonmodel microorganisms can assimilate different pentoses and hexoses individually, but they preferably use glucose first and more efficiently than xylose. When both glucose and xylose are present, microorganisms show diauxic growth phases because of carbon catabolite repression (CCR).[372,378] To reduce CCR for efficient co-utilization of mixed sugars, several strategies have

https://doi.org/10.1021/acs.chemrev.2c00403
Chem. Rev. 2023, 123, 5521−5570

been implemented, including inactivation or deletion of *ptsG*, replacement of native cAMP with a cAMP-independent mutant, overexpression of *xylR* (i.e., xylose transcriptional activator), deletion of *araC* (i.e., L-arabinose transcriptional regulator), inactivation of *ptsHIcrr* operon and overexpression of *galP*, and directed evolution of transporter proteins in *E. coli* (Figure 17).[379] Wang and co-workers adopted ALE to achieve higher efficiency of xylose fermentation.[380] After one round of ALE, evolved strains showed a point mutation in a transcriptional activator for xylose catabolic operons, either CRP or XylR (R121C and P363S, respectively). XylR mutants showed a higher affinity to their DNA binding sites, leading to a xylose catabolic activation independent of catabolite repression control. Introducing this mutant into the D-lactate producing *E. coli* TG114 improved the product titer by 50%. In another study, deletion of *ptsG* in *E. coli* and further optimization of glucose and xylose ratio led to consumption of 93% of xylose and 97% of glucose in 24 h of fermentation and increased succinate titer to 107.0 g/L.[380] A similar strategy was adopted to produce 15.8 g/L of 4-hydroxymandelic acid[381] and 5.2 g/L of *n*-butanol from a mixed carbon source in *E. coli*.[382] Because *ptsG* inactivation usually alleviates the glucose transportation rate which affects the growth rate, the Zmglf gene from *Z. mobiliz* encoding a glucose facilitator was overexpressed to stimulate the coutilization of sugars.[382] For other organisms which cannot assimilate xylose as a sole carbon source, the xylose utilization pathway needs to be introduced to them. *S. cerevisiae* can utilize glucose but not xylose. In one study, xylose reductase (XR), xylitol dehydrogenase (XDH), and xylulokinase (XK) were overexpressed along with the modular expression of HXT1 transporter to produce ethanol from mixed sugars. However, xylose consumption was still not efficient. In another study, mutations were studied in yeast hexose transporters Hxt7 and Gal2 by calculating the kinetic properties of the mutant transporters, and a mutation in Gal2 (N376F) was found to yield aglucose-insensitive xylose transporter.[383] Moreover, the xylose utilization pathway alongside the deletion of RPE1 (i.e., D-ribulose-5-phosphate 3-epimerase) is another way to achieve the coutilization of glucose and xylose for ethanol production.[384]

3.3.2. C1 Chemical Utilization. Concerns regarding climate change and global warming have motivated the engineering of microorganisms for the assimilation and utilization of greenhouse gases as a carbon source to produce molecules. Direct carbon capture from the environment by recalcification is not cost-efficient, generates secondary pollution, and often requires advanced technology for carbon capture. On the other hand, microorganisms such as photoautotrophs and aerobic chemoautotrophs and plants can capture C1 gases and convert them into a range of molecules efficiently. So far, seven natural CO_2-fixing pathways have been discovered: Calvin Benson Bassham cycle (CBB), Wood−Ljungdahl (WL) pathway, reductive TCA (rTCA) cycle, 3-hydroxypropionate-4-hydroxybutyrate (HP/HB) cycle, dicarboxylate/4-hydroxybutyrate (DC/HB) cycle, 3-hydroxypropionate (3-HP) bicycle, and reductive glycine pathway (rGlyP).[385] These pathways are widely distributed in different organisms and capable of utilizing C1 sources with varying efficiency.

3.3.2.1. Natural C1 Utilizing Microorganisms. Naturally, C1 fermenting organisms include anaerobic clostridial acetogens, aerobic chemolithoautotrophic like *Cupriavidus necator*, and various photoautotrophic cyanobacteria that can utilize CO_2 and produce different chemicals. As an example, the case study on 2-propanol/acetone discussed earlier uses acetogenic bacteria. Carbon sources derived from industrial waste gas (e.g., CO), municipal solid waste (e.g., CO, CO_2, H_2), biogas (e.g., CO + H_2), and biomass (e.g., CO_2 + H_2) were used for production of more than 30 chemicals including ethanol, 1−3 butadiene, isobutene, and long-chain alcohols. Acetogenic bacteria employ the WL pathway, which utilizes either CO or CO_2 plus H_2 to synthesize acetyl-CoA that can be further channeled toward biomass or biochemical production. Nonengineered *C. autoethanogenum* has been utilized to convert waste gas into ethanol in continuous fermentation at the scale of more than 90,000 tons per year.[386] Acetogens can produce ethanol natively but have also been genetically engineered for the synthesis of a range of non-native commodity chemicals.

Apart from acetogens, *C. necator* can assimilate CO_2 by using oxygen or hydrogen as the energy source. Under nutrient limitation, *C. necator* can direct its reduced carbon into the synthesis of polyhydroxybutyrate (PHB) and accumulate PHB at 70% of total cell weight.[387] *C. necator* was further engineered to produce 2-propanol at 3.44 g/L in heterotrophic batch conditions by overexpressing codon-optimized clostridial genes in a *phaB*/*phaC* double mutant.[388] By utilizing only auxotrophic growth, 250 mg/L of 2-propanol was produced. The alkane biosynthesis pathway was also expressed in *C. necator* by heterologously expressing an acyl-ACP reductase and an aldehyde-deformylating oxygenase in the PHA knockout strain, which led to the production of up to 4.4 mg/L alkanes.[389]

3.3.2.2. Synthetic C1 Carbon Assimilating Microorganisms. Because natural C1 carbon assimilating microorganisms grow slowly, synthetic microorganisms can be more beneficial. These heterotrophic hosts can attain better growth and are easier to engineer because of the availability of a genetic toolbox and detailed knowledge of the cellular metabolisms. Among them are *E. coli*, *S. cerevisiae*, and *P. pastoris*, which have been engineered for CO_2 fixing as well as utilization of C1 substrates like methanol and formate.[385]

So far, the CBB cycle, starting with the carboxylation reaction, where the RuBisCo enzyme fixes C1 in the form of CO_2 in the ribulose biphosphate (RuBP) molecules, has been explored most extensively for carbon fixation in heterotrophs. In 2016, Milo and co-workers achieved partial success in the biosynthesis of sugars and another intermediate responsible for making biomass by expressing a fully functional CBB cycle in *E. coli*. They could only obtained semi-autotrophic growth due to dependency on the other carbon sources and organic compounds for reducing power and energy.[390] In this particular study, carbon fixation was performed via a CBB cycle in which pyruvate was used to produce reducing power and energy and xylose was used to support a productive CBB module. Recently, full autotrophy in *E. coli* was achieved by employing rational metabolic engineering and ALE. In the study, all biomass carbon was derived from CO_2, and energy and reducing power were supplied through formate.[391] In the process, RuBisCo/phosphoribulokinase enzymes were coexpressed with NAD^+ dependent formate dehydrogenase with additional genetic modification to enable CO_2 fixation and reduction via the CBB cycle. Also, successful autotrophy using CO_2 has been achieved in methylotrophic yeast *P. pastoris*, where native genes responsible for methanol assimilation were blocked by deleting the *DAS1*, *DAS2*, and *AOX1* genes to

https://doi.org/10.1021/acs.chemrev.2c00403
Chem. Rev. 2023, 123, 5521−5570

Table 5. List of Robustness Phenotypes Engineered through Metabolic Engineering

specific conditions	host	improvements/discoveries	metabolic engineering strategies	ref
		Thermotolerance		
ethanol fermentation at ≥40 °C	*S. cerevisiae*	• 1.91 ± 0.12 fold growth improvement	• ALE	92
		• 1.50 ± 0.2 fold glucose consumption	• GSM	
		• 1.6 ± 0.09 fold ethanol excretion	• MFA	
		• 1.3 ± 0.08 fold glycerol excretion	• transcriptomics	
			• metabolite quantification	
			• genome sequencing	
		Ethanol		
ethanol fermentations ≥120 g/L	*S. cerevisiae*	• 1.80 ± 1.3 fold ethanol titer/tolerance over wild type by adding K+ to media	• media optimization	406
		• 1.27 ± 2.2 fold ethanol titer/tolerance with K+ H+ transporter engineering	• transporter engineering	
			• overexpression	
			• metabolite quantification	
		Furfural and HMF		
20 mM furfural and 20 mM HMF	*S. cerevisiae*	• increased tolerance to furfural and 5-hydroxymethyl-2-furaldehyde (HMF)	• ALE	411−413
		• gene target identification	• metabolite quantification	
			• transcriptomics (microarray)	
		Thermotolerance		
fermentation at 40 °C	*S. cerevisiae*	• gene target identification	• identifying regulatory network targets	414
		• long-term thermotolerance regulatory network	• transcriptomics (RNA-seq)	
			• gene knockout	
		Thermotolerance		
fermentation at 40 °C	*E. coli*	• production of lysine was increased 5-fold at 40 °C	• quorum sensing and regulatory response for temperature control	415
		Low pH		
pH of 5.5	*Clostridium cellulovorans*, *Clostridium beijerinckii*	• 5-fold butanol titer improvement when compared against wild consortia	• Co-fermentation	416
			• heterologous gene expression	
			• gene knockout	
			• gene overexpression	
			• ALE	
		Butanol[440]		
(v/v): butanol 0.20%	*Synechocystis* sp. PCC 6803	• identification of gene targets for butanol tolerance engineering	• NGS transcriptomics	417
		• 3-phosphoglycerate, glycine, serine and urea related to stress response	• GC-MS metabolomics	
			• gene knockout	
up to 6% ethanol and 120 g/L glucose	*S. cerevisiae*	• glucose and ethanol tolerance	• gTME	29
		• 98% of ethanol theoretical yield achieve; 15% improvement from control	• gene knockout	
		• 3 mutations on SPT15 responsible for conferred tolerance	• gene overexpression	
			• transcriptomics (microarray)	
		Biofuel		
(v/v):	*E. coli*	• identified efflux pumps that could improve biofuel tolerance for: geranyl acetate, geraniol, α-pinene, limonene, farnesyl hexanoate	• competitive growth assays	86
• 2% geranyl acetate		• >50% limonene titer production in limonene production strain	• heterologous expression of efflux pumps	
• 0.05% geraniol				
• 2% a-pinene				
• 0.025% limonene				
• 2.5% farnesyl hexanoate				
		Short-Chain Alcohol		
(v/v):	*E. coli*	• 25% enhanced growth rate under	• directed evolution of efflux pump	418
		• improved growth for other	• whole gene random mutagenesis library	
• 0.7% *n*-butanol		• alcohols	• growth assay	
• 6% ethanol			• metabolite quantification	
• 0.8% isobutanol				
• 0.2% *n*-pentanol				

https://doi.org/10.1021/acs.chemrev.2c00403
Chem. Rev. 2023, 123, 5521−5570

Table 5. continued

specific conditions	host	improvements/discoveries	metabolic engineering strategies	ref
		Short-Chain Alcohol		
• 0.1% *n*-hexanol • 0.04% *n*-heptanol				

reduce the formaldehyde formation rate while generating reducing power and energy through methanol's dissimilatory pathway supported by AOX2.[35] Additionally, the enzymatic machinery involved in the CBB pathway was introduced into the peroxisome alongside methanol assimilation. The resulting strain grew continuously with CO_2 as a sole carbon source with the maximum specific growth rate of 0.008 h^{-1}, which was further improved to 0.018 h^{-1} by ALE. In a subsequent study, strains that were further evolved through ALE showed that mutations affected the activity and transcript levels of the genes that affect the availability of ATP and NADH.[392] In another report, to increase the rate of CO_2 fixation via the Calvin cycle, 20 enzymes involved in CO_2 concentrating mechanism (CCM) were overexpressed from *Halothiobacillus neapolitanus*, which enabled *E. coli* to grow by fixing CO_2 from the ambient air into biomass.[393] Compared to CBB, other native pathways have not been extensively explored to generate autotrophy in heterotrophs. Kondo and co-workers constructed a CO_2 and formate fixation pathway that converts two formate molecules and one CO_2 molecule to one pyruvate via glycine and L-serine in *E. coli*.[394]

3.3.3. Plastics as Substrate. Another alarming environmental concern is the growing accumulation of plastic wastes that are difficult to degrade. Many studies have developed innovative approaches for the disposal of plastic wastes in a more sustainable manner. Plastics are mainly polymers, like polyethylene (PE), polystyrene (PS), polypropylene (PP), polyvinyl chloride (PVC), polyurethane (PUR), and polyethylene terephthalate (PET). Natural degradation of these plastics takes thousands of years, which is not sustainable considering the rate of plastics production. However, natural microorganisms carrying microbial metabolic pathways for plastic depolymerization offer an alternative solution. Several species of *Pseudomonas* have been of particular interest due to their capabilities to degrade and metabolize synthetic plastics; however, the rate of plastic degradation and its use for making other materials are still limited.[395] Huang and co-workers degraded plastics using chemical technology called alkane metathesis, where long alkanes present in PE are covalently rearranged to give a new molecular distribution creating new products used in liquid fuels and waxes.[396] Recently, Wei and co-workers utilized an enzymatic degradation process where TfCut2 (i.e., thermophilic polyester hydrolases) from *Thermobifida fusca* expressed in *B. subtilis* was used for the treatment of PET-based food packaging containers and showed weight reduction of more than 50% after 96 h of incubation at 70 °C.[397] In 2020, characterization of *Ideonella sakaiensis* revealed the presence of a two-enzyme system PETase and MHETase, which was capable of deconstructing PET to terephthalate and ethylene glycol.[398] Chen and co-workers worked on hydrolysis of PET films using a thermostable polyester hydrolase, which yielded terephthalate and ethylene glycol. Later the obtained hydrolysate was used directly as a feedstock by the bacterium *Pseudomonas umsongensis* GO16 capable of degrading terephthalate.[399] Also, *P. umsongensis* was further evolved to efficiently metabolize ethylene glycol to produce polyhydroxyalkanoate (PHA). With metabolic engineering, synthetic plastics can be directly employed as a feedstock to alleviate the global challenge of plastic accumulation.

3.4. Robustness

The concept of engineering organism robustness originated in developing robust crops that are resistant to harsh conditions like drought,[400] and in the last few decades, more attention has been given to engineering robustness in microbial cell factories.[401] The use of lignocellulosic hydrolysates as a substrate has motivated much of this research as it contains a multitude of inhibitors including furans, phenols, and aliphatic acids.[373] Robustness covers a broad range of common phenotypes that can be characterized as maintenance of fitness under the pressure of a particular type of stress. Common stressors include temperature, pH, osmotic pressure, substrate, and product chemicals. Engineering robustness is an industrially relevant phenotype, as large-scale bioreactors create a harsh environment for cell proliferation. The most common strategy for engineering robustness is ALE due to its compatibility with growth related phenotypes and its ability to achieve tolerant mutants within a single experiment while also often evolving other beneficial phenotypes in the process.[94] Common rational approaches include engineering regulation factors, transporters, membranes, protection pathways, competition pathways, and balancing pathways.[402] Similar to growth phenotypes, tolerance is a polygenic trait and often draws on many different cellular subsystems making it a difficult rational engineering target.[403] Furthermore, tolerance to individual stressors does not often behave additively, ruling out most rational design strategies in the multi-inhibitor setting.[404] Robustness is typically studied in laboratory conditions at milliliter to liter scales in conditions that do not necessarily replicate in an industrial setting, making it difficult to bridge the gap between academic methods and industrial applications.[405] As there are numerous engineering strategies and different inhibitor environments, we primarily focus on temperature and alcohol stress and the diverse set of tools that have been used to overcome tolerance. Table 5 provides additional examples of tolerance engineering.

3.4.1. Temperature Tolerance. Nielsen and co-workers conducted a notable study related to robustness using a breadth of tools to engineer thermotolerant yeast strains (TTSs) through ALE.[92] Temperature tolerance is of general interest due to reduced contamination risk and higher operating temperature of industrial fermenters. Nielsen and co-workers grew three individual clonal populations of haploid *S. cerevisiae* at 39.5 ± 0.3 °C for more than 90 days, equating to more than 300 generations. Nine strains were selected from these three clonal populations, 3 from each, and 7 TTS behaved similarly at 40 ± 0.1 °C in fully aerobic conditions growing an average of 1.91 ± 0.12 times faster, consuming glucose on average 1.50 ± 0.2 times faster and excreting ethanol and glycerol on average 1.6 ± 0.09 and 1.3 ± 0.08 times faster respectively compared to the parent strain. Genome sequencing and whole-genome transcriptome profil-

https://doi.org/10.1021/acs.chemrev.2c00403
Chem. Rev. 2023, 123, 5521−5570

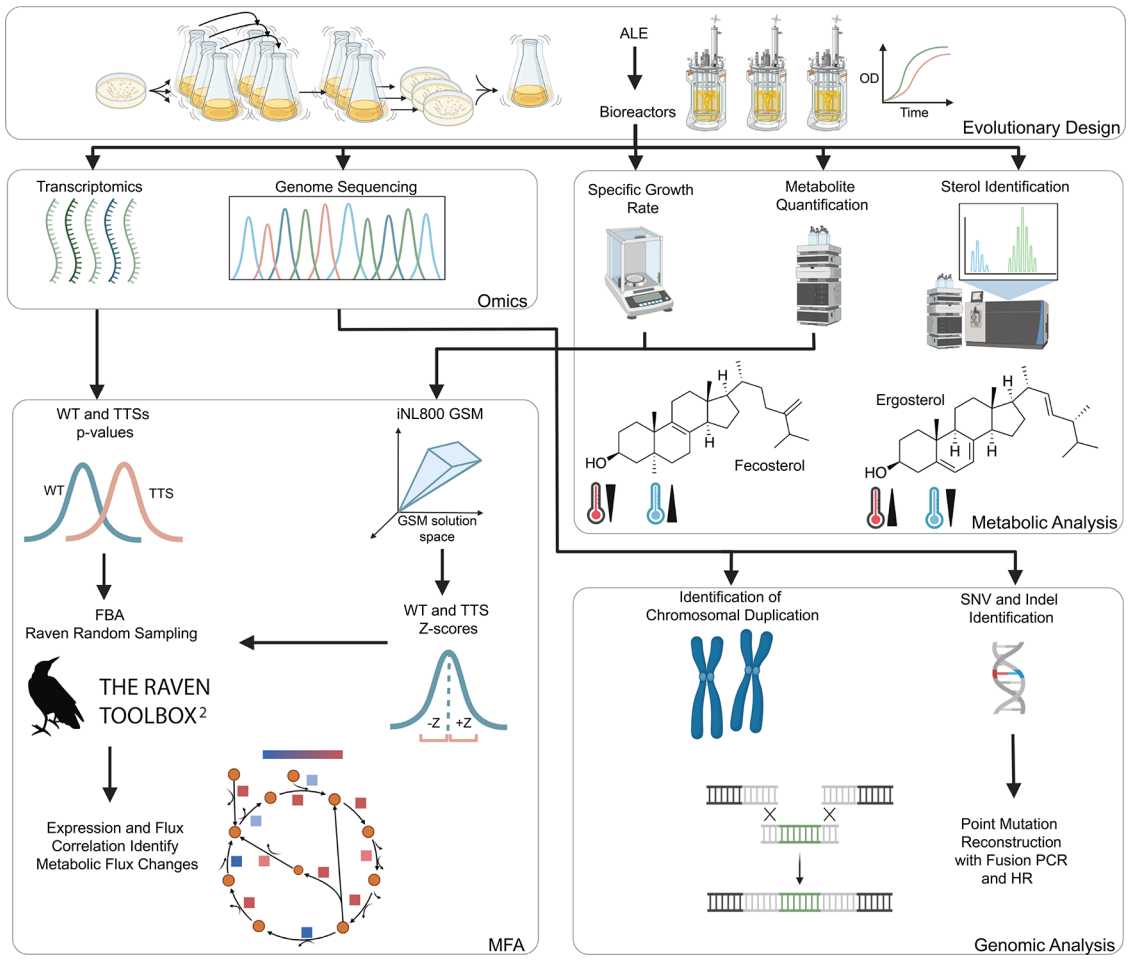


Figure 18. Robustness case study on thermotolerance. After evolutionary design with ALE, subsequent analysis including omics, MFA, and genomic analysis helped deduce that increased tolerance at 40 °C was associated with increased concentration of fecosterol as opposed to ergosterol.

ing were performed on the strains. Genome sequencing identified chromosomal duplication in the evolved strain and single nucleotide variants (SNVs) and indels responsible for improved thermotolerance. Specifically, duplication of genes on a region ChrIII previously reported for thermotolerance was reidentified. To verify the causal effect of SNVs on thermotolerance, point mutation reconstruction with fusion PCR was performed on genes *ATP3* and *ERG3*, and it was found that *ATP3* mutations negatively affect thermotolerance, but *ERG3* mutations conferred up to 86% of the specific growth rate of the evolved TTSs. The mutations in *ERG3* have been reported to augment the activity of sterol methyltransferase Erg6, causing higher production of fecosterol, or "bent" sterol as opposed to "flat sterols," like ergosterol, which have been previously reported to have no effect on thermotolerance, leaving researchers to hypothesize that "bent" sterols could be responsible for improving thermotolerance. Whole-genome transcriptome profiling was performed on yeast during exponential growth in bioreactors and the moderated t-statistic was applied to identify pairwise differences between gene expression in each TTS and the parental strain. External fluxes from HPLC, including glucose and ethanol and specific growth, were used to constrain GEM iIN800 and calculate metabolic fluxes. Transcriptome profiles and calculated metabolic fluxes were used with the FBA RAVEN random sampling algorithm to identify probable changes in metabolic flux. From the MFA results, TTSs were identified to have residual TCA activity which in combination with TTSs having ~30% increased oxygen uptake explains the oxidation of cytosolic NADH generated by increased biomass yield. Additionally, gene transcription and metabolic flux were found to be upregulated in the mevalonate pathway and in sterol biosynthesis in select TTSs when compared to parental strains. This along with the knowledge that sterols are responsible for membrane fluidity allowed the researchers to hypothesize that "bent" sterols could be optimal for membrane fluidity at higher temperatures (Figure 18).[92]

3.4.2. Alcohol Tolerance. Model yeast *S. cerevisiae* is the most common microbial ethanol producer, but ethanol in high concentrations is toxic to yeast cells because it permeates cell membranes and allows ion leakage disturbing carefully controlled ion gradients. Stephanopoulos and co-workers hypothesized that supplementing the growth media could help stabilize the plasma membrane, so they added different salts into high gravity, or high sugar, media which mimics the osmotic stress found in industrial bioreactors. The addition of 40 mM KCl and 10 mM KOH to media with wild-type *S. cerevisiae* brought the strain up to industrial production levels by increasing intracellular K^+ ions and increasing pH. The addition of ions alone improved titers by 30 ± 1.2%. Media supplementation similarly improved cell viability in the presence of common antiseptic 2-propanol and alternative

https://doi.org/10.1021/acs.chemrev.2c00403
Chem. Rev. 2023, 123, 5521−5570

biofuel isobutanol, suggesting that ion composition affects higher-order cellular processes that provide tolerance to small alcohols and other membrane permeating molecules in general. The media supplementation also improved ethanol fermentation on xylose for a modified xylose assimilation strain, improving titer by 54 ± 5.7%, showing that improvements have little to do with genetic background. Next, the ATP pumps including K^+ importer TRK1 and H^+ exporter PMA1 were rationally engineered by deleting PPZ1 and PPZ2 to activate TRK1 and upregulate PMA1 to help maintain the ion gradient across the plasma membrane. These changes resulted in 27 ± 2.2% and ∼6% titer improvement from the wild-type strain and the Brazilian ethanol production strain PE-2, respectively. Modification of the ion pumps alone improved cell tolerance, but the largest improvements came from KCl and KOH supplementation, which suggests that certain types of tolerance are primarily driven by physical instead of genetic determinants.[406] While altering the environment is a feasible solution to engineering robustness, metabolic engineering efforts typically focus on genetic interventions.

