ACS chemical biology

Reviews

pubs.acs.org/acschemicalbiology

Integrating the Protein and Metabolic Engineering Toolkits for Next-Generation Chemical Biosynthesis

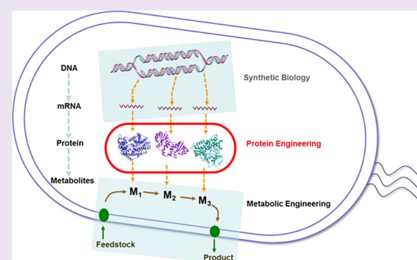
Christopher M. Pirie,[†] Marjan De Mey,[†,‡] Kristala L. Jones Prather,[§] and Parayil Kumaran Ajikumar*,[†]

[†]Manus Biosynthesis Inc., Suite 102, 790 Memorial Drive, Cambridge, Massachusetts 02139, United States

[‡]Centre of Expertise−Industrial Biotechnology and Biocatalysis, Ghent University, Coupure links 653, 9000 Ghent, Belgium

[§]Department of Chemical Engineering, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, United States

ABSTRACT: Through microbial engineering, biosynthesis has the potential to produce thousands of chemicals used in everyday life. Metabolic engineering and synthetic biology are fields driven by the manipulation of genes, genetic regulatory systems, and enzymatic pathways for developing highly productive microbial strains. Fundamentally, it is the biochemical characteristics of the enzymes themselves that dictate flux through a biosynthetic pathway toward the product of interest. As metabolic engineers target sophisticated secondary metabolites, there has been little recognition of the reduced catalytic activity and increased substrate/product promiscuity of the corresponding enzymes compared to those of central metabolism. Thus, fine-tuning these enzymatic characteristics through protein engineering is paramount for developing high-productivity microbial strains for secondary metabolites. Here, we describe the importance of protein engineering for advancing metabolic engineering of secondary metabolism pathways. This pathway integrated enzyme optimization can enhance the collective toolkit of microbial engineering to shape the future of chemical manufacturing.



Over the past decade, the chemical manufacturing industry witnessed the beginnings of a new technological transformation in industrial chemical production. This revolution was spurred by advances in metabolic engineering and synthetic biology that enabled sophisticated engineering of microbes for sustainable chemical manufacturing. (Bio)chemicals derived from the primary metabolism of biological systems, including ethanol, butanol, propanediol, lactic acid, acetic acid, citric acid, and succinic acid, were the first set of chemicals to be targeted for microbial overproduction and are being commercially produced from renewable carbon sources.[1−8] Secondary metabolites, which are products of complex chemistries found in nature, represent a vast number of chemical candidates (>200,000) for a myriad of applications such as drugs, food additives, consumer products, industrial chemicals, and biofuels.[9−14] Developing technologies for sustainable production of these chemicals at commercial scales has the opportunity to drastically alter the chemical industry's landscape. Traditionally sourced as natural extracts from plants or as derivatives of petrochemical processes, secondary metabolites produced in engineered microbes can alleviate market volume and price constraints with clean, renewable, and sustainable technology. However, only a few of these chemicals have been produced using engineered pathways in microbes and are still in transition from lab to commercial manufacturing.[15,16]

Microbes engineered to enable these processes are manipulated such that the fluxes of enzymatic reactions in the pathway are balanced to limit the accumulation of intermediates and byproducts to increase the conversion of feedstock through intermediate metabolites to a desired product.[17] In addition, several competing and off target pathways have been altered to channel cellular flux through desired biosynthetic pathways.[18] Metabolic engineers have traditionally used rational and combinatorial genetic engineering to improve the balance of reactions in the biochemical pathways responsible for carbon flux, energy/cofactor supply, and accumulation of intermediates.[19−23] With the emergence of synthetic biology, researchers are now trying to design molecular level control mechanisms intended to improve flexibility and capability in the engineering of biological processes.[24,25] At a fundamental level, all of these engineering interventions are changing the concentration of enzymes in a cell but largely ignore the properties of the enzymes themselves. Little attention has been paid to the kinetics of enzyme-catalyzed biochemical reactions and how they might be improved. At the same time, advances in enzyme engineering facilitating design of improved natural and synthetic enzymes that catalyze both existing and novel reactions have yet to be fully explored for microbial biosynthesis of naturally occurring[26] or tailor-made chemicals.[27,28] To achieve sustainable manufacturing of secondary metabolites, one must combine metabolic engineering (i.e., introducing biochemical pathways from various organisms into a microbial host and balancing pathways in conjunction with the native metabolism for overproduction) with protein engineering of secondary metabolism enzymes that are specifically hindered by reduced catalytic activity and increased substrate/product promiscuity[29]