It is known that adaptation to different environments through a few cellular generations is governed by the regulatory network and not protein mutations,[407] and because of this genes can be identified in different stress environments associated with key transcription factors. The gene regulatory network is constructed as a hierarchy with global transcription factors sitting on top of the hierarchy, making it easy to affect system-wide changes with few gene modifications.[408] Using this idea, at a time when it was still difficult to introduce multiple simultaneous gene modifications, Stephanopoulos and co-workers developed the tool of global transcription machinery engineering (gTME), which used random mutagenesis to mutate key residues in the main sigma factor, σ^{70} which in turn altered RNA polymerase promoter preferences. This tool was used to engineer ethanol tolerance, lycopene production, and a multiphenotype tolerance to ethanol and sodium dodecyl sulfate.[409] In another effort to engineer alcohol tolerance, Papoutsakis and co-workers engineered *E. coli* for improved tolerance against *i*-, *n*-, and 2-butanol, 1,2,4-butanetriol, and ethanol and other alcohols. Heat shock proteins have been previously associated with robustness in $Gram^+$ bacteria due to their ability to maintain normal protein folding and protein transport, but their ability to confer solvent tolerance in $Gram^-$ bacteria like *E. coli* was still unknown. Overexpression of the two-part chaperonin system GroESL was overexpressed and consistently increased cell viability for all studied chemicals.[410]

4. FUTURE PERSPECTIVES

Development of microbial cell factories to produce value-added chemicals has been aided by the rapid progress and remarkable accomplishments in the field of metabolic engineering. Nevertheless, due to the lack of fundamental knowledge, multiple rounds of the DBTL cycle are necessary to achieve efficient cell factories, and several obstacles remain. Here we discuss the future directions in metabolic engineering, including engineering of nonmodel organisms, development of biofoundries and AI/ML tools, utilization of nonsugar substrates, engineering of microbial consortia, and scale-up of fermentation processes. We focus on the existing barriers and the possible solutions and future developments that could potentially address such challenges.

4.1. Non-model Organisms

E. coli and *S. cerevisiae* are considered as the traditional workhorses of metabolic engineering because of their well-understood physiologies and because there are abundant genetic engineering tools. Nevertheless, *S. cerevisiae* and *E. coli* lack some desired characteristics, such as tolerance to harsh conditions, including low pH or high oxidative stress, the ability to utilize simple or complex carbon sources, or the capability to accumulate target products or precursors of target molecules. Thus, non-model microorganisms possessing these unique traits are being increasingly explored for production of chemicals. For example, *Issatchenkia orientalis* is renowned for its superior tolerance to highly acidic conditions, which makes it a potential host for the economic production of organic acids, and *R. toruloides* is an oleaginous yeast capable of accumulating lipids up to 70% of its dry cellular weight and is thus a preferred host for the production of lipid-based chemicals, such as oleochemicals and diesel-like fuels.[142,419,420] Moreover, cyanobacteria can perform oxygenic photosynthesis and carbon fixation, enabling production of chemicals from carbon dioxide, water, and light. Filamentous fungi like *Aspergillus niger* and *Fusarium fujikuroi* can naturally secrete organic acids and plant growth hormones, respectively, making them attractive hosts for overproduction of these compounds.[421−423]

Although non-model organisms are increasingly explored in metabolic engineering, genetic engineering in non-conventional microorganisms is still hindered by several challenges. In *E. coli* and *S. cerevisiae*, gene expressions using episomal plasmids are widely used due to the availability of stable and varying copy number vectors. Nevertheless, in non-model organisms, gene expression is often accomplished through chromosomal integration because of the lack of stable episomal plasmids. While episomal plasmids harboring autonomously replicating sequences are available for some non-model organisms, they tend to be unstable and have low copy numbers. Moreover, plasmids lacking centromeric sequences exhibit variable expression across cells in a single population due to imperfect partitioning of plasmids upon cell division. Some non-model organisms also suffer from low transformation efficiencies making genome-scale engineering methods, which require high transformation efficiency, challenging in these species. Furthermore, metabolic engineering in non-conventional microorganisms is limited by the incomplete understanding of their metabolism, genetics, and physiology. Thus, to harness the full potential of non-model microorganisms, new systems biology and synthetic biology tools need to be developed to increase the genetic toolbox for rapid and efficient strain engineering and to expand our understanding of their unique biochemistry and native physiology.

4.2. Biofoundry Development

Metabolic engineering of biological systems such as proteins, pathways, genetic circuits, and genomes involves the repetitive execution of the DBTL cycle to achieve the desirable engineering objective. However, manual operations are prone to human errors when numerous iterations are required, which reduces the experimental consistency and objectiveness. Moreover, the number of tasks that can be explored in parallel in each cycle is limited due to low throughput and turnover rate.[365,366] Biofoundries provide an integrated infrastructure for rapid design, construction, and characterization of synthetic

https://doi.org/10.1021/acs.chemrev.2c00403
Chem. Rev. 2023, 123, 5521−5570

biosystems through computer-aided design (CAD) software, design-of-experiments (DOE), liquid handlers, and high-throughput screening equipment.[424−426] One aspirational goal of biofoundries is to accelerate and enhance both academic and translational research in engineering/synthetic biology by promoting and enabling the beneficial use of lab automation. Over the past 5 years, multiple biofoundries have been built to expand the biotechnology development capacities. The Edinburgh Genome Foundry can process over 2,000 DNA assembly reactions per week.[427] The Illinois Biological Foundry for Advanced Biomanufacturing (iBioFAB) can automatically generate 400 pairs of TALENs on a daily basis.[428] Furthermore, iBioFAB has been able to perform multiplexed genome-scale engineering of *S. cerevisiae*[429] and construct any plasmidin a high-throughput manner.[430] Other biofoundries such as the London DNA Foundry,[431] Singapore SynCTI Foundry,[432] and US DOE Agile BioFoundry[433] have been developing similar automated workflows to benefit scientists across various research fields as well. In order to share experiences and resources and work together to overcome common challenges and unmet scientific and engineering needs, the Global Biofoundry Alliance (GBA) was launched in May 2019.[426]

Within this context, we believe that one of the next waves of innovation in metabolic engineering would be a fully automated DBTL cycle with reduced turnaround time. A flexible, versatile, and data-driven biofoundry framework for metabolic engineering would become possible with improved laboratory automation, data and protocols standardization, and better integration of the wet-lab and cloud-based operations.[424−426,434] To develop such a biofoundry, there are three fundamental requirements: (1) combining rational models and ML tools for Design and Learn, (2) investigating solutions to efficiently use cloud repositories and exchange workflows among biofoundries for Build, and (3) exploiting the potential of real-time high-resolution omics data analysis for Test. Currently, lab robotics and integrated software are mostly suited for automating standard workflows where constructed mutants can be separated and characterized via fluorescence-based protein tags or an NGS-based approach.[435−437] However, analysis of the small molecules and bioproducts produced by engineered microorganisms often includes quantification and comparison of TRY. Because most of these molecules are difficult to quantify and whole-genome sequencing of strains is expensive and time-consuming, there remains a significant gap for automatically screening synthesized compounds in a high-throughput manner. Droplet-based microfluidics has emerged as an attractive platform as the integration of automated sample preparation and mass spectrometry analysis becomes possible.[438−442] Online real-time detection and analysis of produced molecules with limited human interventions would be ideal in the future. With the recent achievements in ML and artificial intelligence (AI)-based prediction of protein structures and functionsof interest, standardized computational tools and exchange of information among pipelines could be reached.[443] We envision iterative rounds of the DBTL cycle being fully automated, with the robot making predictions and decisions about the gene targets to choose in future experimental rounds until the desired microorganisms are constructed.[266,444] A data-driven closed-loop metabolic engineering biofoundry would solve the challenges of labor-intensive strain library construction as well as library screening and characterization.

4.3. AI/ML Tools

ML in metabolic engineering has demonstrated its effectiveness in improving production in a reduced number of DBTL iterations and in analyzing data to predict new biological interactions or characterize component parts. For example, Bioautomata[266] improved lycopene production by 77% compared to random screening in just 3 rounds, and the automated recommendation tool (ART)[272] demonstrated improved design predictions for fatty acids and tryptophan.[271] A recent example has used sequence information and cell sorting to characterize all of the promoters in *S. cerevisiae*, creating a model that in principle could allow for promoter design in an engineered pathway.[273] Major limitations of ML specific to metabolic engineering include integration of different data modalities and interpretability.

Integrating data modalities is a primary concern because it is now feasible to collect data at every level of cell processing. Traditionally, specific omics layers have been studied in isolation, or with small amounts of additional data. Because all omics layers are interconnected, it is difficult to know a priori which elements across the different modalities are important to measure. Modeling multiple data modalities will allow for the identification of important covariates across layers to help identify gene targets for engineering that would be otherwise undetected from a single modality. The inclusion of output from mechanistic models as an additional data modality has also received some attention with some success.[446,447] Lastly, there is an abundance of unused data in the literature that has not been codified into databases due to the difficulty in extracting, verifying, and unifying experimental information. While ML has shown use in predicting synthesis parameters in material sciences,[445] mining of literature for data curation useful for metabolic engineering remains an open question, with current methods only able to capture general metabolic trends.[282] It remains to be seen whether it is possible to use ML for the extraction and curation of high-quality biological data.

A fully predictive mechanistic model of cellular metabolism is ideal for design, as it allows for reasoning about predictions. Constructing interpretable ML models, sometimes called white box or glass box models, would allow for refinement of biological knowledge and subsequent elaboration of more accurate mechanistic models. Without interpretability of salient features, we risk collecting an abundant amount of costly data that provides little to no improvement in biological understanding. The lack of interpretability in ML models is in opposition to the traditional mechanistic modeling paradigm. In the last five years, there has been a strong effort in taking advantage of the knowledge-based representation of the cell from reaction networks to ontologies to build ML models with more interpretable architectures. One example uses the hierarchical representation of the gene ontology graph as the model architecture to build a visible neural network. While this model's ability is not directly useful for the prediction of target product TRY, it demonstrates that cell growth phenotypes can be predicted from gene knockouts, and it provides explanations by allowing the affected cellular compartments to be inspected for meaningful changes.[279] Metabolic representation of the cell in GEMs and in ML models do not currently share a unified representation. There have been attempts to combine them,[446] but typically the output of a GEM is used as input to a ML model or vice versa.[447] In an attempt to further merge this gap, Faulon and co-workers developed an artificial metabolic

https://doi.org/10.1021/acs.chemrev.2c00403
Chem. Rev. 2023, 123, 5521−5570

network (AMN) that allows for the backpropagation of errors within the metabolic network, blending important elements of mechanistic and ML-based models. The AMN model was first trained on simulated FBA data sets and then used to improve predictions of MFA on growth rate data sets in *E. coli*. This construction allowed for the uncovering of regulation between growth media and the steady-state metabolic phenotype.[448] GEMs and kinetic models have become more popular as engineering tools, but their predictive performance is still insufficient. We expect to see an increased effort in building new hybrid models that combine mechanistic models and ML models to achieve both high predictive performance and interpretability.

4.4. Nonsugar Substrates

For most metabolic engineering studies, glucose has been used as the substrate for microbial production of molecules at laboratory and industrial scales. Nevertheless, it is not considered a sustainable source for large-scale fermentations. To address this limitation, attention has shifted to the utilization of other sustainable carbon sources such as lignocellulosic biomass and glycerol derived from petrochemical industries. However, mixed sugars and inhibitors present in the lignocellulosic hydrolysate often expose cells to carbon catabolite repression issues, which can lower cell density and reduce product yields. Fortunately, "Y-shaped" microbial consortium reactors with a shared biosynthetic pathway and distinctive sugar catabolic pathways, for example glucose and xylose, can be used in simultaneous cofermentation to overcome microbial substrate preferences.[449] This strategy could be implemented for the coutilization of other substrates.

Other nonsugar, inexpensive substrates are organic industrial waste and greenhouse gases, which have been shown to be a great alternative to feed microorganisms. LanzaTech has been intensively engineering acetogenic bacteria to produce more than 50 chemicals from gaseous species at the lab level and a few molecules, such as ethanol, acetone, and 2-propanol, at the pilot scale. One of the challenges with acetogens is cellular energy limitation, which affects carbon assimilation. Although acetogens have a native ability to utilize the C1 substrate and low value-added chemicals, an extensive exploration of this bacteria is still hampered by the scarcity of advanced genetic tools. Researchers are still focusing on engineering autotrophic model organisms, and they recently made the first autotrophic *E. coli*, which was still not comparable to natural C1 utilizing hosts. Metabolizing CO_2 also requires an additional energy source that can be provided by light, H_2, CO, and organic or inorganic compounds. We expect further advancement of genetic tools related to non-model organisms will have the soonest impact on C1 utilization.

Apart from the mentioned carbon substrates, marine macroalgae including red, green, and brown macroalgae have huge potential to serve as feedstocks for the production of a broad range of chemicals.[450] Macroalgae production does not compete with the production of land-based food crops, as it does not require arable land, freshwater, or fertilizer. These macroalgae contain a unique composition of carbohydrates and are shown to produce a variety of chemicals including ethanol, β-carotene, lactate, acetoin, and BDO. However, these macroalgae can be further rewired to achieve a particular carbohydrate composition. The ability to utilize different carbon sources by the native or engineered host can reduce the dependency on one substrate and make them available sustainably.

4.5. Microbial Consortia

A barrier when using genetically engineered microorganisms that carry heterologous genes is the increase in metabolic burden from the activity of the corresponding enzymes.[451] This can decrease the overall productivity of the microorganisms toward target molecule production. To counter the low metabolic efficiencies in such scenarios, division of labor of complex metabolic pathways among multiple microorganisms has been considered as a viable alternative. The past decade has seen an increased focus on the application of communities of multiple microorganisms, also known as microbial consortia, to produce target molecules. The artificial or synthetic microbial consortia were inspired by naturally occurring microbial consortia, such as the gut microbiome.[452] Besides the division of labor leading to a distributed metabolic burden, microbial consortia offer additional advantages in biochemical production. The compartmentalized nature of metabolic pathways for production results in a modular organization. The presence of symbiotic or coexisting microorganisms also reduces the chances of contamination compared to a monoculture.

So far, applications of microbial consortia in biochemical production have seen very little work on engineering the individual microorganisms of the consortium. Due to the lack of effective predictive tools at the community level, the design space of microbial communities for biochemical production consists of numerous unknowns. The most common design decisions involve tuning the ratios of different cell populations, geometry of the bioreactors, and growth conditions suitable for all the chosen microorganisms. Recent work on the use of methods such as FBA in a microbial community setting has the potential to inform the engineering of individual microorganisms in a community or consortium context.[453] Variations of such algorithms have even enabled the time-based prediction of community interactions and may play a role in the future in designing dynamic control of target molecule production in a microbial consortium.[454] One particular method, SteadyCom, is capable of predicting steady-state compositions of microbial communities.[455] This may prove particularly useful in systematically tailoring the initial ratios of the different microorganisms used in a synthetic consortium to achieve a desired microbial composition, which previously relied on intensive trial-and-error experiments.

The modeling of microbiome community dynamics has received greater attention in the context of the human gut microbiome.[456] Efforts in building community models can aid in identifying metabolite exchanges among gut microorganisms. The use of GEMs along with constraint-based analysis has been used to inform hypotheses on metabolic interactions within gut microbiomes. Such modeling efforts have also been used to infer rules about the organization of the gut microbiome at a community level.[457] In the future, this accrued knowledge generated in the field of human gut microbiome modeling can be translated to applications in metabolic engineering of microbial consortia as well. Community models can be applied to engineer microbial species that work in tandem within a consortium for efficient conversion of substrates to a target molecule.

https://doi.org/10.1021/acs.chemrev.2c00403
Chem. Rev. 2023, 123, 5521−5570

4.6. Scale-up Fermentations

The goal of metabolic engineering is to develop efficient processes for production of chemicals and fuels from renewable biomass at large scales. Semisynthetic artemisinin was the first pharmaceutical agent that was produced biologically at an industrial scale.[322] Genomatica has established a commercial route for the conversion of glucose into 1,4-BDO. The process was successfully conducted in over 50 runs to produce over 4,000 tons of 1,4-BDO at a commercial scale.[295] Recently, LanzaTech developed an industrial carbon-negative fermentation process, which was discussed in the 2-propanol and acetone case study.[297]

Despite the outstanding accomplishments in metabolic engineering over the past 30 years, there have been very few successful transitions from lab-scale fermentations into commercial scale processes and marketed products. The main technical bottleneck with scale-up is lack of access to facilities for pilot and large-scale productions. Even if entry to a manufacturing site is possible, performing experiments at scale to optimize the fermentation conditions is often infeasible due to cost. Furthermore, because the environment of a large-scale bioreactor is vastly different from that of small-scale cultures such as shake flasks, it is expected that most strains do not perform similarly at these conditions. Several scale-dependent physical, chemical, and biological parameters can hinder microbial growth and product formation in commercial bioreactors.[458] To facilitate scale-up processing, it is necessary to perform lab-scale experiments under conditions imitating the intended large-scale production conditions.[459] Computational tools to simulate large-scale conditions in smaller-scale reactors can also be utilized.

5. CONCLUSIONS

Metabolic engineering has come a long way from its early days of manipulating a few genes for production of simple molecules to being successfully employed for efficient production of chemicals and fuels from a diverse range of carbon sources. Furthermore, motivated by concerns over energy and sustainability, metabolic engineering is currently regarded as a sustainable alternative to conventional petroleum-based production processes. Advances in computational tools, synthetic biology tools, and analytical tools have further facilitated metabolic engineering endeavors, including the development of efficient microbial cell factories and the expansion of the range of substrates that can be utilized and the products that can be produced. Despite all of these achievements, lack of fundamental knowledge on cellular metabolism and physiology remains a major roadblock to the advancement of the field. Therefore, it is necessary to improve the understanding of biological systems to realize the full potential of metabolic engineering.

AUTHOR INFORMATION

Corresponding Author

Huimin Zhao − *Department of Chemical and Biomolecular Engineering, Carl R. Woese Institute for Genomic Biology, DOE Center for Advanced Bioenergy and Bioproducts Innovation, and Department of Chemistry, University of Illinois at Urbana−Champaign, Urbana, Illinois 61801, United States;* orcid.org/0000-0002-9069-6739; Phone: (217) 333-2631; Email: zhao5@illinois.edu; Fax: (217) 333-5052

Authors

Michael J. Volk − *Department of Chemical and Biomolecular Engineering, Carl R. Woese Institute for Genomic Biology, and DOE Center for Advanced Bioenergy and Bioproducts Innovation, University of Illinois at Urbana−Champaign, Urbana, Illinois 61801, United States*

Vinh G. Tran − *Department of Chemical and Biomolecular Engineering, Carl R. Woese Institute for Genomic Biology, and DOE Center for Advanced Bioenergy and Bioproducts Innovation, University of Illinois at Urbana−Champaign, Urbana, Illinois 61801, United States*

Shih-I Tan − *Department of Chemical and Biomolecular Engineering and DOE Center for Advanced Bioenergy and Bioproducts Innovation, University of Illinois at Urbana−Champaign, Urbana, Illinois 61801, United States; Department of Chemical Engineering, National Cheng Kung University, Tainan 70101, Taiwan*

Shekhar Mishra − *Department of Chemical and Biomolecular Engineering, Carl R. Woese Institute for Genomic Biology, and DOE Center for Advanced Bioenergy and Bioproducts Innovation, University of Illinois at Urbana−Champaign, Urbana, Illinois 61801, United States*

Zia Fatma − *Department of Chemical and Biomolecular Engineering and DOE Center for Advanced Bioenergy and Bioproducts Innovation, University of Illinois at Urbana−Champaign, Urbana, Illinois 61801, United States*

Aashutosh Boob − *Department of Chemical and Biomolecular Engineering, Carl R. Woese Institute for Genomic Biology, and DOE Center for Advanced Bioenergy and Bioproducts Innovation, University of Illinois at Urbana−Champaign, Urbana, Illinois 61801, United States*

Hongxiang Li − *Department of Chemical and Biomolecular Engineering, Carl R. Woese Institute for Genomic Biology, and Department of Chemistry, University of Illinois at Urbana−Champaign, Urbana, Illinois 61801, United States*

Pu Xue − *Department of Chemical and Biomolecular Engineering, Carl R. Woese Institute for Genomic Biology, and DOE Center for Advanced Bioenergy and Bioproducts Innovation, University of Illinois at Urbana−Champaign, Urbana, Illinois 61801, United States;* orcid.org/0000-0001-6023-1298

Teresa A. Martin − *Department of Chemical and Biomolecular Engineering, Carl R. Woese Institute for Genomic Biology, and DOE Center for Advanced Bioenergy and Bioproducts Innovation, University of Illinois at Urbana−Champaign, Urbana, Illinois 61801, United States*

Complete contact information is available at: https://pubs.acs.org/10.1021/acs.chemrev.2c00403

Author Contributions

#M.J.V., V.G.T., and S.I.T. contributed equally

Notes

The authors declare no competing financial interest.

Biographies

Michael J. Volk graduated from the University of Arizona with B.S. degrees in Chemical Engineering and Mathematics in 2018. He is pursuing a Ph.D. degree at UIUC under Dr. Huimin Zhao. His thesis research focuses on applying machine learning strategies to systems biology and metabolic engineering.

https://doi.org/10.1021/acs.chemrev.2c00403
Chem. Rev. 2023, 123, 5521−5570

Vinh G. Tran earned his B.S. degree in Chemical Engineering from the University of California, Berkeley, in 2016. He is currently pursuing a Ph.D. at UIUC under the tutelage of Dr. Huimin Zhao, and his research focuses on exploring *Issatchenkia orientalis*' potential as a platform organism for organic acid production.

Shih-I Tan received his B.S. degree in Chemical Engineering from National Cheng Kung University (NCKU), Taiwan, in 2016. He pursued his Ph.D. degree from NCKU with Prof. I-Son Ng. Currently, he is a visiting scholar in Prof. Huimin Zhao's laboratory, and his research focuses on developing genetic tools and genome engineering tools to advance metabolic engineering of non-model yeasts for chemical production.

Shekhar Mishra received his B.Tech and M.Tech from the Indian Institute of Technology, Bombay, in 2014. He is currently pursuing his Ph.D. degree at UIUC under Dr. Huimin Zhao. His thesis research focuses on the development of computational tools to aid in the design of improved metabolic engineering strategies in yeast.

Zia Fatma received her Ph.D. degree in Biotechnology from the International Centre for Genetic Engineering and Biotechnology, Delhi, (ICGEB) in 2016. She is now a postdoctoral fellow in the Department of Energy Center for Advanced Bioenergy and Bioproducts Innovation (CABBI) at UIUC. Her research focuses on genetic tool development and metabolic engineering of non-conventional yeasts to produce value-added chemicals.

Aashutosh Boob received his B.Tech with Honours degree from the Indian Institute of Technology, Bombay in 2019. He is currently pursuing his Ph.D. degree at UIUC, where his thesis research focuses on implementing machine learning strategies to design synthetic biology toolkits for enhanced biochemical production.

Hongxiang Li received her B.S. degree in Chemistry from China Agricultural University in 2021. She is currently pursuing her Ph.D. degree at UIUC with Dr. Huimin Zhao. Her research focuses on retrosynthesis planning.

Pu Xue (Mason) graduated from the Rensselaer Polytechnic Institute (RPI) with a B.S. degree in Chemical and Biomolecular Engineering and joined the graduate program in Chemical and Biomolecular Engineering at UIUC in 2016. His Ph.D. thesis research focuses on developing a fully automated workflow to speed up the design–build–test–learn cycle for engineering of microbial cell factories. He has applied multiple synthetic biology and directed evolution concepts/tools for high-throughput library construction, and screening during his Ph.D. thesis research.

Teresa A. Martin received her Ph.D. in Chemical Biology from the University of Illinois at Urbana–Champaign in 2011. She joined the lab of Dr. Huimin Zhao in 2016 as the Coordinator of Research Programs. Her research focuses on the metabolic engineering of non-conventional yeasts to produce value-added chemicals.

Dr. Huimin Zhao is the Steven L. Miller Chair of Chemical and Biomolecular Engineering, and Professor of Chemistry, Biochemistry, and Bioengineering at UIUC. He received his B.S. degree in Biology from the University of Science and Technology of China in 1992 and his Ph.D. degree in Chemistry from the California Institute of Technology in 1998 under the guidance of Dr. Frances Arnold. He joined UIUC in 2000 after a stint as a project leader at The Dow Chemical Company and was promoted to full professor in 2008. Dr. Zhao has authored and coauthored over 380 research articles and over 30 issued and pending patent applications. In addition, he has given over 450 plenary, keynote, or invited lectures. His primary research interests are in the development and applications of directed evolution, synthetic biology, machine learning, and laboratory automation tools to address society's most daunting challenges in health, energy, and sustainability, and in the fundamental aspects of enzyme catalysis, cell metabolism, and gene regulation.

ACKNOWLEDGMENTS

We thank financial support from the U.S. Department of Energy (DE-SC0018420 and DE-SC0018260) for various metabolic engineering projects (H.Z.). The online tool of BioRender (BioRender.com) was used to create parts of the TOC Figures 2–18.

REFERENCES

(1) Cohen, S. N.; Chang, A. C. Y.; Boyer, H. W.; Helling, R. B. Construction of Biologically Functional Bacterial Plasmids In Vitro. *Proc. Natl. Acad. Sci. U. S. A.* **1973**, *70*, 3240–3244.
(2) Chang, C. N.; Key, M.; Bochner, B.; Heyneker, H.; Gray, G. High-Level Secretion of Human Growth Hormone by *Escherichia coli*. *Gene* **1987**, *55*, 189–196.
(3) Bräu, B.; Sahm, H. Cloning and Expression of the Structural Gene for Pyruvate Decarboxylase of *Zymomonas mobilis* in *Escherichia coli*. *Arch. Microbiol.* **1986**, *144*, 296–301.
(4) Cameron, D. C.; Tong, I.-T. Cellular and Metabolic Engineering. *Appl. Biochem. Biotechnol.* **1993**, *38*, 105.
(5) Farmer, W. R.; Liao, J. C. Progress in Metabolic Engineering. *Curr. Opin. Biotechnol.* **1996**, *7*, 198–204.
(6) Bailey, J. E. Toward a Science of Metabolic Engineering. *Science* **1991**, *252*, 1668–1675.
(7) Vemuri, G. N.; Aristidou, A. A. Metabolic Engineering in the -Omics Era: Elucidating and Modulating Regulatory Networks. *Microbiol. Mol. Biol. Rev.* **2005**, *69*, 197–216.
(8) Borodina, I.; Nielsen, J. Advances in Metabolic Engineering of Yeast *Saccharomyces cerevisiae* for Production of Chemicals. *Biotechnol. J.* **2014**, *9*, 609–620.
(9) Lee, S. Y.; Kim, H. U.; Chae, T. U.; Cho, J. S.; Kim, J. W.; Shin, J. H.; Kim, D. I.; Ko, Y.-S.; Jang, W. D.; Jang, Y.-S. A Comprehensive Metabolic Map for Production of Bio-Based Chemicals. *Nat. Catal.* **2019**, *2*, 18–33.
(10) Clomburg, J. M.; Crumbley, A. M.; Gonzalez, R. Industrial Biomanufacturing: The Future of Chemical Production. *Science* **2017**, *355*, aag0804.
(11) Lee, J. W.; Na, D.; Park, J. M.; Lee, J.; Choi, S.; Lee, S. Y. Systems Metabolic Engineering of Microorganisms for Natural and Non-Natural Chemicals. *Nat. Chem. Biol.* **2012**, *8*, 536–546.
(12) Ko, Y.-S.; Kim, J. W.; Lee, J. A.; Han, T.; Kim, G. B.; Park, J. E.; Lee, S. Y. Tools and Strategies of Systems Metabolic Engineering for the Development of Microbial Cell Factories for Chemical Production. *Chem. Soc. Rev.* **2020**, *49*, 4615–4636.
(13) Galanie, S.; Thodey, K.; Trenchard, I. J.; Filsinger Interrante, M.; Smolke, C. D. Complete Biosynthesis of Opioids in Yeast. *Science* **2015**, *349*, 1095–1100.
(14) Sharma, B.; Shukla, P. Futuristic Avenues of Metabolic Engineering Techniques in Bioremediation. *Biotechnol. Appl. Biochem.* **2022**, *69*, 51–60.
(15) Wang, F.; Raab, R. M.; Washabaugh, M. W.; Buckland, B. C. Gene Therapy and Metabolic Engineering. *Metab. Eng.* **2000**, *2*, 126–139.
(16) Khosla, C.; Keasling, J. D. Metabolic Engineering for Drug Discovery and Development. *Nat. Rev. Drug Discovery* **2003**, *2*, 1019–1025.
(17) He, F.; Murabito, E.; Westerhoff, H. V. Synthetic Biology and Regulatory Networks: Where Metabolic Systems Biology Meets Control Engineering. *J. R. Soc. Interface* **2016**, *13*, 20151046.
(18) Ikeda, M.; Katsumata, R. Metabolic Engineering To Produce Tyrosine or Phenylalanine in a Tryptophan-Producing *Corynebacterium glutamicum* Strain. *Appl. Environ. Microbiol.* **1992**, *58*, 781–785.

https://doi.org/10.1021/acs.chemrev.2c00403
Chem. Rev. 2023, 123, 5521–5570

(19) Katsumata, R.; Ikeda, M. Hyperproduction of Tryptophan in *Corynebacterium glutamicum* by Pathway Engineering. *Bio/Technology* **1993**, *11*, 921−925.
(20) Zhang, M.; Eddy, C.; Deanda, K.; Finkelstein, M.; Picataggio, S. Metabolic Engineering of a Pentose Metabolism Pathway in Ethanologenic *Zymomonas mobilis*. *Science* **1995**, *267*, 240−243.
(21) Stephanopoulos, G.; Vallino, J. J. Network Rigidity and Metabolic Engineering in Metabolite Overproduction. *Science* **1991**, *252*, 1675−1681.
(22) Burgard, A. P.; Pharkya, P.; Maranas, C. D. Optknock: A Bilevel Programming Framework for Identifying Gene Knockout Strategies for Microbial Strain Optimization. *Biotechnol. Bioeng.* **2003**, *84*, 647−657.
(23) Förster, J.; Famili, I.; Fu, P.; Palsson, B. Ø.; Nielsen, J. Genome-Scale Reconstruction of the *Saccharomyces cerevisiae* Metabolic Network. *Genome Res.* **2003**, *13*, 244−253.
(24) Marx, A.; de Graaf, A. A.; Wiechert, W.; Eggeling, L.; Sahm, H. Determination of the Fluxes in the Central Metabolism of *Corynebacterium glutamicum* by Nuclear Magnetic Resonance Spectroscopy Combined with Metabolite Balancing. *Biotechnol. Bioeng.* **1996**, *49*, 111−129.
(25) Varma, A.; Palsson, B. O. Metabolic Flux Balancing: Basic Concepts, Scientific and Practical Use. *Bio/Technology* **1994**, *12*, 994−998.
(26) Edwards, J. S.; Palsson, B. O. The *Escherichia coli* MG1655 *in Silico* Metabolic Genotype: Its Definition, Characteristics, and Capabilities. *Proc. Natl. Acad. Sci. U. S. A.* **2000**, *97*, 5528−5533.
(27) Farmer, W. R.; Liao, J. C. Improving Lycopene Production in *Escherichia coli* by Engineering Metabolic Control. *Nat. Biotechnol.* **2000**, *18*, 533−537.
(28) Leonard, E.; Ajikumar, P. K.; Thayer, K.; Xiao, W.-H.; Mo, J. D.; Tidor, B.; Stephanopoulos, G.; Prather, K. L. J. Combining Metabolic and Protein Engineering of a Terpenoid Biosynthetic Pathway for Overproduction and Selectivity Control. *Proc. Natl. Acad. Sci. U. S. A.* **2010**, *107*, 13654−13659.
(29) Alper, H.; Moxley, J.; Nevoigt, E.; Fink, G. R.; Stephanopoulos, G. Engineering Yeast Transcription Machinery for Improved Ethanol Tolerance and Production. *Science* **2006**, *314*, 1565−1568.
(30) Atsumi, S.; Liao, J. C. Directed Evolution of *Methanococcus jannaschii* Citramalate Synthase for Biosynthesis of 1-Propanol and 1-Butanol by *Escherichia coli*. *Appl. Environ. Microbiol.* **2008**, *74*, 7802−7808.
(31) Guterl, J.-K.; Garbe, D.; Carsten, J.; Steffler, F.; Sommer, B.; Reiße, S.; Philipp, A.; Haack, M.; Rühmann, B.; Koltermann, A.; et al. Cell-Free Metabolic Engineering: Production of Chemicals by Minimized Reaction Cascades. *ChemSusChem* **2012**, *5*, 2165−2172.
(32) Jakočiūnas, T.; Bonde, I.; Herrgård, M.; Harrison, S. J.; Kristensen, M.; Pedersen, L. E.; Jensen, M. K.; Keasling, J. D. Multiplex Metabolic Pathway Engineering Using CRISPR/Cas9 in *Saccharomyces cerevisiae*. *Metab. Eng.* **2015**, *28*, 213−222.
(33) Li, Y.; Lin, Z.; Huang, C.; Zhang, Y.; Wang, Z.; Tang, Y.; Chen, T.; Zhao, X. Metabolic Engineering of *Escherichia coli* Using CRISPR−Cas9Meditated Genome Editing. *Metab. Eng.* **2015**, *31*, 13−21.
(34) Kim, S.; Lindner, S. N.; Aslan, S.; Yishai, O.; Wenk, S.; Schann, K.; Bar-Even, A. Growth of *E. coli* on Formate and Methanol via the Reductive Glycine Pathway. *Nat. Chem. Biol.* **2020**, *16*, 538−545.
(35) Gassler, T.; Sauer, M.; Gasser, B.; Egermeier, M.; Troyer, C.; Causon, T.; Hann, S.; Mattanovich, D.; Steiger, M. G. The Industrial Yeast *Pichia pastoris* Is Converted from a Heterotroph into an Autotroph Capable of Growth on CO_2. *Nat. Biotechnol.* **2020**, *38*, 210−216.
(36) Robinson, C. J.; Carbonell, P.; Jervis, A. J.; Yan, C.; Hollywood, K. A.; Dunstan, M. S.; Currin, A.; Swainston, N.; Spiess, R.; Taylor, S.; et al. Rapid Prototyping of Microbial Production Strains for the Biomanufacture of Potential Materials Monomers. *Metab. Eng.* **2020**, *60*, 168−182.
(37) Costello, Z.; Martin, H. G. A Machine Learning Approach to Predict Metabolic Pathway Dynamics from Time-Series Multiomics Data. *Npj Syst. Biol. Appl.* **2018**, *4*, 19.
(38) Shen, C. R.; Liao, J. C. Metabolic Engineering of *Escherichia coli* for 1-Butanol and 1-Propanol Production via the Keto-Acid Pathways. *Metab. Eng.* **2008**, *10*, 312−320.
(39) Fischer, M. J. C.; Meyer, S.; Claudel, P.; Bergdoll, M.; Karst, F. Metabolic Engineering of Monoterpene Synthesis in Yeast. *Biotechnol. Bioeng.* **2011**, *108*, 1883−1892.
(40) Wang, K.; Shi, T.-Q.; Wang, J.; Wei, P.; Ledesma-Amaro, R.; Ji, X.-J. Engineering the Lipid and Fatty Acid Metabolism in *Yarrowia lipolytica* for Sustainable Production of High Oleic Oils. *ACS Synth. Biol.* **2022**, *11*, 1542−1554.
(41) Vanhercke, T.; Petrie, J. R.; Singh, S. P. Energy Densification in Vegetative Biomass through Metabolic Engineering. *Biocatal. Agric. Biotechnol.* **2014**, *3*, 75−80.
(42) Winichayakul, S.; Scott, R. W.; Roldan, M.; Hatier, J.-H. B.; Livingston, S.; Cookson, R.; Curran, A. C.; Roberts, N. J. *In Vivo* Packaging of Triacylglycerols Enhances Arabidopsis Leaf Biomass and Energy Density. *Plant Physiol.* **2013**, *162*, 626−639.
(43) Li, Y.; Li, S.; Thodey, K.; Trenchard, I.; Cravens, A.; Smolke, C. D. Complete Biosynthesis of Noscapine and Halogenated Alkaloids in Yeast. *Proc. Natl. Acad. Sci. U. S. A.* **2018**, *115*, No. E3922-E3931.
(44) Lian, J.; Zhao, H. Functional Reconstitution of a Pyruvate Dehydrogenase in the Cytosol of *Saccharomyces cerevisiae* through Lipoylation Machinery Engineering. *ACS Synth. Biol.* **2016**, *5*, 689−697.
(45) Chen, Y.; Xu, D.; Fan, L.; Zhang, X.; Tan, T. Manipulating Multi-System of NADPH Regulation in *Escherichia coli* for Enhanced S-Adenosylmethionine Production. *RSC Adv.* **2015**, *5*, 41103−41111.
(46) Vemuri, G. N.; Eiteman, M. A.; Altman, E. Effects of Growth Mode and Pyruvate Carboxylase on Succinic Acid Production by Metabolically Engineered Strains of *Escherichia coli*. *Appl. Environ. Microbiol.* **2002**, *68*, 1715−1727.
(47) Toya, Y.; Nakahigashi, K.; Tomita, M.; Shimizu, K. Metabolic Regulation Analysis of Wild-Type and ArcA Mutant *Escherichia coli* under Nitrate Conditions Using Different Levels of Omics Data. *Mol. Biosyst.* **2012**, *8*, 2593−2604.
(48) Deng, C.; Lv, X.; Li, J.; Zhang, H.; Liu, Y.; Du, G.; Amaro, R. L.; Liu, L. Synergistic Improvement of N-Acetylglucosamine Production by Engineering Transcription Factors and Balancing Redox Cofactors. *Metab. Eng.* **2021**, *67*, 330−346.
(49) Cahn, J. K. B.; Werlang, C. A.; Baumschlager, A.; Brinkmann-Chen, S.; Mayo, S. L.; Arnold, F. H. A General Tool for Engineering the NAD/NADP Cofactor Preference of Oxidoreductases. *ACS Synth. Biol.* **2017**, *6*, 326−333.
(50) Wang, M.; Chen, B.; Fang, Y.; Tan, T. Cofactor Engineering for More Efficient Production of Chemicals and Biofuels. *Biotechnol. Adv.* **2017**, *35*, 1032−1039.
(51) Chen, R.; Yang, S.; Zhang, L.; Zhou, Y. J. Advanced Strategies for Production of Natural Products in Yeast. *iScience* **2020**, *23*, 100879.
(52) Wang, C.; Yoon, S.-H.; Jang, H.-J.; Chung, Y.-R.; Kim, J.-Y.; Choi, E.-S.; Kim, S.-W. Metabolic Engineering of *Escherichia coli* for α-Farnesene Production. *Metab. Eng.* **2011**, *13*, 648−655.
(53) Lee, D.; Lloyd, N. D. R.; Pretorius, I. S.; Borneman, A. R. Heterologous Production of Raspberry Ketone in the Wine Yeast *Saccharomyces cerevisiae* via Pathway Engineering and Synthetic Enzyme Fusion. *Microb. Cell Factories* **2016**, *15*, 49.
(54) Kummer, M. J.; Lee, Y. S.; Yuan, M.; Alkotaini, B.; Zhao, J.; Blumenthal, E.; Minteer, S. D. Substrate Channeling by a Rationally Designed Fusion Protein in a Biocatalytic Cascade. *JACS Au* **2021**, *1*, 1187−1197.
(55) Sachdeva, G.; Garg, A.; Godding, D.; Way, J. C.; Silver, P. A. *In Vivo* Co-Localization of Enzymes on RNA Scaffolds Increases Metabolic Production in a Geometrically Dependent Manner. *Nucleic Acids Res.* **2014**, *42*, 9493−9503.
(56) Dueber, J. E.; Wu, G. C.; Malmirchegini, G. R.; Moon, T. S.; Petzold, C. J.; Ullal, A. V.; Prather, K. L. J.; Keasling, J. D. Synthetic