Received: November 20, 2012
Accepted: February 4, 2013
Published: February 4, 2013

© 2013 American Chemical Society

dx.doi.org/10.1021/cb300634b | ACS Chem. Biol. 2013, 8, 662−672

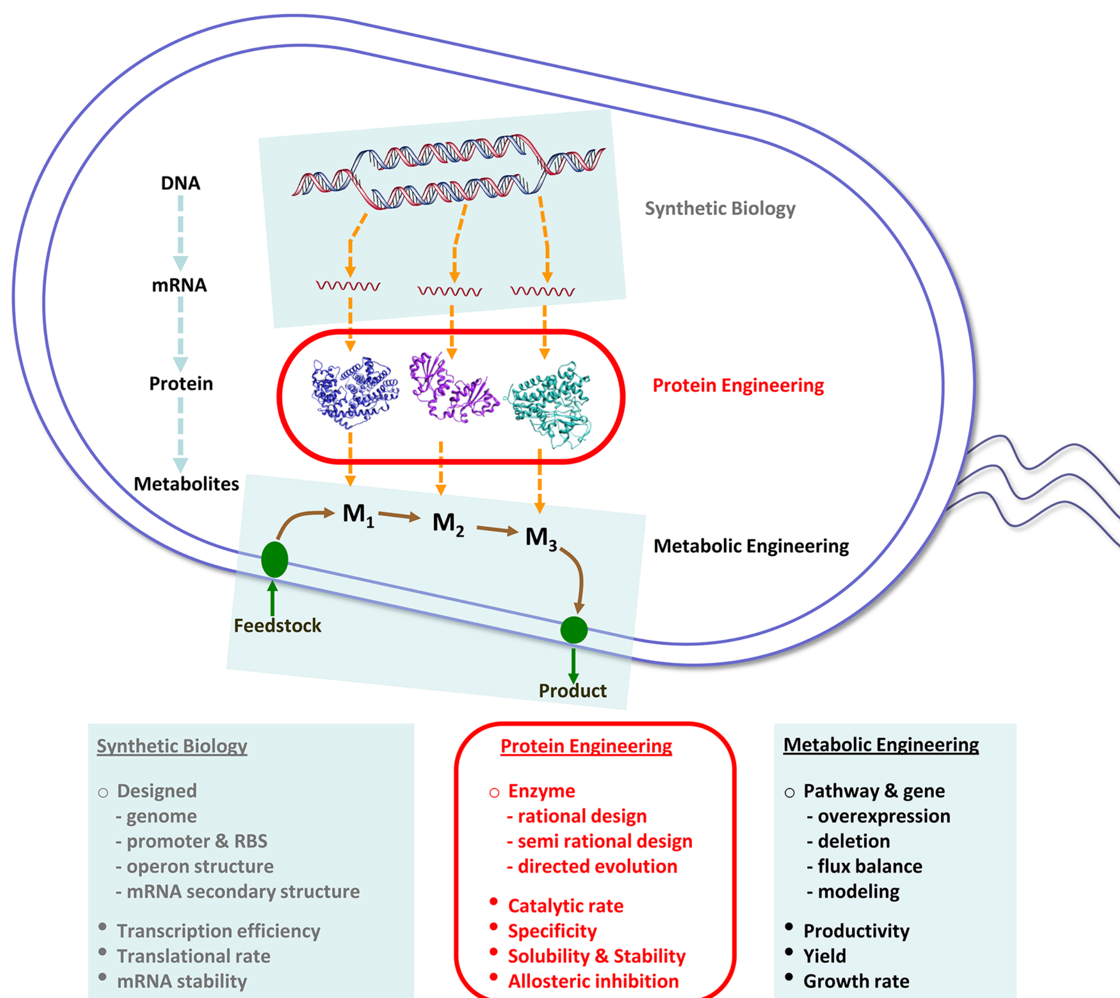


Figure 1. Cellular engineering for secondary metabolite biosynthesis. The synthetic biology toolkit and components are designed to manipulate the rate and efficiency of gene transcription, including very complex designed genetic circuits. These controls can also extend to translational rate through alteration of ribosome binding site (RBS) strength or mRNA secondary structure. Protein engineering of biosynthetic enzymes (blue, purple, and teal) aims to optimize catalytic rate, solubility, stability, specificity, product inhibition, and degradation rate. Metabolic engineering considers pathway and gene expression tools in an attempt to control feedstock, intermediate metabolites/substrates (M_1, M_2, and M_3), and product concentrations by balancing steady-state levels of the engineered enzymes to optimize productivity/flux. Open circle symbols represent approaches that can be used to modify specific characteristics shown in filled circles.