https://doi.org/10.1021/acs.chemrev.2c00403
Chem. Rev. 2023, 123, 5521−5570

Protein Scaffolds Provide Modular Control over Metabolic Flux. *Nat. Biotechnol.* **2009**, *27*, 753−759.
(57) Conrado, R. J.; Wu, G. C.; Boock, J. T.; Xu, H.; Chen, S. Y.; Lebar, T.; Turnšek, J.; Tomšič, N.; Avbelj, M.; Gaber, R.; et al. DNA-Guided Assembly of Biosynthetic Pathways Promotes Improved Catalytic Efficiency. *Nucleic Acids Res.* **2012**, *40*, 1879−1889.
(58) Poirier, Y.; Erard, N.; Petétot, J. M.-C. Synthesis of Polyhydroxyalkanoate in the Peroxisome of *Saccharomyces cerevisiae* by Using Intermediates of Fatty Acid β-Oxidation. *Appl. Environ. Microbiol.* **2001**, *67*, 5254−5260.
(59) Li, J.; Mutanda, I.; Wang, K.; Yang, L.; Wang, J.; Wang, Y. Chloroplastic Metabolic Engineering Coupled with Isoprenoid Pool Enhancement for Committed Taxanes Biosynthesis in *Nicotiana benthamiana*. *Nat. Commun.* **2019**, *10*, 4850.
(60) Malhotra, K.; Subramaniyan, M.; Rawat, K.; Kalamuddin, M.; Qureshi, M. I.; Malhotra, P.; Mohmmed, A.; Cornish, K.; Daniell, H.; Kumar, S. Compartmentalized Metabolic Engineering for Artemisinin Biosynthesis and Effective Malaria Treatment by Oral Delivery of Plant Cells. *Mol. Plant* **2016**, *9*, 1464−1477.
(61) Avalos, J. L.; Fink, G. R.; Stephanopoulos, G. Compartmentalization of Metabolic Pathways in Yeast Mitochondria Improves the Production of Branched-Chain Alcohols. *Nat. Biotechnol.* **2013**, *31*, 335−341.
(62) Hammer, S. K.; Avalos, J. L. Harnessing Yeast Organelles for Metabolic Engineering. *Nat. Chem. Biol.* **2017**, *13*, 823−832.
(63) Lv, X.; Wang, F.; Zhou, P.; Ye, L.; Xie, W.; Xu, H.; Yu, H. Dual Regulation of Cytoplasmic and Mitochondrial Acetyl-CoA Utilization for Improved Isoprene Production in *Saccharomyces cerevisiae*. *Nat. Commun.* **2016**, *7*, 12851.
(64) Liu, G.-S.; Li, T.; Zhou, W.; Jiang, M.; Tao, X.-Y.; Liu, M.; Zhao, M.; Ren, Y.-H.; Gao, B.; Wang, F.-Q.; et al. The Yeast Peroxisome: A Dynamic Storage Depot and Subcellular Factory for Squalene Overproduction. *Metab. Eng.* **2020**, *57*, 151−161.
(65) Ma, Y.; Li, J.; Huang, S.; Stephanopoulos, G. Targeting Pathway Expression to Subcellular Organelles Improves Astaxanthin Synthesis in *Yarrowia lipolytica*. *Metab. Eng.* **2021**, *68*, 152−161.
(66) Yang, K.; Qiao, Y.; Li, F.; Xu, Y.; Yan, Y.; Madzak, C.; Yan, J. Subcellular Engineering of Lipase Dependent Pathways Directed towards Lipid Related Organelles for Highly Effectively Compartmentalized Biosynthesis of Triacylglycerol Derived Products in *Yarrowia lipolytica*. *Metab. Eng.* **2019**, *55*, 231−238.
(67) Thodey, K.; Galanie, S.; Smolke, C. D. A Microbial Biomanufacturing Platform for Natural and Semisynthetic Opioids. *Nat. Chem. Biol.* **2014**, *10*, 837−844.
(68) Parsons, J. B.; Dinesh, S. D.; Deery, E.; Leech, H. K.; Brindley, A. A.; Heldt, D.; Frank, S.; Smales, C. M.; Lünsdorf, H.; Rambach, A.; et al. Biochemical and Structural Insights into Bacterial Organelle Form and Biogenesis. *J. Biol. Chem.* **2008**, *283*, 14366−14375.
(69) Li, L.; Zhang, R.; Tian, X.; Li, T.; Pu, B.; Ma, C.; Ba, F.; Xiong, C.; Shi, Y.; Li, J. Configurable Compartmentation Enables *In Vitro* Reconstitution of Sustained Synthetic Biology Systems. *bioRxiv* 2022, DOI: 10.1101/2022.03.19.484972.
(70) Zhao, E. M.; Suek, N.; Wilson, M. Z.; Dine, E.; Pannucci, N. L.; Gitai, Z.; Avalos, J. L.; Toettcher, J. E. Light-Based Control of Metabolic Flux through Assembly of Synthetic Organelles. *Nat. Chem. Biol.* **2019**, *15*, 589−597.
(71) Kim, J.-E.; Jang, I.-S.; Son, S.-H.; Ko, Y.-J.; Cho, B.-K.; Kim, S. C.; Lee, J. Y. Tailoring the *Saccharomyces cerevisiae* Endoplasmic Reticulum for Functional Assembly of Terpene Synthesis Pathway. *Metab. Eng.* **2019**, *56*, 50−59.
(72) Du, J.; Li, S.; Zhao, H. Discovery and Characterization of Novel D-Xylose-Specific Transporters from *Neurospora crassa* and *Pichia stipitis*. *Mol. Biosyst.* **2010**, *6*, 2150−2156.
(73) Havukainen, S.; Pujol-Giménez, J.; Valkonen, M.; Hediger, M. A.; Landowski, C. P. Functional Characterization of a Highly Specific L-Arabinose Transporter from *Trichoderma reesei*. *Microb. Cell Factories* **2021**, *20*, 177.
(74) Protzko, R. J.; Latimer, L. N.; Martinho, Z.; de Reus, E.; Seibert, T.; Benz, J. P.; Dueber, J. E. Engineering *Saccharomyces cerevisiae* for Co-Utilization of D-Galacturonic Acid and D-Glucose from Citrus Peel Waste. *Nat. Commun.* **2018**, *9*, 5059.
(75) Podolsky, I. A.; Seppälä, S.; Xu, H.; Jin, Y.-S.; O'Malley, M. A. A SWEET Surprise: Anaerobic Fungal Sugar Transporters and Chimeras Enhance Sugar Uptake in Yeast. *Metab. Eng.* **2021**, *66*, 137−147.
(76) Li, S.; Du, J.; Sun, J.; Galazka, J. M.; Glass, N. L.; Cate, J. H. D.; Yang, X.; Zhao, H. Overcoming Glucose Repression in Mixed Sugar Fermentation by Co-Expressing a Cellobiose Transporter and a β-Glucosidase in *Saccharomyces cerevisiae*. *Mol. Biosyst.* **2010**, *6*, 2129−2132.
(77) Eriksen, D. T.; Hsieh, P. C. H.; Lynn, P.; Zhao, H. Directed Evolution of a Cellobiose Utilization Pathway in *Saccharomyces cerevisiae* by Simultaneously Engineering Multiple Proteins. *Microb. Cell Factories* **2013**, *12*, 61.
(78) Thomik, T.; Wittig, I.; Choe, J.; Boles, E.; Oreb, M. An Artificial Transport Metabolon Facilitates Improved Substrate Utilization in Yeast. *Nat. Chem. Biol.* **2017**, *13*, 1158−1163.
(79) Li, W.; Ma, L.; Shen, X.; Wang, J.; Feng, Q.; Liu, L.; Zheng, G.; Yan, Y.; Sun, X.; Yuan, Q. Targeting Metabolic Driving and Intermediate Influx in Lysine Catabolism for High-Level Glutarate Production. *Nat. Commun.* **2019**, *10*, 3337.
(80) Kim, J.; Yoo, H.-W.; Kim, M.; Kim, E.-J.; Sung, C.; Lee, P.-G.; Park, B. G.; Kim, B.-G. Rewiring FadR Regulon for the Selective Production of ω-Hydroxy Palmitic Acid from Glucose in *Escherichia coli*. *Metab. Eng.* **2018**, *47*, 414−422.
(81) Ferreira, R.; Teixeira, P. G.; Gossing, M.; David, F.; Siewers, V.; Nielsen, J. Metabolic Engineering of *Saccharomyces cerevisiae* for Overproduction of Triacylglycerols. *Metab. Eng. Commun.* **2018**, *6*, 22−27.
(82) Rodriguez, A.; Chen, Y.; Khoomrung, S.; Özdemir, E.; Borodina, I.; Nielsen, J. Comparison of the Metabolic Response to Over-Production of p-Coumaric Acid in Two Yeast Strains. *Metab. Eng.* **2017**, *44*, 265−272.
(83) Srinivasan, P.; Smolke, C. D. Biosynthesis of Medicinal Tropane Alkaloids in Yeast. *Nature* **2020**, *585*, 614−619.
(84) Srinivasan, P.; Smolke, C. D. Engineering Cellular Metabolite Transport for Biosynthesis of Computationally Predicted Tropane Alkaloid Derivatives in Yeast. *Proc. Natl. Acad. Sci. U. S. A.* **2021**, *118*, e2104460118.
(85) Yuzbasheva, E. Y.; Scarcia, P.; Yuzbashev, T. V.; Messina, E.; Kosikhina, I. M.; Palmieri, L.; Shutov, A. V.; Taratynova, M. O.; Amaro, R. L.; Palmieri, F.; et al. Engineering *Yarrowia lipolytica* for the Selective and High-Level Production of Isocitric Acid through Manipulation of Mitochondrial Dicarboxylate−Tricarboxylate Carriers. *Metab. Eng.* **2021**, *65*, 156−166.
(86) Dunlop, M. J; Dossani, Z. Y; Szmidt, H. L; Chu, H. C.; Lee, T. S.; Keasling, J. D; Hadi, M. Z; Mukhopadhyay, A. Engineering Microbial Biofuel Tolerance and Export Using Efflux Pumps. *Mol. Syst. Biol.* **2011**, *7*, 487.
(87) Wu, J.; Wang, Z.; Zhang, X.; Zhou, P.; Xia, X.; Dong, M. Improving Medium Chain Fatty Acid Production in *Escherichia coli* by Multiple Transporter Engineering. *Food Chem.* **2019**, *272*, 628−634.
(88) Kurgan, G.; Kurgan, L.; Schneider, A.; Onyeabor, M.; Rodriguez-Sanchez, Y.; Taylor, E.; Martinez, R.; Carbonell, P.; Shi, X.; Gu, H.; et al. Identification of Major Malate Export Systems in an Engineered Malate-Producing *Escherichia coli* Aided by Substrate Similarity Search. *Appl. Microbiol. Biotechnol.* **2019**, *103*, 9001−9011.
(89) van der Hoek, S. A.; Borodina, I. Transporter Engineering in Microbial Cell Factories: The Ins, the Outs, and the in-Betweens. *Curr. Opin. Biotechnol.* **2020**, *66*, 186−194.
(90) Zhu, Y.; Zhou, C.; Wang, Y.; Li, C. Transporter Engineering for Microbial Manufacturing. *Biotechnol. J.* **2020**, *15*, 1900494.
(91) Cakar, Z.; Seker, U.; Tamerler, C.; Sonderegger, M.; Sauer, U. Evolutionary Engineering of Multiple-Stress Resistant. *FEMS Yeast Res.* **2005**, *5*, 569−578.
(92) Caspeta, L.; Chen, Y.; Ghiaci, P.; Feizi, A.; Buskov, S.; Hallström, B. M.; Petranovic, D.; Nielsen, J. Altered Sterol

Composition Renders Yeast Thermotolerant. *Science* **2014**, *346*, 75−78.
(93) Lenski, R. E.; Rose, M. R.; Simpson, S. C.; Tadler, S. C. Long-Term Experimental Evolution in *Escherichia coli*. I. Adaptation and Divergence during 2,000 Generations. *Am. Nat.* **1991**, *138*, 1315−1341.
(94) Sandberg, T. E.; Salazar, M. J.; Weng, L. L.; Palsson, B. O.; Feist, A. M. The Emergence of Adaptive Laboratory Evolution as an Efficient Tool for Biological Discovery and Industrial Biotechnology. *Metab. Eng.* **2019**, *56*, 1−16.
(95) Wang, Y.; Xue, P.; Cao, M.; Yu, T.; Lane, S. T.; Zhao, H. Directed Evolution: Methodologies and Applications. *Chem. Rev.* **2021**, *121*, 12384−12444.
(96) Ravikumar, A.; Arzumanyan, G. A.; Obadi, M. K. A.; Javanpour, A. A.; Liu, C. C. Scalable, Continuous Evolution of Genes at Mutation Rates above Genomic Error Thresholds. *Cell* **2018**, *175*, 1946−1957.
(97) Halperin, S. O.; Tou, C. J.; Wong, E. B.; Modavi, C.; Schaffer, D. V.; Dueber, J. E. CRISPR-Guided DNA Polymerases Enable Diversification of All Nucleotides in a Tunable Window. *Nature* **2018**, *560*, 248−252.
(98) Moore, C. L.; Papa, L. J.; Shoulders, M. D. A Processive Protein Chimera Introduces Mutations across Defined DNA Regions *In Vivo*. *J. Am. Chem. Soc.* **2018**, *140*, 11560−11564.
(99) Molina, R. S.; Rix, G.; Mengiste, A. A.; Álvarez, B.; Seo, D.; Chen, H.; Hurtado, J. E.; Zhang, Q.; García-García, J. D.; Heins, Z. J.; et al. *In Vivo* Hypermutation and Continuous Evolution. *Nat. Rev. Methods Primer* **2022**, *2*, 1−22.
(100) Wang, H. H.; Kim, H.; Cong, L.; Jeong, J.; Bang, D.; Church, G. M. Genome-Scale Promoter Engineering by Coselection MAGE. *Nat. Methods* **2012**, *9*, 591−593.
(101) Wang, H. H.; Isaacs, F. J.; Carr, P. A.; Sun, Z. Z.; Xu, G.; Forest, C. R.; Church, G. M. Programming Cells by Multiplex Genome Engineering and Accelerated Evolution. *Nature* **2009**, *460*, 894−898.
(102) Du, J.; Yuan, Y.; Si, T.; Lian, J.; Zhao, H. Customized Optimization of Metabolic Pathways by Combinatorial Transcriptional Engineering. *Nucleic Acids Res.* **2012**, *40*, No. e142.
(103) Lian, J.; Schultz, C.; Cao, M.; HamediRad, M.; Zhao, H. Multi-Functional Genome-Wide CRISPR System for High Throughput Genotype−Phenotype Mapping. *Nat. Commun.* **2019**, *10*, 5794.
(104) Lee, S. Y.; Nielsen, J.; Stephanopoulos, G. *Metabolic Engineering: Concepts and Applications*; John Wiley & Sons, 2021.
(105) Coley, C. W.; Rogers, L.; Green, W. H.; Jensen, K. F. SCScore: Synthetic Complexity Learned from a Reaction Corpus. *J. Chem. Inf. Model.* **2018**, *58*, 252−261.
(106) Kumar, A.; Wang, L.; Ng, C. Y.; Maranas, C. D. Pathway Design Using de Novo Steps through Uncharted Biochemical Spaces. *Nat. Commun.* **2018**, *9*, 184.
(107) Koch, M.; Duigou, T.; Faulon, J.-L. Reinforcement Learning for Bioretrosynthesis. *ACS Synth. Biol.* **2020**, *9*, 157−168.
(108) Finnigan, W.; Hepworth, L. J.; Flitsch, S. L.; Turner, N. J. RetroBioCat as a Computer-Aided Synthesis Planning Tool for Biocatalytic Reactions and Cascades. *Nat. Catal.* **2021**, *4*, 98−104.
(109) Ding, S.; Tian, Y.; Cai, P.; Zhang, D.; Cheng, X.; Sun, D.; Yuan, L.; Chen, J.; Tu, W.; Wei, D.-Q.; et al. NovoPathFinder: A Webserver of Designing Novel-Pathway with Integrating GEM-Model. *Nucleic Acids Res.* **2020**, *48*, W477−W487.
(110) Probst, D.; Manica, M.; Nana Teukam, Y. G.; Castrogiovanni, A.; Paratore, F.; Laino, T. Biocatalysed Synthesis Planning Using Data-Driven Learning. *Nat. Commun.* **2022**, *13*, 964.
(111) Kim, G. B.; Kim, W. J.; Kim, H. U.; Lee, S. Y. Machine Learning Applications in Systems Metabolic Engineering. *Curr. Opin. Biotechnol.* **2020**, *64*, 1−9.
(112) Kim, D. I.; Chae, T. U.; Kim, H. U.; Jang, W. D.; Lee, S. Y. Microbial Production of Multiple Short-Chain Primary Amines via Retrobiosynthesis. *Nat. Commun.* **2021**, *12*, 173.
(113) Gu, P.; Yang, F.; Su, T.; Wang, Q.; Liang, Q.; Qi, Q. A Rapid and Reliable Strategy for Chromosomal Integration of Gene(s) with Multiple Copies. *Sci. Rep.* **2015**, *5*, 9684.
(114) Datsenko, K. A.; Wanner, B. L. One-Step Inactivation of Chromosomal Genes in *Escherichia coli* K-12 Using PCR Products. *Proc. Natl. Acad. Sci. U. S. A.* **2000**, *97*, 6640−6645.
(115) Gueldener, U.; Heinisch, J.; Koehler, G. J.; Voss, D.; Hegemann, J. H. A Second Set of LoxP Marker Cassettes for Cre-Mediated Multiple Gene Knockouts in Budding Yeast. *Nucleic Acids Res.* **2002**, *30*, No. e23.
(116) Güldener, U.; Heck, S.; Fiedler, T.; Beinhauer, J.; Hegemann, J. H. A New Efficient Gene Disruption Cassette for Repeated Use in Budding Yeast. *Nucleic Acids Res.* **1996**, *24*, 2519−2524.
(117) Murphy, K. C. Use of Bacteriophage λ Recombination Functions To Promote Gene Replacement in *Escherichia coli*. *J. Bacteriol.* **1998**, *180*, 2063−2071.
(118) Escalante, A.; Calderón, R.; Valdivia, A.; de Anda, R.; Hernández, G.; Ramírez, O. T.; Gosset, G.; Bolívar, F. Metabolic Engineering for the Production of Shikimic Acid in an Evolved *Escherichia coli* Strain Lacking the Phosphoenolpyruvate: Carbohydrate Phosphotransferase System. *Microb. Cell Factories* **2010**, *9*, 21.
(119) Peirú, S.; Rodríguez, E.; Menzella, H. G.; Carney, J. R.; Gramajo, H. Metabolically Engineered *Escherichia coli* for Efficient Production of Glycosylated Natural Products. *Microb. Biotechnol.* **2008**, *1*, 476−486.
(120) Zhang, X.; Shanmugam, K. T.; Ingram, L. O. Fermentation of Glycerol to Succinate by Metabolically Engineered Strains of *Escherichia coli*. *Appl. Environ. Microbiol.* **2010**, *76*, 2397.
(121) Maeda, T.; Sanchez-Torres, V.; Wood, T. K. Enhanced Hydrogen Production from Glucose by Metabolically Engineered *Escherichia coli*. *Appl. Microbiol. Biotechnol.* **2007**, *77*, 879−890.
(122) Maeda, T.; Sanchez-Torres, V.; Wood, T. K. *Escherichia coli* Hydrogenase 3 Is a Reversible Enzyme Possessing Hydrogen Uptake and Synthesis Activities. *Appl. Microbiol. Biotechnol.* **2007**, *76*, 1035−1042.
(123) Chen, N.; Wang, J.; Zhao, Y.; Deng, Y. Metabolic Engineering of *Saccharomyces cerevisiae* for Efficient Production of Glucaric Acid at High Titer. *Microb. Cell Factories* **2018**, *17*, 67.
(124) Hegemann, J. H.; Heick, S. B. Delete and Repeat: A Comprehensive Toolkit for Sequential Gene Knockout in the Budding Yeast *Saccharomyces cerevisiae*. In *Strain Engineering: Methods and Protocols*; Methods in Molecular Biology; Williams, J. A., Ed.; Humana Press: Totowa, NJ, 2011; pp 189−206.
(125) Su, B.; Song, D.; Zhu, H. Homology-Dependent Recombination of Large Synthetic Pathways into *E. coli* Genome via λ-Red and CRISPR/Cas9 Dependent Selection Methodology. *Microb. Cell Factories* **2020**, *19*, 108.
(126) Ng, C. Y.; Farasat, I.; Maranas, C. D.; Salis, H. M. Rational Design of a Synthetic Entner−Doudoroff Pathway for Improved and Controllable NADPH Regeneration. *Metab. Eng.* **2015**, *29*, 86−96.
(127) Merrick, C. A.; Zhao, J.; Rosser, S. J. Serine Integrases: Advancing Synthetic Biology. *ACS Synth. Biol.* **2018**, *7*, 299−310.
(128) Morita, K.; Morimura, K.; Fusada, N.; Komatsu, M.; Ikeda, H.; Hirano, N.; Takahashi, H. Site-Specific Genome Integration in Alphaproteobacteria Mediated by TG1 Integrase. *Appl. Microbiol. Biotechnol.* **2012**, *93*, 295−304.
(129) Li, L.; Zheng, G.; Chen, J.; Ge, M.; Jiang, W.; Lu, Y. Multiplexed Site-Specific Genome Engineering for Overproducing Bioactive Secondary Metabolites in Actinomycetes. *Metab. Eng.* **2017**, *40*, 80−92.
(130) Huang, H.; Chai, C.; Yang, S.; Jiang, W.; Gu, Y. Phage Serine Integrase-Mediated Genome Engineering for Efficient Expression of Chemical Biosynthetic Pathway in Gas-Fermenting Clostridium Ljungdahlii. *Metab. Eng.* **2019**, *52*, 293−302.
(131) Sikorski, R. S.; Hieter, P. A System of Shuttle Vectors and Yeast Host Strains Designed for Efficient Manipulation of DNA in *Saccharomyces cerevisiae*. *Genetics* **1989**, *122*, 19−27.
(132) Shao, Z.; Zhao, H.; Zhao, H. DNA Assembler, an *in Vivo* Genetic Method for Rapid Construction of Biochemical Pathways. *Nucleic Acids Res.* **2009**, *37*, No. e16.
(133) Mikkelsen, M. D.; Buron, L. D.; Salomonsen, B.; Olsen, C. E.; Hansen, B. G.; Mortensen, U. H.; Halkier, B. A. Microbial Production

https://doi.org/10.1021/acs.chemrev.2c00403
Chem. Rev. 2023, 123, 5521−5570

of Indolylglucosinolate through Engineering of a Multi-Gene Pathway in a Versatile Yeast Expression Platform. *Metab. Eng.* **2012**, *14*, 104−111.
(134) Siddiqui, M. S.; Choksi, A.; Smolke, C. D. A System for Multilocus Chromosomal Integration and Transformation-Free Selection Marker Rescue. *FEMS Yeast Res.* **2014**, *14*, 1171−1185.
(135) Maury, J.; Germann, S. M.; Baallal Jacobsen, S. A.; Jensen, N. B.; Kildegaard, K. R.; Herrgård, M. J.; Schneider, K.; Koza, A.; Forster, J.; Nielsen, J.; Borodina, I EasyCloneMulti: A Set of Vectors for Simultaneous and Multiple Genomic Integrations in *Saccharomyces cerevisiae*. *PLoS One* **2016**, *11*, No. e0150394.
(136) Lian, J.; HamediRad, M.; Zhao, H. Advancing Metabolic Engineering of *Saccharomyces cerevisiae* Using the CRISPR/Cas System. *Biotechnol. J.* **2018**, *13*, 1700601.
(137) Schwartz, C. M.; Hussain, M. S.; Blenner, M.; Wheeldon, I. Synthetic RNA Polymerase III Promoters Facilitate High-Efficiency CRISPR−Cas9-Mediated Genome Editing in *Yarrowia lipolytica*. *ACS Synth. Biol.* **2016**, *5*, 356−359.
(138) Koh, C. M. J.; Liu, Y.; Moehninsi; Du, M.; Ji, L. Molecular Characterization of KU70 and KU80 Homologues and Exploitation of a KU70-Deficient Mutant for Improving Gene Deletion Frequency in *Rhodosporidium toruloides*. *BMC Microbiol.* **2014**, *14*, 50.
(139) Lu, L.; Shen, X.; Sun, X.; Yan, Y.; Wang, J.; Yuan, Q. CRISPR-Based Metabolic Engineering in Non-Model Microorganisms. *Curr. Opin. Biotechnol.* **2022**, *75*, 102698.
(140) DiCarlo, J. E.; Norville, J. E.; Mali, P.; Rios, X.; Aach, J.; Church, G. M. Genome Engineering in *Saccharomyces cerevisiae* Using CRISPR-Cas Systems. *Nucleic Acids Res.* **2013**, *41*, 4336−4343.
(141) Bao, Z.; Xiao, H.; Liang, J.; Zhang, L.; Xiong, X.; Sun, N.; Si, T.; Zhao, H. Homology-Integrated CRISPR−Cas (HI-CRISPR) System for One-Step Multigene Disruption in *Saccharomyces cerevisiae*. *ACS Synth. Biol.* **2015**, *4*, 585−594.
(142) Schultz, J. C.; Cao, M.; Zhao, H. Development of a CRISPR/Cas9 System for High Efficiency Multiplexed Gene Deletion in *Rhodosporidium toruloides*. *Biotechnol. Bioeng.* **2019**, *116*, 2103−2109.
(143) Wu, H.; Li, Y.; Ma, Q.; Li, Q.; Jia, Z.; Yang, B.; Xu, Q.; Fan, X.; Zhang, C.; Chen, N.; et al. Metabolic Engineering of *Escherichia coli* for High-Yield Uridine Production. *Metab. Eng.* **2018**, *49*, 248−256.
(144) Lee, J. W.; Kwak, S.; Liu, J.-J.; Yu, S.; Yun, E. J.; Kim, D. H.; Liu, C.; Kim, K. H.; Jin, Y.-S. Enhanced 2′-Fucosyllactose Production by Engineered *Saccharomyces cerevisiae* Using Xylose as a Co-Substrate. *Metab. Eng.* **2020**, *62*, 322−329.
(145) Shi, S.; Liang, Y.; Zhang, M. M.; Ang, E. L.; Zhao, H. A Highly Efficient Single-Step, Markerless Strategy for Multi-Copy Chromosomal Integration of Large Biochemical Pathways in *Saccharomyces cerevisiae*. *Metab. Eng.* **2016**, *33*, 19−27.
(146) Alonso-Gutierrez, J.; Koma, D.; Hu, Q.; Yang, Y.; Chan, L. J. G.; Petzold, C. J.; Adams, P. D.; Vickers, C. E.; Nielsen, L. K.; Keasling, J. D.; et al. Toward Industrial Production of Isoprenoids in *Escherichia coli*: Lessons Learned from CRISPR-Cas9 Based Optimization of a Chromosomally Integrated Mevalonate Pathway. *Biotechnol. Bioeng.* **2018**, *115*, 1000−1013.
(147) Heo, M.-J.; Jung, H.-M.; Um, J.; Lee, S.-W.; Oh, M.-K. Controlling Citrate Synthase Expression by CRISPR/Cas9 Genome Editing for n-Butanol Production in *Escherichia coli*. *ACS Synth. Biol.* **2017**, *6*, 182−189.
(148) Keasling, J. D. Gene-Expression Tools for the Metabolic Engineering of Bacteria. *Trends Biotechnol.* **1999**, *17*, 452−460.
(149) Holtz, W. J.; Keasling, J. D. Engineering Static and Dynamic Control of Synthetic Pathways. *Cell* **2010**, *140*, 19−23.
(150) Liu, D.; Mannan, A. A.; Han, Y.; Oyarzún, D. A.; Zhang, F. Dynamic Metabolic Control: Towards Precision Engineering of Metabolism. *J. Ind. Microbiol. Biotechnol.* **2018**, *45*, 535−543.
(151) Hartline, C. J.; Schmitz, A. C.; Han, Y.; Zhang, F. Dynamic Control in Metabolic Engineering: Theories, Tools, and Applications. *Metab. Eng.* **2021**, *63*, 126−140.
(152) Cazier, A. P.; Blazeck, J. Advances in Promoter Engineering: Novel Applications and Predefined Transcriptional Control. *Biotechnol. J.* **2021**, *16*, 2100239.
(153) Zhu, J.; Zhang, M. Q. SCPD: A Promoter Database of the Yeast *Saccharomyces cerevisiae*. *Bioinformatics* **1999**, *15*, 607−611.
(154) *Promoters/Catalog/Anderson*; IGEM, 2022; http://parts.igem.org/Promoters/Catalog/Anderson (accessed 2022-08-30).
(155) Tang, H.; Wu, Y.; Deng, J.; Chen, N.; Zheng, Z.; Wei, Y.; Luo, X.; Keasling, J. D. Promoter Architecture and Promoter Engineering in *Saccharomyces cerevisiae*. *Metabolites* **2020**, *10*, 320.
(156) Blazeck, J.; Liu, L.; Redden, H.; Alper, H. Tuning Gene Expression in *Yarrowia lipolytica* by a Hybrid Promoter Approach. *Appl. Environ. Microbiol.* **2011**, *77*, 7905−7914.
(157) Deng, J.; Wu, Y.; Zheng, Z.; Chen, N.; Luo, X.; Tang, H.; Keasling, J. D. A Synthetic Promoter System for Well-Controlled Protein Expression with Different Carbon Sources in *Saccharomyces cerevisiae*. *Microb. Cell Factories* **2021**, *20*, 202.
(158) Zhou, C.; Ye, B.; Cheng, S.; Zhao, L.; Liu, Y.; Jiang, J.; Yan, X. Promoter Engineering Enables Overproduction of Foreign Proteins from a Single Copy Expression Cassette in *Bacillus subtilis*. *Microb. Cell Factories* **2019**, *18*, 111.
(159) Redden, H.; Alper, H. S. The Development and Characterization of Synthetic Minimal Yeast Promoters. *Nat. Commun.* **2015**, *6*, 7810.
(160) Jin, L.-Q.; Jin, W.-R.; Ma, Z.-C.; Shen, Q.; Cai, X.; Liu, Z.-Q.; Zheng, Y.-G. Promoter Engineering Strategies for the Overproduction of Valuable Metabolites in Microbes. *Appl. Microbiol. Biotechnol.* **2019**, *103*, 8725−8736.
(161) Alper, H.; Fischer, C.; Nevoigt, E.; Stephanopoulos, G. Tuning Genetic Control through Promoter Engineering. *Proc. Natl. Acad. Sci. U. S. A.* **2005**, *102*, 12678−12683.
(162) Yuan, Y.; Zhao, H. Directed Evolution of a Highly Efficient Cellobiose Utilizing Pathway in an Industrial *Saccharomyces cerevisiae* Strain. *Biotechnol. Bioeng.* **2013**, *110*, 2874−2881.
(163) Xu, N.; Wei, L.; Liu, J. Recent Advances in the Applications of Promoter Engineering for the Optimization of Metabolite Biosynthesis. *World J. Microbiol. Biotechnol.* **2019**, *35*, 33.
(164) Wang, Y.; Cheng, H.; Liu, Y.; Liu, Y.; Wen, X.; Zhang, K.; Ni, X.; Gao, N.; Fan, L.; Zhang, Z.; et al. *In-Situ* Generation of Large Numbers of Genetic Combinations for Metabolic Reprogramming via CRISPR-Guided Base Editing. *Nat. Commun.* **2021**, *12*, 678.
(165) Salis, H. M.; Mirsky, E. A.; Voigt, C. A. Automated Design of Synthetic Ribosome Binding Sites to Control Protein Expression. *Nat. Biotechnol.* **2009**, *27*, 946−950.
(166) Lin, P.-C.; Zhang, F.; Pakrasi, H. B. Enhanced Limonene Production in a Fast-Growing Cyanobacterium through Combinatorial Metabolic Engineering. *Metab. Eng. Commun.* **2021**, *12*, No. e00164.
(167) Cetnar, D. P.; Salis, H. M. Systematic Quantification of Sequence and Structural Determinants Controlling MRNA Stability in Bacterial Operons. *ACS Synth. Biol.* **2021**, *10*, 318−332.
(168) Li, H.; Gao, S.; Zhang, S.; Zeng, W.; Zhou, J. Effects of Metabolic Pathway Gene Copy Numbers on the Biosynthesis of (2S)-Naringenin in *Saccharomyces cerevisiae*. *J. Biotechnol.* **2021**, *325*, 119−127.
(169) Saleski, T. E.; Chung, M. T.; Carruthers, D. N.; Khasbaatar, A.; Kurabayashi, K.; Lin, X. N. Optimized Gene Expression from Bacterial Chromosome by High-Throughput Integration and Screening. *Sci. Adv.* **2021**, *7*, No. eabe1767.
(170) Lian, J.; Jin, R.; Zhao, H. Construction of Plasmids with Tunable Copy Numbers in *Saccharomyces cerevisiae* and Their Applications in Pathway Optimization and Multiplex Genome Integration. *Biotechnol. Bioeng.* **2016**, *113*, 2462−2473.
(171) Chen, Y.; Guo, E.; Zhang, J.; Si, T. Advances in RNAi-Assisted Strain Engineering in *Saccharomyces cerevisiae*. *Front. Bioeng. Biotechnol.* **2020**, *8*, 731.
(172) Si, T.; Luo, Y.; Bao, Z.; Zhao, H. RNAi-Assisted Genome Evolution in *Saccharomyces cerevisiae* for Complex Phenotype Engineering. *ACS Synth. Biol.* **2015**, *4*, 283−291.

https://doi.org/10.1021/acs.chemrev.2c00403
Chem. Rev. 2023, 123, 5521−5570

(173) HamediRad, M.; Lian, J.; Li, H.; Zhao, H. RNAi Assisted Genome Evolution Unveils Yeast Mutants with Improved Xylose Utilization. *Biotechnol. Bioeng.* **2018**, *115*, 1552−1560.
(174) Wang, G.; Björk, S. M.; Huang, M.; Liu, Q.; Campbell, K.; Nielsen, J.; Joensson, H. N.; Petranovic, D. RNAi Expression Tuning, Microfluidic Screening, and Genome Recombineering for Improved Protein Production in *Saccharomyces cerevisiae*. *Proc. Natl. Acad. Sci. U. S. A.* **2019**, *116*, 9324−9332.
(175) Na, D.; Yoo, S. M.; Chung, H.; Park, H.; Park, J. H.; Lee, S. Y. Metabolic Engineering of *Escherichia coli* Using Synthetic Small Regulatory RNAs. *Nat. Biotechnol.* **2013**, *31*, 170−174.
(176) Yang, D.; Yoo, S. M.; Gu, C.; Ryu, J. Y.; Lee, J. E.; Lee, S. Y. Expanded Synthetic Small Regulatory RNA Expression Platforms for Rapid and Multiplex Gene Expression Knockdown. *Metab. Eng.* **2019**, *54*, 180−190.
(177) Qi, L. S.; Larson, M. H.; Gilbert, L. A.; Doudna, J. A.; Weissman, J. S.; Arkin, A. P.; Lim, W. A. Repurposing CRISPR as an RNA-Guided Platform for Sequence-Specific Control of Gene Expression. *Cell* **2013**, *152*, 1173−1183.
(178) MacLeod, R. S.; Cawley, K. M.; Gubrij, I.; Nookaew, I.; Onal, M.; O'Brien, C. A. Effective CRISPR Interference of an Endogenous Gene via a Single Transgene in Mice. *Sci. Rep.* **2019**, *9*, 17312.
(179) Choi, S. Y.; Woo, H. M. CRISPRi-DCas12a: A DCas12a-Mediated CRISPR Interference for Repression of Multiple Genes and Metabolic Engineering in Cyanobacteria. *ACS Synth. Biol.* **2020**, *9*, 2351−2361.
(180) Kiattisewee, C.; Dong, C.; Fontana, J.; Sugianto, W.; Peralta-Yahya, P.; Carothers, J. M.; Zalatan, J. G. Portable Bacterial CRISPR Transcriptional Activation Enables Metabolic Engineering in *Pseudomonas putida*. *Metab. Eng.* **2021**, *66*, 283−295.
(181) Lian, J.; HamediRad, M.; Hu, S.; Zhao, H. Combinatorial Metabolic Engineering Using an Orthogonal Tri-Functional CRISPR System. *Nat. Commun.* **2017**, *8*, 1688.
(182) Dong, C.; Schultz, J. C.; Liu, W.; Lian, J.; Huang, L.; Xu, Z.; Zhao, H. Identification of Novel Metabolic Engineering Targets for S-Adenosyl-L-Methionine Production in *Saccharomyces cerevisiae* via Genome-Scale Engineering. *Metab. Eng.* **2021**, *66*, 319−327.
(183) Lalwani, M. A.; Zhao, E. M.; Avalos, J. L. Current and Future Modalities of Dynamic Control in Metabolic Engineering. *Curr. Opin. Biotechnol.* **2018**, *52*, 56−65.
(184) Ni, C.; Dinh, C. V.; Prather, K. L. J. Dynamic Control of Metabolism. *Annu. Rev. Chem. Biomol. Eng.* **2021**, *12*, 519−541.
(185) Soma, Y.; Tsuruno, K.; Wada, M.; Yokota, A.; Hanai, T. Metabolic Flux Redirection from a Central Metabolic Pathway toward a Synthetic Pathway Using a Metabolic Toggle Switch. *Metab. Eng.* **2014**, *23*, 175−184.
(186) Reizman, I. M. B.; Stenger, A. R.; Reisch, C. R.; Gupta, A.; Connors, N. C.; Prather, K. L. J. Improvement of Glucaric Acid Production in *E. coli* via Dynamic Control of Metabolic Fluxes. *Metab. Eng. Commun.* **2015**, *2*, 109−116.
(187) Cress, B. F.; Leitz, Q. D.; Kim, D. C.; Amore, T. D.; Suzuki, J. Y.; Linhardt, R. J.; Koffas, M. A. G. CRISPRi-Mediated Metabolic Engineering of *E. coli* for O-Methylated Anthocyanin Production. *Microb. Cell Factories* **2017**, *16*, 10.
(188) Dong, H.; Tao, W.; Zhang, Y.; Li, Y. Development of an Anhydrotetracycline-Inducible Gene Expression System for Solvent-Producing *Clostridium acetobutylicum*: A Useful Tool for Strain Engineering. *Metab. Eng.* **2012**, *14*, 59−67.
(189) Castillo-Hair, S. M.; Fujita, M.; Igoshin, O. A.; Tabor, J. J. An Engineered *B. subtilis* Inducible Promoter System with over 10 000-Fold Dynamic Range. *ACS Synth. Biol.* **2019**, *8*, 1673−1678.
(190) Paradise, E. M.; Kirby, J.; Chan, R.; Keasling, J. D. Redirection of Flux through the FPP Branch-Point in *Saccharomyces cerevisiae* by down-Regulating Squalene Synthase. *Biotechnol. Bioeng.* **2008**, *100*, 371−378.
(191) Xie, W.; Ye, L.; Lv, X.; Xu, H.; Yu, H. Sequential Control of Biosynthetic Pathways for Balanced Utilization of Metabolic Intermediates in *Saccharomyces cerevisiae*. *Metab. Eng.* **2015**, *28*, 8−18.
(192) Sathesh-Prabu, C.; Kim, D.; Lee, S. K. Metabolic Engineering of *Escherichia coli* for 2,3-Butanediol Production from Cellulosic Biomass by Using Glucose-Inducible Gene Expression System. *Bioresour. Technol.* **2020**, *309*, 123361.
(193) Love, C. A.; Lilley, P. E.; Dixon, N. E. Stable High-Copy-Number Bacteriophage λ Promoter Vectors for Overproduction of Proteins in *Escherichia coli*. *Gene* **1996**, *176*, 49−53.
(194) Elvin, C. M.; Thompson, P. R.; Argall, M. E.; Philip Hendr, N.; Stamford, P. J.; Lilley, P. E.; Dixon, N. E. Modified Bacteriophage Lambda Promoter Vectors for Overproduction of Proteins in *Escherichia coli*. *Gene* **1990**, *87*, 123−126.
(195) Zhou, L.; Niu, D.-D.; Tian, K.-M.; Chen, X.-Z.; Prior, B. A.; Shen, W.; Shi, G.-Y.; Singh, S.; Wang, Z.-X. Genetically Switched D-Lactate Production in *Escherichia coli*. *Metab. Eng.* **2012**, *14*, 560−568.
(196) Zhou, L.; Deng, C.; Cui, W.-J.; Liu, Z.-M.; Zhou, Z.-M. Efficient L-Alanine Production by a Thermo-Regulated Switch in *Escherichia coli*. *Appl. Biochem. Biotechnol.* **2016**, *178*, 324−337.
(197) Harder, B.-J.; Bettenbrock, K.; Klamt, S. Temperature-Dependent Dynamic Control of the TCA Cycle Increases Volumetric Productivity of Itaconic Acid Production by *Escherichia coli*. *Biotechnol. Bioeng.* **2018**, *115*, 156−164.
(198) Hwang, H. J.; Kim, J. W.; Ju, S. Y.; Park, J. H.; Lee, P. C. Application of an Oxygen-Inducible Nar Promoter System in Metabolic Engineering for Production of Biochemicals in *Escherichia coli*. *Biotechnol. Bioeng.* **2017**, *114*, 468−473.
(199) Weitzman, M.; Hahn, K. M. Optogenetic Approaches to Cell Migration and Beyond. *Curr. Opin. Cell Biol.* **2014**, *30*, 112−120.
(200) Tandar, S. T.; Senoo, S.; Toya, Y.; Shimizu, H. Optogenetic Switch for Controlling the Central Metabolic Flux of *Escherichia coli*. *Metab. Eng.* **2019**, *55*, 68−75.
(201) Zhao, E. M.; Zhang, Y.; Mehl, J.; Park, H.; Lalwani, M. A.; Toettcher, J. E.; Avalos, J. L. Optogenetic Regulation of Engineered Cellular Metabolism for Microbial Chemical Production. *Nature* **2018**, *555*, 683−687.
(202) Zhang, F.; Carothers, J. M.; Keasling, J. D. Design of a Dynamic Sensor-Regulator System for Production of Chemicals and Fuels Derived from Fatty Acids. *Nat. Biotechnol.* **2012**, *30*, 354−359.
(203) Wu, Y.; Chen, T.; Liu, Y.; Tian, R.; Lv, X.; Li, J.; Du, G.; Chen, J.; Ledesma-Amaro, R.; Liu, L. Design of a Programmable Biosensor-CRISPRi Genetic Circuits for Dynamic and Autonomous Dual-Control of Metabolic Flux in *Bacillus subtilis*. *Nucleic Acids Res.* **2020**, *48*, 996−1009.
(204) Yang, Y.; Lin, Y.; Wang, J.; Wu, Y.; Zhang, R.; Cheng, M.; Shen, X.; Wang, J.; Chen, Z.; Li, C.; et al. Sensor-Regulator and RNAi Based Bifunctional Dynamic Control Network for Engineered Microbial Synthesis. *Nat. Commun.* **2018**, *9*, 3043.
(205) Yang, X.; Liu, J.; Zhang, J.; Shen, Y.; Qi, Q.; Bao, X.; Hou, J. Quorum Sensing-Mediated Protein Degradation for Dynamic Metabolic Pathway Control in *Saccharomyces cerevisiae*. *Metab. Eng.* **2021**, *64*, 85−94.
(206) Dinh, C. V.; Prather, K. L. J. Development of an Autonomous and Bifunctional Quorum-Sensing Circuit for Metabolic Flux Control in Engineered *Escherichia coli*. *Proc. Natl. Acad. Sci. U. S. A.* **2019**, *116*, 25562−25568.
(207) Gill, R. T. Enabling Inverse Metabolic Engineering through Genomics. *Curr. Opin. Biotechnol.* **2003**, *14*, 484−490.
(208) Phaneuf, P. V.; Yurkovich, J. T.; Heckmann, D.; Wu, M.; Sandberg, T. E.; King, Z. A.; Tan, J.; Palsson, B. O.; Feist, A. M. Causal Mutations from Adaptive Laboratory Evolution Are Outlined by Multiple Scales of Genome Annotations and Condition-Specificity. *BMC Genomics* **2020**, *21*, 514.
(209) Phaneuf, P. V.; Gosting, D.; Palsson, B. O.; Feist, A. M. ALEdb 1.0: A Database of Mutations from Adaptive Laboratory Evolution Experimentation. *Nucleic Acids Res.* **2019**, *47*, D1164−D1171.
(210) Maurer, M. J.; Sutardja, L.; Pinel, D.; Bauer, S.; Muehlbauer, A. L.; Ames, T. D.; Skerker, J. M.; Arkin, A. P. Quantitative Trait Loci (QTL)-Guided Metabolic Engineering of a Complex Trait. *ACS Synth. Biol.* **2017**, *6*, 566−581.