(Figure 1). This review highlights the exciting possibilities as pathway integrated enzyme optimization approaches become fully assimilated into the metabolic engineering/synthetic biology toolkit.

■ ENZYME CHALLENGES FOR METABOLIC ENGINEERING

Recently, Bar-Even et al. reported a detailed meta-analysis of the natural enzymes that catalyze the reactions of primary, intermediate, and secondary metabolism (Box 1.).[30] Their work clearly frames the uniquely slow kinetic characteristics of enzymes in pathways associated with various levels of metabolism (Figure 2a). When the microbial engineering goal is high flux through these enzymes, the traditional metabolic engineering approach would be to express 30-fold more of the slow enzymes. However, this approach is suboptimal, as doing so may come at the cost of a significant metabolic burden that can be detrimental to productivity (Figure 2b).[31] If you consider pathways of inefficient enzymes, which are often more than five steps long, it quickly becomes impossible to simply overexpress the enzymes, as metabolic resources and physical space in the cell become hard constraints. In this context, protein engineering techniques are capable of overcoming the deficiencies of secondary metabolism enzymes to enable the development of microbial strains with commercial level productivity. Therefore, there is an urgent need to systematize and exploit protein engineering techniques to improve activity but also to control specificity, optimize stability, direct localization, improve solubility, and/or modulate regulatory domains in metabolic engineering for next generation chemical synthesis.[32−34]

Enzymes drive flux through metabolic pathways. In the end, the productivity of a pathway is dictated by the kinetic characteristics of the enzymes involved. The enzymatic properties influencing flux noted above (solubility, localization, cofactor dependence, product inhibition, substrate/product specificity, catalytic rate, and stability) are a few of the biochemical protein characteristics that can be manipulated for any particular enzyme and may thus be addressed using protein engineering techniques.[35] Some of these properties are also influenced by metabolic balance and the levels of other metabolites, cofactors, or inhibitors. The degree to which each of these factors hinders

Box 1.

Meta-analysis of Natural Enzyme Kinetics. Biologists have always had an appreciation for the importance of enzyme kinetics in cellular metabolism of different organisms. But until recently, none had endeavored to pursue a comprehensive meta-analysis of the numerous natural enzymes whose kinetics had been characterized in independent studies. Bar-Even et al.[30] studied the k_{cat} and k_{cat}/K_M of several thousand natural enzymes and discovered that the median rates were $\sim 10\ s^{-1}$ and $\sim 10^5\ s^{-1}\ M^{-1}$, respectively. Such kinetics are far from diffusion-limited ($>10^8\ s^{-1}\ M^{-1}$) and indicative of evolutionary selective pressures insufficient to elicit maximal catalytic efficiencies. The foremost implication for design of biosynthetic microbial strains is that even pathways producing primary metabolites have room for protein engineering efforts to facilitate process improvements. When the relative kinetics of different classes of enzymes were compared, it was found that the median rates for secondary metabolism enzymes are ~30-fold less than those for primary metabolism of carbohydrate catabolism and energy production (Figure 2a). Differences between primary and secondary metabolism enzymes stem from the conflicting selective pressures and diversity requirements of secondary enzymes and metabolites. Because biological activity is a rare trait in chemical compounds, nature's search for protein−ligand interactions demands diversity of secondary metabolites generated by enzymatic promiscuity (in substrate and product). This increased promiscuity is correlated with reduced activity.[106,107] *Thus, metabolic engineers seeking to construct strains producing secondary metabolites face a requirement rather than an opportunity to apply protein engineering techniques to overcome the catalytic inefficiency of the relevant secondary metabolism enzymes* as described by Bar-Even et al. as well as the substrate/product promiscuity selected for in nature.

direct metabolic engineering of a heterologous pathway depends upon the characteristics of the native enzyme.

CONSTRAINTS ON SECONDARY METABOLITE BIOSYNTHESIS

Research efforts in metabolic engineering often emulate biosynthetic processes utilized in nature. Evolution has fashioned multiple routes for producing individual biochemicals from various starting points (i.e., glucose, glycerol, ethanol, xylose, carbon dioxide). Some metabolic pathways are highly targeted (phospholipid or tryptophan biosynthesis), while others are phenomenally diverse (terpenoid or flavonoid biosynthesis).[11] In these diverse pathways, biosynthesis has evolved to utilize sequential coupling of smaller building blocks followed by cyclization reactions to create hydrocarbon frameworks. These chemical skeletons can then be hydroxylated, oxidized, or reduced to increase diversity or enable further functionalization such as acetylation, glycosylation, and amidation[12,36,37] (Figure 3a). Framework and functionalized diversity can arise from the presence of either multiple enzymes acting on the same substrate or one promiscuous active site that can act on multiple substrates. The diverse pathways can also display network-like characteristics, with identical enzymes acting in alternative sequences to produce similar products (Figure 3b).

Over time, selective evolutionary pressure has induced more stringent evolution of primary metabolism enzymes resulting in improved specificity and activity. As a result, engineering of microbial strains overproducing primary metabolites leads to fewer activity/specificity challenges and can be effectively modulated by changing enzyme concentration.[38,39] In contrast, secondary metabolite biosynthesis pathways show more enzymatic promiscuity, which is biochemically linked to reduced catalytic activity.[29,30] Because biological activity is a generally uncommon characteristic in the chemical landscape, nature has evolved enzymes for secondary metabolite biosynthesis with greater substrate and product diversity, in contrast to their primary metabolism counterparts.[29] This makes secondary metabolite pathway engineering a difficult task. Regardless of whether the limiting factor is detrimental promiscuity or insufficient activity, there is room in these pathways for substantial enzyme optimization.[31] We will discuss this with two examples: terpenoid and flavonoid metabolic pathways.

The terpenoid pathway is one of nature's most diverse secondary metabolism pathways with greater than 55,000 members (Figure 3a). Many terpenoids have been identified as having a wide array of useful applications.[40] Structurally and functionally similar terpene synthase and cyclase enzymes have evolved through gene duplication to convert the pathway's allylic pyrophosphate (PP) backbone molecules, geranyl-PP, farnesyl-PP, and geranylgeranyl-PP, into dephosphorylated and often cyclized or branched compounds referred to as monoterpenes, sesquiterpenes, and diterpenes, respectively.[41] Nature's defect/default chemistry in terpenoid cyclization (allylic substrates with highly unstable carbo-cation transition states) plays a huge role in creating the chemical diversity. At the enzyme level, many terpenoid synthases have evolved catalytic sites that coordinate loosely bound transition states and exhibit drastic product promiscuity due to a lack of carbo-cation intermediate stabilization. This serves to further increase molecular diversity without necessitating the metabolic burden of numerous unique enzymes.[42] Further, many cytochrome P450s exist to oxidize and otherwise functionalize the various carbon skeletons,[43] which can subsequently be modified by enzymes transferring methyl, acetyl, or glucose units. Enzymes in the terpenoid pathway deviate from the canonical single substrate/single product dogma and instead have either a range of possible substrates or can synthesize a mixture of products from a given substrate. Engineering such a pathway is not amenable to traditional metabolic engineering approaches, because byproducts and pathway intermediates are produced by the same enzyme. This means altering enzyme concentration will have little effect on the product yield, and only protein engineering approaches will be effective in improving the pathway.

Another unique example of the aforementioned primary/secondary metabolism principle is the phenylalanine/flavonoid biosynthesis pathway. Beginning with primary metabolism intermediates, phosphoenolpyruvate from glycolysis and D-erythrose-4-phosphate from pentose phosphate, independent pathways with chorismate as the common precursor lead to the formation of three different cyclic amino acids: tryptophan, tyrosine, and phenylalanine. Two of these products, phenylalanine and tyrosine, are precursors for flavonoid biosynthesis through coumarate (Figure 3b). Flavonoids are an important class of molecules for plant defense mechanisms and may have significant health benefits in humans.[44−46] Luteolin and quercetin are two downstream secondary metabolites that are reported to have pharmaceutical effects[47,48] and along with their flavonoid relatives constitute a very diverse class of over 9,000 compounds including anthocyanins and flavonols, whose