https://doi.org/10.1021/acs.chemrev.2c00403
Chem. Rev. 2023, 123, 5521−5570

(211) Wang, Z.; Danziger, S. A.; Heavner, B. D.; Ma, S.; Smith, J. J.; Li, S.; Herricks, T.; Simeonidis, E.; Baliga, N. S.; Aitchison, J. D.; et al. Combining Inferred Regulatory and Reconstructed Metabolic Networks Enhances Phenotype Prediction in Yeast. *PLOS Comput. Biol.* **2017**, *13*, No. e1005489.
(212) Kim, M.; Park, B. G.; Kim, J.; Kim, J. Y.; Kim, B.-G. Exploiting Transcriptomic Data for Metabolic Engineering: Toward a Systematic Strain Design. *Curr. Opin. Biotechnol.* **2018**, *54*, 26−32.
(213) Trauchessec, M.; Jaquinod, M.; Bonvalot, A.; Brun, V.; Bruley, C.; Ropers, D.; de Jong, H.; Garin, J.; Bestel-Corre, G.; Ferro, M. Mass Spectrometry-Based Workflow for Accurate Quantification of *Escherichia coli* Enzymes: How Proteomics Can Play a Key Role in Metabolic Engineering. *Mol. Cell. Proteomics* **2014**, *13*, 954−968.
(214) Alonso-Gutierrez, J.; Kim, E.-M.; Batth, T. S.; Cho, N.; Hu, Q.; Chan, L. J. G.; Petzold, C. J.; Hillson, N. J.; Adams, P. D.; Keasling, J. D.; et al. Principal Component Analysis of Proteomics (PCAP) as a Tool to Direct Metabolic Engineering. *Metab. Eng.* **2015**, *28*, 123−133.
(215) Chang, D.; Islam, Z. U.; Zheng, J.; Zhao, J.; Cui, X.; Yu, Z. Inhibitor Tolerance and Bioethanol Fermentability of Levoglucosan-Utilizing *Escherichia coli* Were Enhanced by Overexpression of Stress-Responsive Gene YcfR: The Proteomics-Guided Metabolic Engineering. *Synth. Syst. Biotechnol.* **2021**, *6*, 384−395.
(216) Canelas, A. B.; ten Pierick, A.; Ras, C.; Seifar, R. M.; van Dam, J. C.; van Gulik, W. M.; Heijnen, J. J. Quantitative Evaluation of Intracellular Metabolite Extraction Techniques for Yeast Metabolomics. *Anal. Chem.* **2009**, *81*, 7379−7389.
(217) Ewald, J. C.; Heux, S.; Zamboni, N. High-Throughput Quantitative Metabolomics: Workflow for Cultivation, Quenching, and Analysis of Yeast in a Multiwell Format. *Anal. Chem.* **2009**, *81*, 3623−3629.
(218) Gold, N. D.; Gowen, C. M.; Lussier, F.-X.; Cautha, S. C.; Mahadevan, R.; Martin, V. J. J. Metabolic Engineering of a Tyrosine-Overproducing Yeast Platform Using Targeted Metabolomics. *Microb. Cell Factories* **2015**, *14*, 73.
(219) Willemsen, A. M.; Hendrickx, D. M.; Hoefsloot, H. C. J.; Hendriks, M. M. W. B.; Wahl, S. A.; Teusink, B.; Smilde, A. K.; van Kampen, A. H. C. MetDFBA: Incorporating Time-Resolved Metabolomics Measurements into Dynamic Flux Balance Analysis. *Mol. Biosyst.* **2015**, *11*, 137−145.
(220) Hashim, Z.; Teoh, S. T.; Bamba, T.; Fukusaki, E. Construction of a Metabolome Library for Transcription Factor-Related Single Gene Mutants of *Saccharomyces cerevisiae*. *J. Chromatogr. B* **2014**, *966*, 83−92.
(221) Kim, H. U.; Kim, T. Y.; Lee, S. Y. Metabolic Flux Analysis and Metabolic Engineering of Microorganisms. *Mol. BioSyst* **2008**, *4*, 113−120.
(222) Orth, J. D.; Thiele, I.; Palsson, B. Ø. What Is Flux Balance Analysis? *Nat. Biotechnol.* **2010**, *28*, 245−248.
(223) Feist, A. M.; Henry, C. S.; Reed, J. L.; Krummenacker, M.; Joyce, A. R.; Karp, P. D.; Broadbelt, L. J.; Hatzimanikatis, V.; Palsson, B. Ø. A Genome-scale Metabolic Reconstruction for *Escherichia coli* K-12 MG1655 that Accounts for 1260 ORFs and Thermodynamic Information. *Mol. Syst. Biol.* **2007**, *3*, 121.
(224) Henry, C. S.; Broadbelt, L. J.; Hatzimanikatis, V. Thermodynamics-Based Metabolic Flux Analysis. *Biophys. J.* **2007**, *92*, 1792−1805.
(225) Niedenführ, S.; Wiechert, W.; Nöh, K. How to Measure Metabolic Fluxes: A Taxonomic Guide for ^{13}C Fluxomics. *Curr. Opin. Biotechnol.* **2015**, *34*, 82−90.
(226) Ando, D.; Garcia Martin, H. Two-Scale 13C Metabolic Flux Analysis for Metabolic Engineering. In *Synthetic Metabolic Pathways*; Methods in Molecular Biology; Jensen, M. K., Keasling, J. D., Eds.; ; Springer: New York, 2018; Vol. *1671*, pp 333−352.
(227) Gopalakrishnan, S.; Maranas, C. D. 13C Metabolic Flux Analysis at a Genome-Scale. *Metab. Eng.* **2015**, *32*, 12−22.
(228) Ghosh, A.; Ando, D.; Gin, J.; Runguphan, W.; Denby, C.; Wang, G.; Baidoo, E. E. K.; Shymansky, C.; Keasling, J. D.; García Martín, H. 13C Metabolic Flux Analysis for Systematic Metabolic Engineering of *S. cerevisiae* for Overproduction of Fatty Acids. *Front. Bioeng. Biotechnol.* **2016**, *4*, 3389.
(229) García Martín, H.; Kumar, V. S.; Weaver, D.; Ghosh, A.; Chubukov, V.; Mukhopadhyay, A.; Arkin, A.; Keasling, J. D. A Method to Constrain Genome-Scale Models with ^{13}C Labeling Data. *PLOS Comput. Biol.* **2015**, *11*, No. e1004363.
(230) Toya, Y.; Shimizu, H. Flux Analysis and Metabolomics for Systematic Metabolic Engineering of Microorganisms. *Biotechnol. Adv.* **2013**, *31*, 818−826.
(231) Liu, Y.; Liu, Y.; Wang, M. Design, Optimization and Application of Small Molecule Biosensor in Metabolic Engineering. *Front. Microbiol.* **2017**, *8*, 2012.
(232) Morgan, S.-A.; Nadler, D. C.; Yokoo, R.; Savage, D. F. Biofuel Metabolic Engineering with Biosensors. *Curr. Opin. Chem. Biol.* **2016**, *35*, 150−158.
(233) Liu, D.; Evans, T.; Zhang, F. Applications and Advances of Metabolite Biosensors for Metabolic Engineering. *Metab. Eng.* **2015**, *31*, 35−43.
(234) Shi, S.; Ang, E. L.; Zhao, H. *In Vivo* Biosensors: Mechanisms, Development, and Applications. *J. Ind. Microbiol. Biotechnol.* **2018**, *45*, 491−516.
(235) Ding, N.; Zhou, S.; Deng, Y. Transcription-Factor-Based Biosensor Engineering for Applications in Synthetic Biology. *ACS Synth. Biol.* **2021**, *10*, 911−922.
(236) Koch, M.; Pandi, A.; Borkowski, O.; Batista, A. C.; Faulon, J.-L. Custom-Made Transcriptional Biosensors for Metabolic Engineering. *Curr. Opin. Biotechnol.* **2019**, *59*, 78−84.
(237) Wang, G.; Øzmerih, S.; Guerreiro, R.; Meireles, A. C.; Carolas, A.; Milne, N.; Jensen, M. K.; Ferreira, B. S.; Borodina, I. Improvement of Cis,Cis-Muconic Acid Production in *Saccharomyces cerevisiae* through Biosensor-Aided Genome Engineering. *ACS Synth. Biol.* **2020**, *9*, 634−646.
(238) Herud-Sikimić, O.; Stiel, A. C.; Kolb, M.; Shanmugaratnam, S.; Berendzen, K. W.; Feldhaus, C.; Höcker, B.; Jürgens, G. A Biosensor for the Direct Visualization of Auxin. *Nature* **2021**, *592*, 768−772.
(239) Romero-Galindo, R.; Torres-Pacheco, I.; Guevara-Gonzalez, G. R.; Contreras-Medina, M. L. Biosensors Used for Quantification of Nitrates in Plants. *J. Sens.* **2016**, *2016*, No. e1630695.
(240) Wu, W.; Zhang, L.; Yao, L.; Tan, X.; Liu, X.; Lu, X. Genetically Assembled Fluorescent Biosensor for *in Situ* Detection of Bio-Synthesized Alkanes. *Sci. Rep.* **2015**, *5*, 10907.
(241) Hossain, G. S.; Saini, M.; Miyake, R.; Ling, H.; Chang, M. W. Genetic Biosensor Design for Natural Product Biosynthesis in Microorganisms. *Trends Biotechnol.* **2020**, *38*, 797−810.
(242) Boada, Y.; Vignoni, A.; Picó, J.; Carbonell, P. Extended Metabolic Biosensor Design for Dynamic Pathway Regulation of Cell Factories. *iScience* **2020**, *23*, 101305.
(243) Reed, B.; Blazeck, J.; Alper, H. Evolution of an Alkane-Inducible Biosensor for Increased Responsiveness to Short-Chain Alkanes. *J. Biotechnol.* **2012**, *158*, 75−79.
(244) Javanpour, A. A.; Liu, C. C. Evolving Small-Molecule Biosensors with Improved Performance and Reprogrammed Ligand Preference Using OrthoRep. *ACS Synth. Biol.* **2021**, *10*, 2705−2714.
(245) Jensen, E. D.; Ambri, F.; Bendtsen, M. B.; Javanpour, A. A.; Liu, C. C.; Jensen, M. K.; Keasling, J. D. Integrating Continuous Hypermutation with High-Throughput Screening for Optimization of Cis,Cis-Muconic Acid Production in Yeast. *Microb. Biotechnol.* **2021**, *14*, 2617−2626.
(246) Ameen, S.; Ahmad, M.; Mohsin, M.; Qureshi, M. I.; Ibrahim, M. M.; Abdin, M. Z.; Ahmad, A. Designing, Construction and Characterization of Genetically Encoded FRET-Based Nanosensor for Real Time Monitoring of Lysine Flux in Living Cells. *J. Nanobiotechnology* **2016**, *14*, 49.
(247) Peroza, E. A.; Ewald, J. C.; Parakkal, G.; Skotheim, J. M.; Zamboni, N. A Genetically Encoded Förster Resonance Energy Transfer Sensor for Monitoring *in Vivo* Trehalose-6-Phosphate Dynamics. *Anal. Biochem.* **2015**, *474*, 1−7.

https://doi.org/10.1021/acs.chemrev.2c00403
Chem. Rev. 2023, 123, 5521−5570

(248) Otten, J.; Tenhaef, N.; Jansen, R. P.; Döbber, J.; Jungbluth, L.; Noack, S.; Oldiges, M.; Wiechert, W.; Pohl, M. A FRET-Based Biosensor for the Quantification of Glucose in Culture Supernatants of ML Scale Microbial Cultivations. *Microb. Cell Factories* **2019**, *18*, 143.
(249) Imani, M.; Mohajeri, N.; Rastegar, M.; Zarghami, N. Recent Advances in FRET-Based Biosensors for Biomedical Applications. *Anal. Biochem.* **2021**, *630*, 114323.
(250) Zhang, X.; Hu, Y.; Yang, X.; Tang, Y.; Han, S.; Kang, A.; Deng, H.; Chi, Y.; Zhu, D.; Lu, Y. FÖrster Resonance Energy Transfer (FRET)-Based Biosensors for Biological Applications. *Biosens. Bioelectron.* **2019**, *138*, 111314.
(251) Serganov, A.; Nudler, E. A Decade of Riboswitches. *Cell* **2013**, *152*, 17−24.
(252) Yan, Q.; Fong, S. S. Biosensors for Metabolic Engineering. In *Systems Biology Application in Synthetic Biology*; Singh, S., Ed.; Springer India: New Delhi, 2016; pp 53−70.
(253) Lenkeit, F.; Eckert, I.; Hartig, J. S.; Weinberg, Z. Discovery and Characterization of a Fourth Class of Guanidine Riboswitches. *Nucleic Acids Res.* **2020**, *48*, 12889−12899.
(254) Bordbar, A.; Monk, J. M.; King, Z. A.; Palsson, B. O. Constraint-Based Models Predict Metabolic and Associated Cellular Functions. *Nat. Rev. Genet.* **2014**, *15*, 107−120.
(255) Lu, H.; Li, F.; Sánchez, B. J.; Zhu, Z.; Li, G.; Domenzain, I.; Marcišauskas, S.; Anton, P. M.; Lappa, D.; Lieven, C.; et al. A Consensus S. Cerevisiae Metabolic Model Yeast8 and Its Ecosystem for Comprehensively Probing Cellular Metabolism. *Nat. Commun.* **2019**, *10*, 3586.
(256) Li, F.; Chen, Y.; Qi, Q.; Wang, Y.; Yuan, L.; Huang, M.; Elsemman, I. E.; Feizi, A.; Kerkhoven, E. J.; Nielsen, J. Improving Recombinant Protein Production by Yeast through Genome-Scale Modeling Using Proteome Constraints. *Nat. Commun.* **2022**, *13*, 2969.
(257) Chowdhury, A.; Zomorrodi, A. R.; Maranas, C. D. K-OptForce: Integrating Kinetics with Flux Balance Analysis for Strain Design. *PLoS Comput. Biol.* **2014**, *10*, No. e1003487.
(258) Fell, D. A. Metabolic Control Analysis: A Survey of Its Theoretical and Experimental Development. *Biochem. J.* **1992**, *286*, 313−330.
(259) Fell, D. A. Increasing the Flux in Metabolic Pathways: A Metabolic Control Analysis Perspective. *Biotechnol. Bioeng.* **1998**, *58*, 121−124.
(260) Tran, L. M.; Rizk, M. L.; Liao, J. C. Ensemble Modeling of Metabolic Networks. *Biophys. J.* **2008**, *95*, 5606−5617.
(261) Khodayari, A.; Maranas, C. D. A Genome-Scale *Escherichia coli* Kinetic Metabolic Model k-Ecoli457 Satisfying Flux Data for Multiple Mutant Strains. *Nat. Commun.* **2016**, *7*, 13806.
(262) Chakrabarti, A.; Miskovic, L.; Soh, K. C.; Hatzimanikatis, V. Towards Kinetic Modeling of Genome-Scale Metabolic Networks without Sacrificing Stoichiometric, Thermodynamic and Physiological Constraints. *Biotechnol. J.* **2013**, *8*, 1043−1057.
(263) Andreozzi, S.; Chakrabarti, A.; Soh, K. C.; Burgard, A.; Yang, T. H.; Van Dien, S.; Miskovic, L.; Hatzimanikatis, V. Identification of Metabolic Engineering Targets for the Enhancement of 1,4-Butanediol Production in Recombinant *E. coli* Using Large-Scale Kinetic Models. *Metab. Eng.* **2016**, *35*, 148−159.
(264) Volk, M. J.; Lourentzou, I.; Mishra, S.; Vo, L. T.; Zhai, C.; Zhao, H. Biosystems Design by Machine Learning. *ACS Synth. Biol.* **2020**, *9*, 1514−1533.
(265) Xu, P.; Rizzoni, E. A.; Sul, S.-Y.; Stephanopoulos, G. Improving Metabolic Pathway Efficiency by Statistical Model-Based Multivariate Regulatory Metabolic Engineering. *ACS Synth. Biol.* **2017**, *6*, 148−158.
(266) HamediRad, M.; Chao, R.; Weisberg, S.; Lian, J.; Sinha, S.; Zhao, H. Towards a Fully Automated Algorithm Driven Platform for Biosystems Design. *Nat. Commun.* **2019**, *10*, 5150.
(267) Carbonell, P.; Jervis, A. J.; Robinson, C. J.; Yan, C.; Dunstan, M.; Swainston, N.; Vinaixa, M.; Hollywood, K. A.; Currin, A.; Rattray, N. J. W.; et al. An Automated Design-Build-Test-Learn Pipeline for Enhanced Microbial Production of Fine Chemicals. *Commun. Biol.* **2018**, *1*, 1−10.
(268) Jervis, A. J.; Carbonell, P.; Vinaixa, M.; Dunstan, M. S.; Hollywood, K. A.; Robinson, C. J.; Rattray, N. J. W.; Yan, C.; Swainston, N.; Currin, A.; et al. Machine Learning of Designed Translational Control Allows Predictive Pathway Optimization in *Escherichia coli*. *ACS Synth. Biol.* **2019**, *8*, 127−136.
(269) Opgenorth, P.; Costello, Z.; Okada, T.; Goyal, G.; Chen, Y.; Gin, J.; Benites, V.; de Raad, M.; Northen, T. R.; Deng, K.; et al. Lessons from Two Design−Build−Test−Learn Cycles of Dodecanol Production in *Escherichia coli* Aided by Machine Learning. *ACS Synth. Biol.* **2019**, *8*, 1337−1351.
(270) Zhou, Y.; Li, G.; Dong, J.; Xing, X.; Dai, J.; Zhang, C. MiYA, an Efficient Machine-Learning Workflow in Conjunction with the YeastFab Assembly Strategy for Combinatorial Optimization of Heterologous Metabolic Pathways in *Saccharomyces cerevisiae*. *Metab. Eng.* **2018**, *47*, 294−302.
(271) Zhang, J.; Petersen, S. D.; Radivojevic, T.; Ramirez, A.; Pérez-Manríquez, A.; Abeliuk, E.; Sánchez, B. J.; Costello, Z.; Chen, Y.; Fero, M. J.; et al. Combining Mechanistic and Machine Learning Models for Predictive Engineering and Optimization of Tryptophan Metabolism. *Nat. Commun.* **2020**, *11*, 4880.
(272) Radivojević, T.; Costello, Z.; Workman, K.; Garcia Martin, H. A Machine Learning Automated Recommendation Tool for Synthetic Biology. *Nat. Commun.* **2020**, *11*, 4879.
(273) Vaishnav, E. D.; de Boer, C. G.; Molinet, J.; Yassour, M.; Fan, L.; Adiconis, X.; Thompson, D. A.; Levin, J. Z.; Cubillos, F. A.; Regev, A. The Evolution, Evolvability and Engineering of Gene Regulatory DNA. *Nature* **2022**, *603*, 455−463.
(274) Alley, E. C.; Khimulya, G.; Biswas, S.; AlQuraishi, M.; Church, G. M. Unified Rational Protein Engineering with Sequence-Based Deep Representation Learning. *Nat. Methods* **2019**, *16*, 1315−1322.
(275) Sanderson, T.; Bileschi, M. L.; Belanger, D.; Colwell, L. J. ProteInfer: Deep Networks for Protein Functional Inference. *bioRxiv* 2021, DOI: 10.1101/2021.09.20.461077.
(276) Kim, G. B.; Gao, Y.; Palsson, B. O.; Lee, S. Y. DeepTFactor: A Deep Learning-Based Tool for the Prediction of Transcription Factors. *Proc. Natl. Acad. Sci. U. S. A.* **2021**, *118*, No. e2021171118.
(277) Dale, J. M.; Popescu, L.; Karp, P. D. Machine Learning Methods for Metabolic Pathway Prediction. *BMC Bioinformatics* **2010**, *11*, 15.
(278) Heckmann, D.; Lloyd, C. J.; Mih, N.; Ha, Y.; Zielinski, D. C.; Haiman, Z. B.; Desouki, A. A.; Lercher, M. J.; Palsson, B. O. Machine Learning Applied to Enzyme Turnover Numbers Reveals Protein Structural Correlates and Improves Metabolic Models. *Nat. Commun.* **2018**, *9*, 5252.
(279) Ma, J.; Yu, M. K.; Fong, S.; Ono, K.; Sage, E.; Demchak, B.; Sharan, R.; Ideker, T. Using Deep Learning to Model the Hierarchical Structure and Function of a Cell. *Nat. Methods* **2018**, *15*, 290−298.
(280) Jansen, R.; Yu, H.; Greenbaum, D.; Kluger, Y.; Krogan, N. J.; Chung, S.; Emili, A.; Snyder, M.; Greenblatt, J. F.; Gerstein, M. A Bayesian Networks Approach for Predicting Protein-Protein Interactions from Genomic Data. *Science* **2003**, *302*, 449−453.
(281) Zelezniak, A.; Vowinckel, J.; Capuano, F.; Messner, C. B.; Demichev, V.; Polowsky, N.; Mülleder, M.; Kamrad, S.; Klaus, B.; Keller, M. A.; et al. Machine Learning Predicts the Yeast Metabolome from the Quantitative Proteome of Kinase Knockouts. *Cell Syst.* **2018**, *7*, 269−283.
(282) Czajka, J. J.; Oyetunde, T.; Tang, Y. J. Integrated Knowledge Mining, Genome-Scale Modeling, and Machine Learning for Predicting *Yarrowia lipolytica* Bioproduction. *Metab. Eng.* **2021**, *67*, 227−236.
(283) Oyetunde, T.; Liu, D.; Martin, H. G.; Tang, Y. J. Machine Learning Framework for Assessment of Microbial Factory Performance. *PLoS One* **2019**, *14*, No. e0210558.
(284) Werpy, T.; Petersen, G. *Top Value Added Chemicals from Biomass: Results of Screening for Potential Candidates from Sugars and Synthesis Gas*; DOE/GO-102004-1992; National Renewable Energy Lab.: Golden, CO, 2004; Vol. *I*, DOI: DOI: 10.2172/15008859.

(285) Harder, B.-J.; Bettenbrock, K.; Klamt, S. Model-Based Metabolic Engineering Enables High Yield Itaconic Acid Production by *Escherichia coli*. *Metab. Eng.* **2016**, *38*, 29−37.
(286) Geiser, E.; Przybilla, S. K.; Friedrich, A.; Buckel, W.; Wierckx, N.; Blank, L. M.; Bölker, M. Ustilago Maydis Produces Itaconic Acid via the Unusual Intermediate Trans-Aconitate. *Microb. Biotechnol.* **2016**, *9*, 116−126.
(287) Elmore, J. R.; Dexter, G. N.; Salvachúa, D.; Martinez-Baird, J.; Hatmaker, E. A.; Huenemann, J. D.; Klingeman, D. M.; Peabody, G. L.; Peterson, D. J.; Singer, C.; et al. Production of Itaconic Acid from Alkali Pretreated Lignin by Dynamic Two Stage Bioconversion. *Nat. Commun.* **2021**, *12*, 2261.
(288) Zhao, C.; Cui, Z.; Zhao, X.; Zhang, J.; Zhang, L.; Tian, Y.; Qi, Q.; Liu, J. Enhanced Itaconic Acid Production in *Yarrowia lipolytica* via Heterologous Expression of a Mitochondrial Transporter MTT. *Appl. Microbiol. Biotechnol.* **2019**, *103*, 2181−2192.
(289) Hanko, E. K. R.; Minton, N. P.; Malys, N. A Transcription Factor-Based Biosensor for Detection of Itaconic Acid. *ACS Synth. Biol.* **2018**, *7*, 1436−1446.
(290) Blazeck, J.; Hill, A.; Jamoussi, M.; Pan, A.; Miller, J.; Alper, H. S. Metabolic Engineering of *Yarrowia lipolytica* for Itaconic Acid Production. *Metab. Eng.* **2015**, *32*, 66−73.
(291) Lu, K. W.; Wang, C. T.; Chang, H.; Wang, R. S.; Shen, C. R. Overcoming Glutamate Auxotrophy in *Escherichia coli* Itaconate Overproducer by the Weimberg Pathway. *Metab. Eng. Commun.* **2021**, *13*, No. e00190.
(292) Cheng, J.; Li, J.; Zheng, L. Achievements and Perspectives in 1,4-Butanediol Production from Engineered Microorganisms. *J. Agric. Food Chem.* **2021**, *69*, 10480−10485.
(293) Wu, M.-Y.; Sung, L.-Y.; Li, H.; Huang, C.-H.; Hu, Y.-C. Combining CRISPR and CRISPRi Systems for Metabolic Engineering of *E. coli* and 1,4-BDO Biosynthesis. *ACS Synth. Biol.* **2017**, *6*, 2350−2361.
(294) Hwang, H. J.; Park, J. H.; Kim, J. H.; Kong, M. K.; Kim, J. W.; Park, J. W.; Cho, K. M.; Lee, P. C. Engineering of a Butyraldehyde Dehydrogenase of *Clostridium saccharoperbut*ylacetonicum to Fit an Engineered 1,4-Butanediol Pathway in *Escherichia coli*. *Biotechnol. Bioeng.* **2014**, *111*, 1374−1384.
(295) Burgard, A.; Burk, M. J.; Osterhout, R.; Van Dien, S.; Yim, H. Development of a Commercial Scale Process for Production of 1,4-Butanediol from Sugar. *Curr. Opin. Biotechnol.* **2016**, *42*, 118−125.
(296) Panjapakkul, W.; El-Halwagi, M. M. Technoeconomic Analysis of Alternative Pathways of Isopropanol Production. *ACS Sustain. Chem. Eng.* **2018**, *6*, 10260−10272.
(297) Liew, F. E.; Nogle, R.; Abdalla, T.; Rasor, B. J.; Canter, C.; Jensen, R. O.; Wang, L.; Strutz, J.; Chirania, P.; De Tissera, S.; et al. Carbon-Negative Production of Acetone and Isopropanol by Gas Fermentation at Industrial Pilot Scale. *Nat. Biotechnol.* **2022**, *40*, 335−344.
(298) Jojima, T.; Inui, M.; Yukawa, H. Production of Isopropanol by Metabolically Engineered *Escherichia coli*. *Appl. Microbiol. Biotechnol.* **2008**, *77*, 1219−1224.
(299) Moon, H. G.; Jang, Y.-S.; Cho, C.; Lee, J.; Binkley, R.; Lee, S. Y. One Hundred Years of Clostridial Butanol Fermentation. *FEMS Microbiol. Lett.* **2016**, *363*, fnw001.
(300) Chae, T. U.; Choi, S. Y.; Ryu, J. Y.; Lee, S. Y. Production of Ethylene Glycol from Xylose by Metabolically Engineered *Escherichia coli*. *AIChE J.* **2018**, *64*, 4193−4200.
(301) Emptage, M.; Haynie, S. L.; Laffend, L. A.; Pucci, J. P.; Whited, G. Process for the Biological Production of 1,3-Propanediol with High Titer. US Patent 6514733B1, February 4, 2003.
(302) Nakamura, C. E.; Whited, G. M. Metabolic Engineering for the Microbial Production of 1,3-Propanediol. *Curr. Opin. Biotechnol.* **2003**, *14*, 454−459.
(303) Niu, W.; Kramer, L.; Mueller, J.; Liu, K.; Guo, J. Metabolic Engineering of *Escherichia coli* for the *De Novo* Stereospecific Biosynthesis of 1,2-Propanediol through Lactic Acid. *Metab. Eng. Commun.* **2019**, *8*, No. e00082.
(304) Niu, W.; Guo, J. Stereospecific Microbial Conversion of Lactic Acid into 1,2-Propanediol. *ACS Synth. Biol.* **2015**, *4*, 378−382.
(305) Kim, T.; Stogios, P. J.; Khusnutdinova, A. N.; Nemr, K.; Skarina, T.; Flick, R.; Joo, J. C.; Mahadevan, R.; Savchenko, A.; Yakunin, A. F. Rational Engineering of 2-Deoxyribose-5-Phosphate Aldolases for the Biosynthesis of (*R*)-1,3-Butanediol. *J. Biol. Chem.* **2020**, *295*, 597−609.
(306) Lee, Y.-G.; Seo, J.-H. Production of 2,3-Butanediol from Glucose and Cassava Hydrolysates by Metabolically Engineered Industrial Polyploid *Saccharomyces cerevisiae*. *Biotechnol. Biofuels* **2019**, *12*, 204.
(307) Wang, J.; Li, C.; Zou, Y.; Yan, Y. Bacterial Synthesis of C3-C5 Diols via Extending Amino Acid Catabolism. *Proc. Natl. Acad. Sci. U. S. A.* **2020**, *117*, 19159−19167.
(308) Chae, T. U.; Kim, W. J.; Choi, S.; Park, S. J.; Lee, S. Y. Metabolic Engineering of *Escherichia coli* for the Production of 1,3-Diaminopropane, a Three Carbon Diamine. *Sci. Rep.* **2015**, *5*, 13040.
(309) Noh, M.; Yoo, S. M.; Kim, W. J.; Lee, S. Y. Gene Expression Knockdown by Modulating Synthetic Small RNA Expression in *Escherichia coli*. *Cell Syst.* **2017**, *5*, 418−426.
(310) Kim, H. T.; Baritugo, K.-A.; Oh, Y. H.; Hyun, S. M.; Khang, T. U.; Kang, K. H.; Jung, S. H.; Song, B. K.; Park, K.; Kim, I.-K.; et al. Metabolic Engineering of Corynebacterium Glutamicum for the High-Level Production of Cadaverine That Can Be Used for the Synthesis of Biopolyamide 510. *ACS Sustain. Chem. Eng.* **2018**, *6*, 5296−5305.
(311) Jiang, X.-R.; Yan, X.; Yu, L.-P.; Liu, X.-Y.; Chen, G.-Q. Hyperproduction of 3-Hydroxypropionate by *Halomonas bluephagenesis*. *Nat. Commun.* **2021**, *12*, 1513.
(312) Zhao, P.; Ma, C.; Xu, L.; Tian, P. Exploiting Tandem Repetitive Promoters for High-Level Production of 3-Hydroxypropionic Acid. *Appl. Microbiol. Biotechnol.* **2019**, *103*, 4017−4031.
(313) Tsuge, Y.; Kato, N.; Yamamoto, S.; Suda, M.; Jojima, T.; Inui, M. Metabolic Engineering of *Corynebacterium glutamicum* for Hyperproduction of Polymer-Grade L- and D-Lactic Acid. *Appl. Microbiol. Biotechnol.* **2019**, *103*, 3381−3391.
(314) Deng, Y.; Ma, N.; Zhu, K.; Mao, Y.; Wei, X.; Zhao, Y. Balancing the Carbon Flux Distributions between the TCA Cycle and Glyoxylate Shunt to Produce Glycolate at High Yield and Titer in *Escherichia coli*. *Metab. Eng.* **2018**, *46*, 28−34.
(315) Song, C. W.; Kim, J. W.; Cho, I. J.; Lee, S. Y. Metabolic Engineering of *Escherichia coli* for the Production of 3-Hydroxypropionic Acid and Malonic Acid through β-Alanine Route. *ACS Synth. Biol.* **2016**, *5*, 1256−1263.
(316) Gao, C.; Yang, X.; Wang, H.; Rivero, C. P.; Li, C.; Cui, Z.; Qi, Q.; Lin, C. S. K. Robust Succinic Acid Production from Crude Glycerol Using Engineered *Yarrowia lipolytica*. *Biotechnol. Biofuels* **2016**, *9*, 179.
(317) Han, T.; Kim, G. B.; Lee, S. Y. Glutaric Acid Production by Systems Metabolic Engineering of an L-Lysine−Overproducing *Corynebacterium glutamicum*. *Proc. Natl. Acad. Sci. U. S. A.* **2020**, *117*, 30328−30334.
(318) Chang, P.; Chen, G. S.; Chu, H.-Y.; Lu, K. W.; Shen, C. R. Engineering Efficient Production of Itaconic Acid from Diverse Substrates in *Escherichia coli*. *J. Biotechnol.* **2017**, *249*, 73−81.
(319) Lee, H.-N.; Shin, W.-S.; Seo, S.-Y.; Choi, S.-S.; Song, J.-S.; Kim, J.-Y.; Park, J.-H.; Lee, D.; Kim, S. Y.; Lee, S. J.; et al. Corynebacterium Cell Factory Design and Culture Process Optimization for Muconic Acid Biosynthesis. *Sci. Rep.* **2018**, *8*, 18041.
(320) Becker, J.; Zelder, O.; Häfner, S.; Schröder, H.; Wittmann, C. From Zero to Hero−Design-Based Systems Metabolic Engineering of *Corynebacterium glutamicum* for L-Lysine Production. *Metab. Eng.* **2011**, *13*, 159−168.
(321) Hale, V.; Keasling, J. D.; Renninger, N.; Diagana, T. T. Microbially Derived Artemisinin: A Biotechnology Solution to the Global Problem of Access to Affordable Antimalarial Drugs. *Am. J. Trop. Med. Hyg.* **2007**, *77*, 198−202.