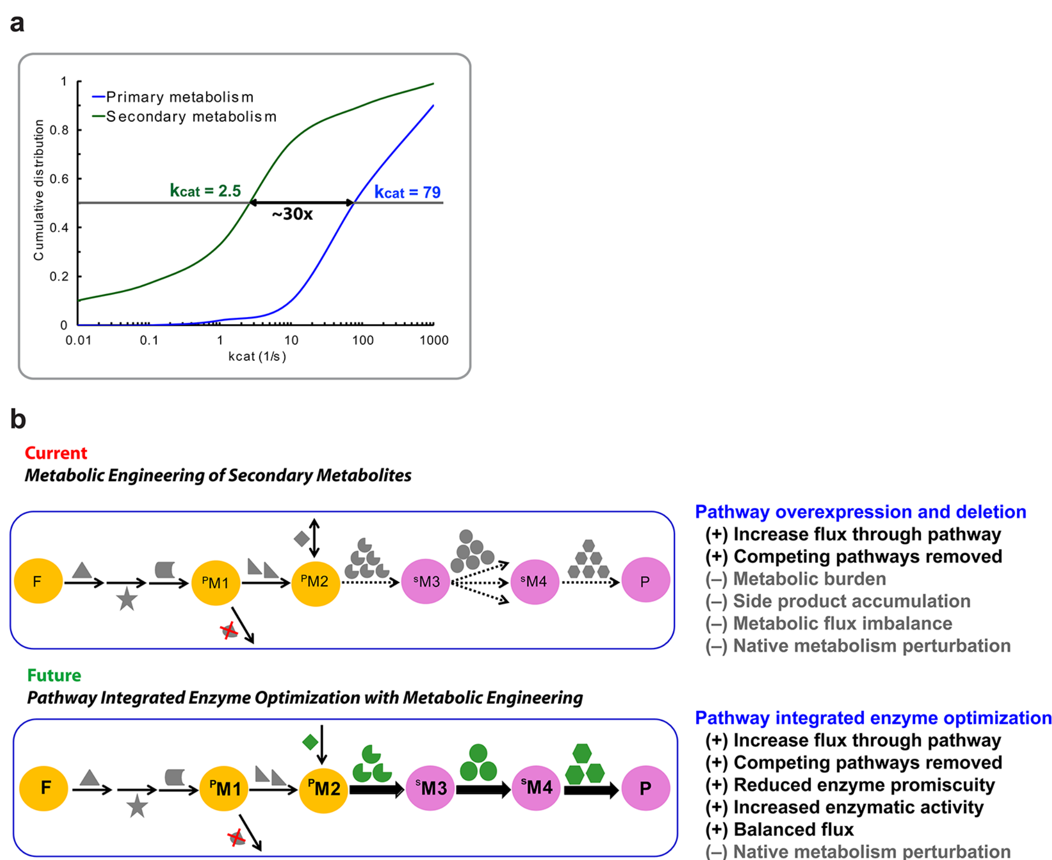


Figure 2. Significance of enzyme engineering in secondary metabolite pathway engineering. (Aa) Evolutionary selective pressures have driven the biosynthetic enzyme activity of primary metabolism higher than that of secondary metabolism. The median k_{cat} for primary metabolism enzymes is approximately 30-fold greater than that of secondary metabolism enzymes (Box 1.). Redrawn from Bar-Even et al.[30] (b) Pathway integrated enzyme optimization is critical to metabolic engineering of secondary metabolite production. Gray symbols represent wild-type enzymes, while larger green symbols are engineered enzymes. Dashed arrows are marginally active and promiscuous enzyme pathways, while thick arrows are highly active and specific engineered pathways. Crossed out enzymes are knocked-out or attenuated. F = feedstock, ^{P}M = primary metabolite, ^{S}M = secondary metabolite, P = product. Traditional (Current) metabolic engineering begins with overexpression of native and/or heterologous pathways and deletion of competing pathways, which tends to generate some modest amount of product. Combined protein engineering and metabolic engineering approaches (Future) for pathway tuning can optimize flux by reducing byproduct formation and increasing enzyme activity.

microbial biosynthesis has been addressed previously.[49−51] This chemical diversity is the product of both a large set of biosynthetic enzymes as well as a moderate level of substrate promiscuity.[52] Substrate promiscuity makes flavonoid secondary metabolism look more like an interconnected network than a linear pathway. This distributed movement of molecules through the network complicates metabolic engineering concepts of pathway optimization.

It is clear that what nature has evolved as a pro-survival phenotype is enzymatic promiscuity leading to chemical diversity, but as scientists and engineers strive to harness nature's chemical biosynthesis capabilities, the advantageous traits of evolution become hurdles for efficiency and economics. Individual enzyme promiscuity at intermediate steps or perhaps most importantly at the penultimate step leads not only to hampered productivity but also to the formation of contaminants in a desired extraction product at levels relative to the lack of specificity and the physio-chemical characteristics of the intermediates and final product. In large-scale commercial operations, purification of target chemicals from contaminating byproducts (a requirement for many applications other than biofuels), especially byproducts that are highly similar to the main product, can reduce yields and increase production costs tremendously. Metabolic engineers have only begun to harness protein engineering tools in the context of strain development,[31,53−58] but to develop commercially viable strains for many secondary metabolite chemicals, it will be necessary to apply the full suite of techniques to engineer native enzymes for enhanced specificity and activity.

THE PROTEIN ENGINEERING TOOLKIT

New techniques in protein engineering have taken the field from site-directed mutagenesis and alanine scanning to computational design and combinatorial libraries. Many tools are now available to assist in rational, semi-rational, and random mutagenesis approaches to enzyme engineering.[59,60] Using these techniques, enzymes have been engineered with enhanced activity, increased/decreased/altered specificity, and increased solubility/stability (Table 1). Selection of an appropriate engineering technique is dependent on available knowledge of enzyme structure, activity/inhibitory domains, and functional or evolutionary relationships balanced against the availability of screening/selection techniques; together these two variables inform the decision of which strategy should be used (Figure 4).

Rational design of enzymes depends on the ability to derive structure−function relationships. The availability of protein structures and optimized homology modeling techniques will