https://doi.org/10.1021/acs.chemrev.2c00403
Chem. Rev. 2023, 123, 5521−5570

(322) Paddon, C. J.; Keasling, J. D. Semi-Synthetic Artemisinin: A Model for the Use of Synthetic Biology in Pharmaceutical Development. *Nat. Rev. Microbiol.* **2014**, *12*, 355−367.
(323) Ro, D.-K.; Paradise, E. M.; Ouellet, M.; Fisher, K. J.; Newman, K. L.; Ndungu, J. M.; Ho, K. A.; Eachus, R. A.; Ham, T. S.; Kirby, J.; et al. Production of the Antimalarial Drug Precursor Artemisinic Acid in Engineered Yeast. *Nature* **2006**, *440*, 940−943.
(324) Lenihan, J. R.; Tsuruta, H.; Diola, D.; Renninger, N. S.; Regentin, R. Developing an Industrial Artemisinic Acid Fermentation Process to Support the Cost-Effective Production of Antimalarial Artemisinin-Based Combination Therapies. *Biotechnol. Prog.* **2008**, *24*, 1026−1032.
(325) Westfall, P. J.; Pitera, D. J.; Lenihan, J. R.; Eng, D.; Woolard, F. X.; Regentin, R.; Horning, T.; Tsuruta, H.; Melis, D. J.; Owens, A.; et al. Production of Amorphadiene in Yeast, and Its Conversion to Dihydroartemisinic Acid, Precursor to the Antimalarial Agent Artemisinin. *Proc. Natl. Acad. Sci. U. S. A.* **2012**, *109*, No. E111.
(326) Zhang, H.; Im, S.-C.; Waskell, L. Cytochrome B5 Increases the Rate of Product Formation by Cytochrome P450 2B4 and Competes with Cytochrome P450 Reductase for a Binding Site on Cytochrome P450 2B4. *J. Biol. Chem.* **2007**, *282*, 29766−29776.
(327) Paddon, C. J.; Westfall, P. J.; Pitera, D. J.; Benjamin, K.; Fisher, K.; McPhee, D.; Leavell, M. D.; Tai, A.; Main, A.; Eng, D.; et al. High-Level Semi-Synthetic Production of the Potent Antimalarial Artemisinin. *Nature* **2013**, *496*, 528−532.
(328) Xue, Z.; Sharpe, P. L.; Hong, S.-P.; Yadav, N. S.; Xie, D.; Short, D. R.; Damude, H. G.; Rupert, R. A.; Seip, J. E.; Wang, J.; et al. Production of Omega-3 Eicosapentaenoic Acid by Metabolic Engineering of *Yarrowia lipolytica*. *Nat. Biotechnol.* **2013**, *31*, 734−740.
(329) Liu, Y.; Liu, Q.; Krivoruchko, A.; Khoomrung, S.; Nielsen, J. Engineering Yeast Phospholipid Metabolism for De Novo Oleoylethanolamide Production. *Nat. Chem. Biol.* **2020**, *16*, 197−205.
(330) Murakami, S.; Shimamoto, T.; Nagano, H.; Tsuruno, M.; Okuhara, H.; Hatanaka, H.; Tojo, H.; Kodama, Y.; Funato, K. Producing Human Ceramide-NS by Metabolic Engineering Using Yeast *Saccharomyces cerevisiae*. *Sci. Rep.* **2015**, *5*, 16319.
(331) Lim, C. G.; Fowler, Z. L.; Hueller, T.; Schaffer, S.; Koffas, M. A. G. High-Yield Resveratrol Production in Engineered *Escherichia coli*. *Appl. Environ. Microbiol.* **2011**, *77*, 3451−3460.
(332) Feng, C.; Chen, J.; Ye, W.; Liao, K.; Wang, Z.; Song, X.; Qiao, M. Synthetic Biology-Driven Microbial Production of Resveratrol: Advances and Perspectives. *Front. Bioeng. Biotechnol.* **2022**, *10*, 833920.
(333) Kang, W.; Ma, T.; Liu, M.; Qu, J.; Liu, Z.; Zhang, H.; Shi, B.; Fu, S.; Ma, J.; Lai, L. T. F.; et al. Modular Enzyme Assembly for Enhanced Cascade Biocatalysis and Metabolic Flux. *Nat. Commun.* **2019**, *10*, 4248.
(334) Li, C.; Swofford, C. A.; Sinskey, A. J. Modular Engineering for Microbial Production of Carotenoids. *Metab. Eng. Commun.* **2020**, *10*, No. e00118.
(335) Larroude, M.; Celinska, E.; Back, A.; Thomas, S.; Nicaud, J.-M.; Ledesma-Amaro, R. A Synthetic Biology Approach to Transform *Yarrowia lipolytica* into a Competitive Biotechnological Producer of β-Carotene. *Biotechnol. Bioeng.* **2018**, *115*, 464−472.
(336) Park, S. Y.; Binkley, R. M.; Kim, W. J.; Lee, M. H.; Lee, S. Y. Metabolic Engineering of *Escherichia coli* for High-Level Astaxanthin Production with High Productivity. *Metab. Eng.* **2018**, *49*, 105−115.
(337) Ko, Y. J.; Kim, M.; You, S. K.; Shin, S. K.; Chang, J.; Choi, H. J.; Jeong, W.-Y.; Lee, M.-E.; Hwang, D.-H.; Han, S. O. Animal-Free Heme Production for Artificial Meat in *Corynebacterium glutamicum* via Systems Metabolic and Membrane Engineering. *Metab. Eng.* **2021**, *66*, 217−228.
(338) Ajikumar, P. K.; Xiao, W.-H.; Tyo, K. E. J.; Wang, Y.; Simeon, F.; Leonard, E.; Mucha, O.; Phon, T. H.; Pfeifer, B.; Stephanopoulos, G. Isoprenoid Pathway Optimization for Taxol Precursor Overproduction in *Escherichia coli*. *Science* **2010**, *330*, 70−74.
(339) Mutanda, I.; Li, J.; Xu, F.; Wang, Y. Recent Advances in Metabolic Engineering, Protein Engineering, and Transcriptome-Guided Insights Toward Synthetic Production of Taxol. *Front. Bioeng. Biotechnol.* **2021**, *9*, 632269.
(340) Luo, X.; Reiter, M. A.; d'Espaux, L.; Wong, J.; Denby, C. M.; Lechner, A.; Zhang, Y.; Grzybowski, A. T.; Harth, S.; Lin, W.; et al. Complete Biosynthesis of Cannabinoids and Their Unnatural Analogues in Yeast. *Nature* **2019**, *567*, 123−126.
(341) Zhang, J.; Hansen, L. G.; Gudich, O.; Viehrig, K.; Lassen, L. M. M.; Schrübbers, L.; Adhikari, K. B.; Rubaszka, P.; Carrasquer-Alvarez, E.; Chen, L.; et al. A Microbial Supply Chain for Production of the Anti-Cancer Drug Vinblastine. *Nature* **2022**, *609*, 341−347.
(342) Liu, Z.; Wang, K.; Chen, Y.; Tan, T.; Nielsen, J. Third-Generation Biorefineries as the Means to Produce Fuels and Chemicals from CO2. *Nat. Catal.* **2020**, *3*, 274−288.
(343) Timilsina, G. R. Biofuels in the Long-Run Global Energy Supply Mix for Transportation. *Philos. Trans. R. Soc. Math. Phys. Eng. Sci.* **2014**, *372*, 20120323.
(344) Chen, C.-T.; Liao, J. C. Frontiers in Microbial 1-Butanol and Isobutanol Production. *FEMS Microbiol. Lett.* **2016**, *363*, fnw020.
(345) Atsumi, S.; Hanai, T.; Liao, J. C. Non-Fermentative Pathways for Synthesis of Branched-Chain Higher Alcohols as Biofuels. *Nature* **2008**, *451*, 86−89.
(346) Atsumi, S.; Wu, T.-Y.; Eckl, E.-M.; Hawkins, S. D.; Buelter, T.; Liao, J. C. Engineering the Isobutanol Biosynthetic Pathway in *Escherichia coli* by Comparison of Three Aldehyde Reductase/Alcohol Dehydrogenase Genes. *Appl. Microbiol. Biotechnol.* **2010**, *85*, 651−657.
(347) Bastian, S.; Liu, X.; Meyerowitz, J. T.; Snow, C. D.; Chen, M. M. Y.; Arnold, F. H. Engineered Ketol-Acid Reductoisomerase and Alcohol Dehydrogenase Enable Anaerobic 2-Methylpropan-1-Ol Production at Theoretical Yield in *Escherichia coli*. *Metab. Eng.* **2011**, *13*, 345−352.
(348) Matsuda, F.; Ishii, J.; Kondo, T.; Ida, K.; Tezuka, H.; Kondo, A. Increased Isobutanol Production in *Saccharomyces cerevisiae* by Eliminating Competing Pathways and Resolving Cofactor Imbalance. *Microb. Cell Factories* **2013**, *12*, 119.
(349) Smith, K. M.; Liao, J. C. An Evolutionary Strategy for Isobutanol Production Strain Development in *Escherichia coli*. *Metab. Eng.* **2011**, *13*, 674−681.
(350) Baez, A.; Cho, K.-M.; Liao, J. C. High-Flux Isobutanol Production Using Engineered *Escherichia coli*: A Bioreactor Study with *in Situ* Product Removal. *Appl. Microbiol. Biotechnol.* **2011**, *90*, 1681−1690.
(351) Tracy, N. I.; Chen, D.; Crunkleton, D. W.; Price, G. L. Hydrogenated Monoterpenes as Diesel Fuel Additives. *Fuel* **2009**, *88*, 2238−2240.
(352) Pitera, D. J.; Paddon, C. J.; Newman, J. D.; Keasling, J. D. Balancing a Heterologous Mevalonate Pathway for Improved Isoprenoid Production in *Escherichia coli*. *Metab. Eng.* **2007**, *9*, 193−207.
(353) Pfleger, B. F.; Pitera, D. J.; Smolke, C. D.; Keasling, J. D. Combinatorial Engineering of Intergenic Regions in Operons Tunes Expression of Multiple Genes. *Nat. Biotechnol.* **2006**, *24*, 1027−1032.
(354) Wu, J.; Cheng, S.; Cao, J.; Qiao, J.; Zhao, G.-R. Systematic Optimization of Limonene Production in Engineered *Escherichia coli*. *J. Agric. Food Chem.* **2019**, *67*, 7087−7097.
(355) Yee, D. A.; DeNicola, A. B.; Billingsley, J. M.; Creso, J. G.; Subrahmanyam, V.; Tang, Y. Engineered Mitochondrial Production of Monoterpenes in *Saccharomyces cerevisiae*. *Metab. Eng.* **2019**, *55*, 76−84.
(356) Peng, B.; Nielsen, L. K.; Kampranis, S. C.; Vickers, C. E. Engineered Protein Degradation of Farnesyl Pyrophosphate Synthase Is an Effective Regulatory Mechanism to Increase Monoterpene Production in *Saccharomyces cerevisiae*. *Metab. Eng.* **2018**, *47*, 83−93.
(357) Ignea, C.; Raadam, M. H.; Motawia, M. S.; Makris, A. M.; Vickers, C. E.; Kampranis, S. C. Orthogonal Monoterpenoid Biosynthesis in Yeast Constructed on an Isomeric Substrate. *Nat. Commun.* **2019**, *10*, 3799.
(358) Cheng, S.; Liu, X.; Jiang, G.; Wu, J.; Zhang, J.; Lei, D.; Yuan, Y.-J.; Qiao, J.; Zhao, G.-R. Orthogonal Engineering of Biosynthetic

https://doi.org/10.1021/acs.chemrev.2c00403
Chem. Rev. 2023, 123, 5521−5570

Pathway for Efficient Production of Limonene in *Saccharomyces cerevisiae*. *ACS Synth. Biol.* **2019**, *8*, 968−975.
(359) Rolf, J.; Julsing, M. K.; Rosenthal, K.; Lütz, S. A Gram-Scale Limonene Production Process with Engineered *Escherichia coli*. *Molecules* **2020**, *25*, 1881.
(360) Yu, T.; Zhou, Y. J.; Huang, M.; Liu, Q.; Pereira, R.; David, F.; Nielsen, J. Reprogramming Yeast Metabolism from Alcoholic Fermentation to Lipogenesis. *Cell* **2018**, *174*, 1549−1558.
(361) Zhou, Y. J.; Buijs, N. A.; Zhu, Z.; Qin, J.; Siewers, V.; Nielsen, J. Production of Fatty Acid-Derived Oleochemicals and Biofuels by Synthetic Yeast Cell Factories. *Nat. Commun.* **2016**, *7*, 11709.
(362) Inokuma, K.; Liao, J. C.; Okamoto, M.; Hanai, T. Improvement of Isopropanol Production by Metabolically Engineered *Escherichia coli* Using Gas Stripping. *J. Biosci. Bioeng.* **2010**, *110*, 696−701.
(363) Hanai, T.; Atsumi, S.; Liao, J. C. Engineered Synthetic Pathway for Isopropanol Production in *Escherichia coli*. *Appl. Environ. Microbiol.* **2007**, *73*, 7814−7818.
(364) Jang, Y.-S.; Lee, J. Y.; Lee, J.; Park, J. H.; Im, J. A.; Eom, M.-H.; Lee, J.; Lee, S.-H.; Song, H.; Cho, J.-H.; et al. Enhanced Butanol Production Obtained by Reinforcing the Direct Butanol-Forming Route in *Clostridium acetobutylicum*. *mBio* **2012**, *3*, No. e00314-12.
(365) Atsumi, S.; Wu, T.-Y.; Machado, I. M. P.; Huang, W.-C.; Chen, P.-Y.; Pellegrini, M.; Liao, J. C. Evolution, Genomic Analysis, and Reconstruction of Isobutanol Tolerance in *Escherichia coli*. *Mol. Syst. Biol.* **2010**, *6*, 449.
(366) Chen, G. S.; Siao, S. W.; Shen, C. R. Saturated Mutagenesis of Ketoisovalerate Decarboxylase V461 Enabled Specific Synthesis of 1-Pentanol via the Ketoacid Elongation Cycle. *Sci. Rep.* **2017**, *7*, 11284.
(367) Runguphan, W.; Sae-Tang, K.; Tanapongpipat, S. Recent Advances in the Microbial Production of Isopentanol (3-Methyl-1-Butanol). *World J. Microbiol. Biotechnol.* **2021**, *37*, 107.
(368) Vogt, M.; Brüsseler, C.; van Ooyen, J.; Bott, M.; Marienhagen, J. Production of 2-Methyl-1-Butanol and 3-Methyl-1-Butanol in Engineered *Corynebacterium glutamicum*. *Metab. Eng.* **2016**, *38*, 436−445.
(369) Wang, F.; Lv, X.; Xie, W.; Zhou, P.; Zhu, Y.; Yao, Z.; Yang, C.; Yang, X.; Ye, L.; Yu, H. Combining Gal4p-Mediated Expression Enhancement and Directed Evolution of Isoprene Synthase to Improve Isoprene Production in *Saccharomyces cerevisiae*. *Metab. Eng.* **2017**, *39*, 257−266.
(370) Cordova, L. T.; Butler, J.; Alper, H. S. Direct Production of Fatty Alcohols from Glucose Using Engineered Strains of *Yarrowia lipolytica*. *Metab. Eng. Commun.* **2020**, *10*, No. e00105.
(371) Kim, J.; Hwang, S.; Lee, S.-M. Metabolic Engineering for the Utilization of Carbohydrate Portions of Lignocellulosic Biomass. *Metab. Eng.* **2022**, *71*, 2−12.
(372) Du, C.; Li, Y.; Xiang, R.; He, Y.; Sun, H.; Wang, H.; Yuan, W. Co-Utilization of Multiple Lignocellulose-Derived Sugars and Inhibitors by the Robust Mutant *Kluyveromyces marxianus* through Adaptive Laboratory Evolution. *Fuel* **2022**, *315*, 122816.
(373) Jönsson, L. J.; Alriksson, B.; Nilvebrant, N.-O. Bioconversion of Lignocellulose: Inhibitors and Detoxification. *Biotechnol. Biofuels* **2013**, *6*, 16.
(374) Chen, Z.; Liu, D. Toward Glycerol Biorefinery: Metabolic Engineering for the Production of Biofuels and Chemicals from Glycerol. *Biotechnol. Biofuels* **2016**, *9*, 205.
(375) Trinh, C. T.; Srienc, F. Metabolic Engineering of *Escherichia coli* for Efficient Conversion of Glycerol to Ethanol. *Appl. Environ. Microbiol.* **2009**, *75*, 6696−6705.
(376) Agu, C. V.; Ujor, V.; Ezeji, T. C. Metabolic Engineering of *Clostridium beijerinckii* to Improve Glycerol Metabolism and Furfural Tolerance. *Biotechnol. Biofuels* **2019**, *12*, 50.
(377) An, N.; Chen, X.; Sheng, H.; Wang, J.; Sun, X.; Yan, Y.; Shen, X.; Yuan, Q. Rewiring the Microbial Metabolic Network for Efficient Utilization of Mixed Carbon Sources. *J. Ind. Microbiol. Biotechnol.* **2021**, *48*, kuab040.
(378) Hua, Y.; Wang, J.; Zhu, Y.; Zhang, B.; Kong, X.; Li, W.; Wang, D.; Hong, J. Release of Glucose Repression on Xylose Utilization in *Kluyveromyces marxianus* to Enhance Glucose-Xylose Co-Utilization and Xylitol Production from Corncob Hydrolysate. *Microb. Cell Factories* **2019**, *18*, 24.
(379) Wu, Y.; Shen, X.; Yuan, Q.; Yan, Y. Metabolic Engineering Strategies for Co-Utilization of Carbon Sources in Microbes. *Bioengineering* **2016**, *3*, 10.
(380) Sievert, C.; Nieves, L. M.; Panyon, L. A.; Loeffler, T.; Morris, C.; Cartwright, R. A.; Wang, X. Experimental Evolution Reveals an Effective Avenue to Release Catabolite Repression via Mutations in XylR. *Proc. Natl. Acad. Sci. U. S. A.* **2017**, *114*, 7349−7354.
(381) Li, F.-F.; Zhao, Y.; Li, B.-Z.; Qiao, J.-J.; Zhao, G.-R. Engineering *Escherichia coli* for Production of 4-Hydroxymandelic Acid Using Glucose−Xylose Mixture. *Microb. Cell Factories* **2016**, *15*, 90.
(382) Saini, M.; Lin, L.-J.; Chiang, C.-J.; Chao, Y.-P. Synthetic Consortium of *Escherichia coli* for N-Butanol Production by Fermentation of the Glucose−Xylose Mixture. *J. Agric. Food Chem.* **2017**, *65*, 10040−10047.
(383) Farwick, A.; Bruder, S.; Schadeweg, V.; Oreb, M.; Boles, E. Engineering of Yeast Hexose Transporters to Transport D-Xylose without Inhibition by D-Glucose. *Proc. Natl. Acad. Sci. U. S. A.* **2014**, *111*, 5159−5164.
(384) Shen, M.-H.; Song, H.; Li, B.-Z.; Yuan, Y.-J. Deletion of D-Ribulose-5-Phosphate 3-Epimerase (RPE1) Induces Simultaneous Utilization of Xylose and Glucose in Xylose-Utilizing *Saccharomyces cerevisiae*. *Biotechnol. Lett.* **2015**, *37*, 1031−1036.
(385) Jiang, W.; Hernández Villamor, D.; Peng, H.; Chen, J.; Liu, L.; Haritos, V.; Ledesma-Amaro, R. Metabolic Engineering Strategies to Enable Microbial Utilization of C1 Feedstocks. *Nat. Chem. Biol.* **2021**, *17*, 845−855.
(386) Takors, R.; Kopf, M.; Mampel, J.; Bluemke, W.; Blombach, B.; Eikmanns, B.; Bengelsdorf, F. R.; Weuster-Botz, D.; Dürre, P. Using Gas Mixtures of CO, CO_2 and H_2 as Microbial Substrates: The Do's and Don'ts of Successful Technology Transfer from Laboratory to Production Scale. *Microb. Biotechnol.* **2018**, *11*, 606−625.
(387) Pan, H.; Wang, J.; Wu, H.; Li, Z.; Lian, J. Synthetic Biology Toolkit for Engineering *Cupriviadus necator H16* as a Platform for CO2 Valorization. *Biotechnol. Biofuels* **2021**, *14*, 212.
(388) Grousseau, E.; Lu, J.; Gorret, N.; Guillouet, S. E.; Sinskey, A. J. Isopropanol Production with Engineered *Cupriavidus necator* as Bioproduction Platform. *Appl. Microbiol. Biotechnol.* **2014**, *98*, 4277−4290.
(389) Miller, J.; MacEachran, D.; Burd, H.; Sathitsuksanoh, N.; Bi, C.; Yeh, Y.-C.; Lee, T. S.; Hillson, N. J.; Chhabra, S. R.; Singer, S. W.; et al. Engineering of *Ralstonia eutropha H16* for Autotrophic and Heterotrophic Production of Methyl Ketones. *Appl. Environ. Microbiol.* **2013**, *79*, 4433−4439.
(390) Antonovsky, N.; Gleizer, S.; Noor, E.; Zohar, Y.; Herz, E.; Barenholz, U.; Zelcbuch, L.; Amram, S.; Wides, A.; Tepper, N.; et al. Sugar Synthesis from CO2 in *Escherichia coli*. *Cell* **2016**, *166*, 115−125.
(391) Gleizer, S.; Ben-Nissan, R.; Bar-On, Y. M.; Antonovsky, N.; Noor, E.; Zohar, Y.; Jona, G.; Krieger, E.; Shamshoum, M.; Bar-Even, A.; et al. Conversion of *Escherichia coli* to Generate All Biomass Carbon from CO_2. *Cell* **2019**, *179*, 1255−1263.
(392) Gassler, T.; Baumschabl, M.; Sallaberger, J.; Egermeier, M.; Mattanovich, D. Adaptive Laboratory Evolution and Reverse Engineering Enhances Autotrophic Growth in *Pichia pastoris*. *Metab. Eng.* **2022**, *69*, 112−121.
(393) Flamholz, A. I.; Dugan, E.; Blikstad, C.; Gleizer, S.; Ben-Nissan, R.; Amram, S.; Antonovsky, N.; Ravishankar, S.; Noor, E.; Bar-Even, A.; et al. Functional Reconstitution of a Bacterial CO_2 Concentrating Mechanism in *Escherichia coli*. *eLife* **2020**, *9*, No. e59882.
(394) Tashiro, Y.; Hirano, S.; Matson, M. M.; Atsumi, S.; Kondo, A. Electrical-Biological Hybrid System for CO_2 Reduction. *Metab. Eng.* **2018**, *47*, 211−218.

https://doi.org/10.1021/acs.chemrev.2c00403
Chem. Rev. 2023, 123, 5521−5570

(395) Wilkes, R. a.; Aristilde, L. Degradation and Metabolism of Synthetic Plastics and Associated Products by *Pseudomonas* Sp.: Capabilities and Challenges. *J. Appl. Microbiol.* **2017**, *123*, 582−593.
(396) Jia, X.; Qin, C.; Friedberger, T.; Guan, Z.; Huang, Z. Efficient and Selective Degradation of Polyethylenes into Liquid Fuels and Waxes under Mild Conditions. *Sci. Adv.* **2016**, *2*, No. e1501591.
(397) Tao, W.; Lv, L.; Chen, G.-Q. Engineering Halomonas Species TD01 for Enhanced Polyhydroxyalkanoates Synthesis via CRISPRi. *Microb. Cell Factories* **2017**, *16*, 48.
(398) Knott, B. C.; Erickson, E.; Allen, M. D.; Gado, J. E.; Graham, R.; Kearns, F. L.; Pardo, I.; Topuzlu, E.; Anderson, J. J.; Austin, H. P.; et al. Characterization and Engineering of a Two-Enzyme System for Plastics Depolymerization. *Proc. Natl. Acad. Sci. U. S. A.* **2020**, *117*, 25476−25485.
(399) Tiso, T.; Narancic, T.; Wei, R.; Pollet, E.; Beagan, N.; Schröder, K.; Honak, A.; Jiang, M.; Kenny, S. T.; Wierckx, N.; et al. Towards Bio-Upcycling of Polyethylene Terephthalate. *Metab. Eng.* **2021**, *66*, 167−178.
(400) Zhang, J. Z.; Creelman, R. A.; Zhu, J.-K. From Laboratory to Field. Using Information from Arabidopsis to Engineer Salt, Cold, and Drought Tolerance in Crops. *Plant Physiol.* **2004**, *135*, 615−621.
(401) Keasling, J.; Garcia Martin, H.; Lee, T. S.; Mukhopadhyay, A.; Singer, S. W.; Sundstrom, E. Microbial Production of Advanced Biofuels. *Nat. Rev. Microbiol.* **2021**, *19*, 701−715.
(402) Gong, Z.; Nielsen, J.; Zhou, Y. J. Engineering Robustness of Microbial Cell Factories. *Biotechnol. J.* **2017**, *12*, 1700014.
(403) Russ, D.; Kishony, R. Additivity of Inhibitory Effects in Multidrug Combinations. *Nat. Microbiol.* **2018**, *3*, 1339−1345.
(404) Brandt, B. A.; Jansen, T.; Görgens, J. F.; van Zyl, W. H. Overcoming Lignocellulose-Derived Microbial Inhibitors: Advancing the *Saccharomyces cerevisiae* Resistance Toolbox. *Biofuels Bioprod. Biorefining* **2019**, *13*, 1520−1536.
(405) Albers, E.; Larsson, C. A Comparison of Stress Tolerance in YPD and Industrial Lignocellulose-Based Medium among Industrial and Laboratory Yeast Strains. *J. Ind. Microbiol. Biotechnol.* **2009**, *36*, 1085−1091.
(406) Lam, F. H.; Ghaderi, A.; Fink, G. R.; Stephanopoulos, G. Engineering Alcohol Tolerance in Yeast. *Science* **2014**, *346*, 71−75.
(407) Heckmann, D.; Campeau, A.; Lloyd, C. J.; Phaneuf, P. V.; Hefner, Y.; Carrillo-Terrazas, M.; Feist, A. M.; Gonzalez, D. J.; Palsson, B. O. Kinetic Profiling of Metabolic Specialists Demonstrates Stability and Consistency of *in Vivo* Enzyme Turnover Numbers. *Proc. Natl. Acad. Sci. U. S. A.* **2020**, *117*, 23182−23190.
(408) David, F.; Nielsen, J.; Siewers, V. Flux Control at the Malonyl-CoA Node through Hierarchical Dynamic Pathway Regulation in *Saccharomyces cerevisiae*. *ACS Synth. Biol.* **2016**, *5*, 224−233.
(409) Alper, H.; Stephanopoulos, G. Global Transcription Machinery Engineering: A New Approach for Improving Cellular Phenotype. *Metab. Eng.* **2007**, *9*, 258−267.
(410) Zingaro, K. A.; Terry Papoutsakis, E. GroESL Overexpression Imparts *Escherichia coli* Tolerance to I-, n-, and 2-Butanol, 1,2,4-Butanetriol and Ethanol with Complex and Unpredictable Patterns. *Metab. Eng.* **2013**, *15*, 196−205.
(411) Liu, Z. L.; Slininger, P. J.; Gorsich, S. W. Enhanced Biotransformation of Furfural and Hydroxymethylfurfural by Newly Developed Ethanologenic Yeast Strains. *Appl. Biochem. Biotechnol.* **2005**, *121*, 0451−0460.
(412) Lewis Liu, Z.; Moon, J.; Andersh, B. J.; Slininger, P. J.; Weber, S. Multiple Gene-Mediated NAD(P)H-Dependent Aldehyde Reduction Is a Mechanism of *in Situ* Detoxification of Furfural and 5-Hydroxymethylfurfural by *Saccharomyces cerevisiae*. *Appl. Microbiol. Biotechnol.* **2008**, *81*, 743−753.
(413) Liu, Z. L.; Ma, M. Pathway-Based Signature Transcriptional Profiles as Tolerance Phenotypes for the Adapted Industrial Yeast *Saccharomyces cerevisiae* Resistant to Furfural and HMF. *Appl. Microbiol. Biotechnol.* **2020**, *104*, 3473−3492.
(414) Gan, Y.; Qi, X.; Lin, Y.; Guo, Y.; Zhang, Y.; Wang, Q. A Hierarchical Transcriptional Regulatory Network Required for Long-Term Thermal Stress Tolerance in an Industrial *Saccharomyces cerevisiae* Strain. *Front. Bioeng. Biotechnol.* **2022**, *9*, 826238.
(415) Jia, H.; Sun, X.; Sun, H.; Li, C.; Wang, Y.; Feng, X.; Li, C. Intelligent Microbial Heat-Regulating Engine (IMHeRE) for Improved Thermo-Robustness and Efficiency of Bioconversion. *ACS Synth. Biol.* **2016**, *5*, 312−320.
(416) Wen, Z.; Ledesma-Amaro, R.; Lu, M.; Jiang, Y.; Gao, S.; Jin, M.; Yang, S. Combined Evolutionary Engineering and Genetic Manipulation Improve Low PH Tolerance and Butanol Production in a Synthetic Microbial *Clostridium* Community. *Biotechnol. Bioeng.* **2020**, *117*, 2008−2022.
(417) Zhu, H.; Ren, X.; Wang, J.; Song, Z.; Shi, M.; Qiao, J.; Tian, X.; Liu, J.; Chen, L.; Zhang, W. Integrated OMICS Guided Engineering of Biofuel Butanol-Tolerance in Photosynthetic *Synechocystis sp. PCC 6803*. *Biotechnol. Biofuels* **2013**, *6*, 106.
(418) Fisher, M. A.; Boyarskiy, S.; Yamada, M. R.; Kong, N.; Bauer, S.; Tullman-Ercek, D. Enhancing Tolerance to Short-Chain Alcohols by Engineering the *Escherichia coli* AcrB Efflux Pump to Secrete the Non-Native Substrate n-Butanol. *ACS Synth. Biol.* **2014**, *3*, 30−40.
(419) Cao, M.; Fatma, Z.; Song, X.; Hsieh, P.-H.; Tran, V. G.; Lyon, W. L.; Sayadi, M.; Shao, Z.; Yoshikuni, Y.; Zhao, H. A Genetic Toolbox for Metabolic Engineering of *Issatchenkia orientalis*. *Metab. Eng.* **2020**, *59*, 87−97.
(420) Tran, V. G.; Cao, M.; Fatma, Z.; Song, X.; Zhao, H. Development of a CRISPR/Cas9-Based Tool for Gene Deletion in *Issatchenkia orientalis*. *mSphere* **2019**, *4*, No. e00345-19.
(421) Angermayr, S. A.; Gorchs Rovira, A.; Hellingwerf, K. J. Metabolic Engineering of Cyanobacteria for the Synthesis of Commodity Products. *Trends Biotechnol.* **2015**, *33*, 352−361.
(422) Upton, D. J.; McQueen-Mason, S. J.; Wood, A. J. *In Silico* Evolution of *Aspergillus niger* Organic Acid Production Suggests Strategies for Switching Acid Output. *Biotechnol. Biofuels* **2020**, *13*, 27.
(423) Cen, Y.-K.; Lin, J.-G.; Wang, Y.-L.; Wang, J.-Y.; Liu, Z.-Q.; Zheng, Y.-G. The Gibberellin Producer Fusarium Fujikuroi: Methods and Technologies in the Current Toolkit. *Front. Bioeng. Biotechnol.* **2020**, *8*, 232.
(424) Chao, R.; Mishra, S.; Si, T.; Zhao, H. Engineering Biological Systems Using Automated Biofoundries. *Metab. Eng.* **2017**, *42*, 98−108.
(425) Otero-Muras, I.; Carbonell, P. Automated Engineering of Synthetic Metabolic Pathways for Efficient Biomanufacturing. *Metab. Eng.* **2021**, *63*, 61−80.
(426) Hillson, N.; Caddick, M.; Cai, Y.; Carrasco, J. A.; Chang, M. W.; Curach, N. C.; Bell, D. J.; Le Feuvre, R.; Friedman, D. C.; Fu, X.; et al. Building a Global Alliance of Biofoundries. *Nat. Commun.* **2019**, *10*, 2040.
(427) *Edinburgh Genome Foundry*: University of Edinburgh, 2022; https://www.ed.ac.uk/biology/research/facilities/edinburgh-genome-foundry (accessed 2022-04-08).
(428) Chao, R.; Liang, J.; Tasan, I.; Si, T.; Ju, L.; Zhao, H. Fully Automated One-Step Synthesis of Single-Transcript TALEN Pairs Using a Biological Foundry. *ACS Synth. Biol.* **2017**, *6*, 678−685.
(429) Si, T.; Chao, R.; Min, Y.; Wu, Y.; Ren, W.; Zhao, H. Automated Multiplex Genome-Scale Engineering in Yeast. *Nat. Commun.* **2017**, *8*, 15187.
(430) Enghiad, B.; Xue, P.; Singh, N.; Boob, A. G.; Shi, C.; Petrov, V. A.; Liu, R.; Peri, S. S.; Lane, S. T.; Gaither, E. D.; et al. PlasmidMaker Is a Versatile, Automated, and High Throughput End-to-End Platform for Plasmid Construction. *Nat. Commun.* **2022**, *13*, 2697.
(431) *London Biofoundry: Engineering Biology Technologies for the Advancement of Synthetic Biology*; SynbiCITE, 2022; http://synbicite.com/innovation/londondnafoundry/ (accessed 2022-04-08).
(432) *Singapore BioFoundry*; SynCTI, 2015; https://syncti.org/biofoundry/ (accessed 2022-04-08).
(433) *Agile BioFoundry*; U.S. Department of Energy, Office of Energy Efficiency and Renewable Energy, 2022; https://agilebiofoundry.org/ (accessed 2022-08-04).

https://doi.org/10.1021/acs.chemrev.2c00403
Chem. Rev. 2023, 123, 5521−5570

(434) Zhang, J.; Chen, Y.; Fu, L.; Guo, E.; Wang, B.; Dai, L.; Si, T. Accelerating Strain Engineering in Biofuel Research via Build and Test Automation of Synthetic Biology. *Curr. Opin. Biotechnol.* **2021**, *67*, 88−98.
(435) Wagner, J. M.; Liu, L.; Yuan, S.-F.; Venkataraman, M. V.; Abate, A. R.; Alper, H. S. A Comparative Analysis of Single Cell and Droplet-Based FACS for Improving Production Phenotypes: Riboflavin Overproduction in *Yarrowia lipolytica*. *Metab. Eng.* **2018**, *47*, 346−356.
(436) Cacho, R. A.; Tang, Y.; Chooi, Y.-H. Next-Generation Sequencing Approach for Connecting Secondary Metabolites to Biosynthetic Gene Clusters in Fungi. *Front. Microbiol.* **2015**, *5*, 774.
(437) Yang, Y.; Wang, L.; Wang, B.; Liu, S.; Yu, B.; Wang, T. Application of Next-Generation Sequencing Following Tandem Mass Spectrometry to Expand Newborn Screening for Inborn Errors of Metabolism: A Multicenter Study. *Front. Genet.* **2019**, *10*, 86.
(438) Linshiz, G.; Jensen, E.; Stawski, N.; Bi, C.; Elsbree, N.; Jiao, H.; Kim, J.; Mathies, R.; Keasling, J. D.; Hillson, N. J. End-to-End Automated Microfluidic Platform for Synthetic Biology: From Design to Functional Analysis. *J. Biol. Eng.* **2016**, *10*, 3.
(439) Bowman, E. K.; Alper, H. S. Microdroplet-Assisted Screening of Biomolecule Production for Metabolic Engineering Applications. *Trends Biotechnol.* **2020**, *38*, 701−714.
(440) Wang, B. L.; Ghaderi, A.; Zhou, H.; Agresti, J.; Weitz, D. A.; Fink, G. R.; Stephanopoulos, G. Microfluidic High-Throughput Culturing of Single Cells for Selection Based on Extracellular Metabolite Production or Consumption. *Nat. Biotechnol.* **2014**, *32*, 473−478.
(441) Saito, K.; Ota, Y.; Tourlousse, D. M.; Matsukura, S.; Fujitani, H.; Morita, M.; Tsuneda, S.; Noda, N. Microdroplet-Based System for Culturing of Environmental Microorganisms Using FNAP-Sort. *Sci. Rep.* **2021**, *11*, 9506.
(442) Diefenbach, X. W.; Farasat, I.; Guetschow, E. D.; Welch, C. J.; Kennedy, R. T.; Sun, S.; Moore, J. C. Enabling Biocatalysis by High-Throughput Protein Engineering Using Droplet Microfluidics Coupled to Mass Spectrometry. *ACS Omega* **2018**, *3*, 1498−1508.
(443) Lawson, C. E.; Martí, J. M.; Radivojevic, T.; Jonnalagadda, S. V. R.; Gentz, R.; Hillson, N. J.; Peisert, S.; Kim, J.; Simmons, B. A.; Petzold, C. J.; et al. Machine Learning for Metabolic Engineering: A Review. *Metab. Eng.* **2021**, *63*, 34−60.
(444) Whitehead, E.; Rudolf, F.; Kaltenbach, H.-M.; Stelling, J. Automated Planning Enables Complex Protocols on Liquid-Handling Robots. *ACS Synth. Biol.* **2018**, *7*, 922−932.
(445) Kim, E.; Huang, K.; Saunders, A.; McCallum, A.; Ceder, G.; Olivetti, E. Materials Synthesis Insights from Scientific Literature via Text Extraction and Machine Learning. *Chem. Mater.* **2017**, *29*, 9436−9444.
(446) Zampieri, G.; Vijayakumar, S.; Yaneske, E.; Angione, C. Machine and Deep Learning Meet Genome-Scale Metabolic Modeling. *PLOS Comput. Biol.* **2019**, *15*, No. e1007084.
(447) Antonakoudis, A.; Barbosa, R.; Kotidis, P.; Kontoravdi, C. The Era of Big Data: Genome-Scale Modelling Meets Machine Learning. *Comput. Struct. Biotechnol. J.* **2020**, *18*, 3287−3300.
(448) Faure, L.; Mollet, B.; Liebermeister, W.; Faulon, J.-L. Artificial Metabolic Networks: Enabling Neural Computation with Metabolic Networks. *bioRxiv* 2022; DOI: 10.1101/2022.01.09.475487.
(449) Zhao, C.; Sinumvayo, J. P.; Zhang, Y.; Li, Y. Design and Development of a "Y-Shaped" Microbial Consortium Capable of Simultaneously Utilizing Biomass Sugars for Efficient Production of Butanol. *Metab. Eng.* **2019**, *55*, 111−119.
(450) Sasaki, Y.; Yoshikuni, Y. Metabolic Engineering for Valorization of Macroalgae Biomass. *Metab. Eng.* **2022**, *71*, 42−61.
(451) Tsoi, R.; Wu, F.; Zhang, C.; Bewick, S.; Karig, D.; You, L. Metabolic Division of Labor in Microbial Systems. *Proc. Natl. Acad. Sci. U. S. A.* **2018**, *115*, 2526−2531.
(452) Jawed, K.; Yazdani, S. S.; Koffas, M. AG. Advances in the Development and Application of Microbial Consortia for Metabolic Engineering. *Metab. Eng. Commun.* **2019**, *9*, No. e00095.
(453) Zomorrodi, A. R.; Maranas, C. D. OptCom: A Multi-Level Optimization Framework for the Metabolic Modeling and Analysis of Microbial Communities. *PLOS Comput. Biol.* **2012**, *8*, No. e1002363.
(454) Zomorrodi, A. R.; Islam, M. M.; Maranas, C. D. D-OptCom: Dynamic Multi-Level and Multi-Objective Metabolic Modeling of Microbial Communities. *ACS Synth. Biol.* **2014**, *3*, 247−257.
(455) Chan, S. H. J.; Simons, M. N.; Maranas, C. D. SteadyCom: Predicting Microbial Abundances While Ensuring Community Stability. *PLOS Comput. Biol.* **2017**, *13*, No. e1005539.
(456) Magnúsdóttir, S.; Thiele, I. Modeling Metabolism of the Human Gut Microbiome. *Curr. Opin. Biotechnol.* **2018**, *51*, 90−96.
(457) Levy, R.; Borenstein, E. Metabolic Modeling of Species Interaction in the Human Microbiome Elucidates Community-Level Assembly Rules. *Proc. Natl. Acad. Sci. U. S. A.* **2013**, *110*, 12804−12809.
(458) Wehrs, M.; Tanjore, D.; Eng, T.; Lievense, J.; Pray, T. R.; Mukhopadhyay, A. Engineering Robust Production Microbes for Large-Scale Cultivation. *Trends Microbiol.* **2019**, *27*, 524−537.
(459) Crater, J. S.; Lievense, J. C. Scale-up of Industrial Microbial Processes. *FEMS Microbiol. Lett.* **2018**, *365*, fny138.

https://doi.org/10.1021/acs.chemrev.2c00403
Chem. Rev. 2023, 123, 5521−5570