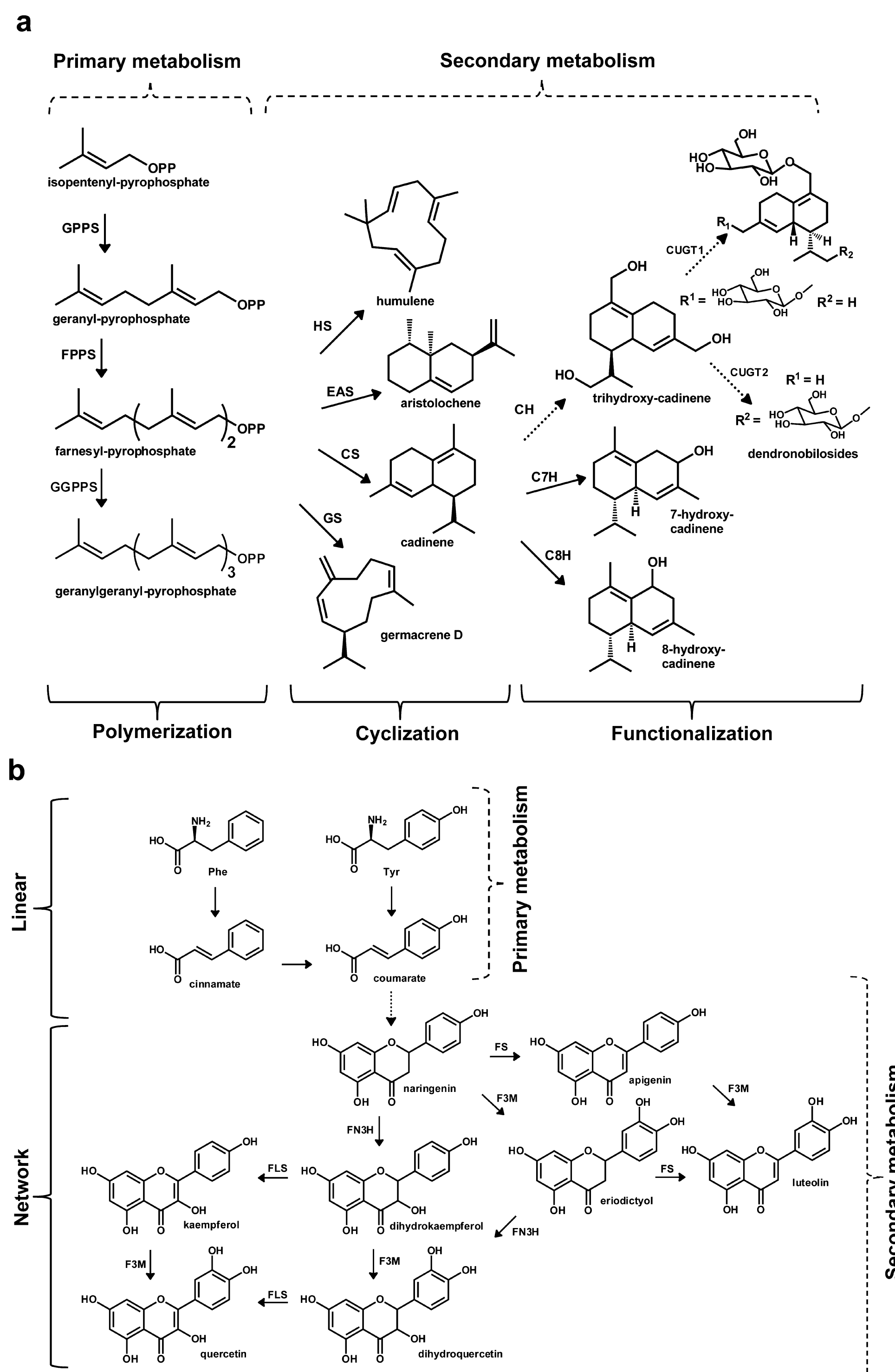


Figure 3. Diversity in secondary metabolism. Chemical diversity in secondary metabolism often comes from a series of chemical processes: polymerization, cyclization, and functionalization with layered product promiscuity. Each of these processes can occur by multiple, distinct, and parallel enzymatic conversions. (a) In the terpenoid pathway, geranyl-pyrophosphate synthase (GPPS), farnesyl-pyrophosphate synthase (FPPS), and geranylgeranyl-pyrophosphate synthase (GGPPS) catalyze the sequential polymerization of isoprenyl-pyrophosphate, a C_5 compound, to C_{10}, C_{15}, C_{20}, and more. Another set of interrelated enzymes catalyze the formation of basic hydrocarbon sesquiterpenes: humulene, aristolochene, cadinene, and germacrene D, for example. Functionalization is exemplified through cadinene hydroxylation and glycosylation. Enzymes: humulene synthase (HS), 5-epi-aristolochene synthase (EAS), cadinene synthase (CS), germacrene D synthase (GS), cadinene-8-hydroxylase (C8H), cadinene-7-hydroxylase (C7H), cadinene hydroxylase (CH), cadinene UDP-glucosyl-transferase 1 and 2 (CUGT1 and CUGT2). Here, dashed lines are used to indicate putative enzymes. (b) The phenylpropanoid pathway exemplifies the interface between primary and secondary metabolism as well as their different traits. Primary metabolism drives carbon building blocks through a series of linear enzymatic steps to form tyrosine (Tyr) and phenylalanine (Phe). The products of primary metabolism are used as precursors to secondary metabolism: Phe and Tyr are converted to flavonoids through the common coumarate and naringenin intermediates. Beyond naringenin, the network architecture of flavonoid secondary metabolism becomes evident as multiple pathways are capable of synthesizing complex chemicals such as quercetin and luteolin. Enzymes: flavone synthase (FS), flavonoid 3′-monooxygenase (F3M), flavonone 3β-hydroxylase (FN3H), flavonol synthase (FLS).[104,105]

Table 1. Examples of Pathway Enzymes Engineered for Activity, Specificity, or Solubility

enzyme (pathway)	engineering technique (tool)	outcome
poplar *O*-methyltransferase-7 (flavonoid)[a]	rational design (Sybyl/FlexX)	6× increase in activity
geranylgeranyl pyrophosphate synthase (terpenoid)[b]	directed evolution (epPCR)	2× increase in activity
premnaspirodiene synthase (terpenoid)[c]	rational design (GOLD)	fully shifted product specificity
horseradish peroxidase (lignin)[d]	directed evolution (epPCR)	enantiomeric selectivity
P450-BM3 (terpenoid)[e]	semi-rational (enriched library)	increased substrate specificity
CYP11A1 (terpenoid)[f]	rational design (visual molecular dynamics)	4× increase in solubility
horseradish peroxidase (lignin)[g]	directed evolution (epPCR/recombination)	7× increase in functional expression, 2× increase in stability at 70 °C
5-epi-aristolochene synthase (terpenoid)[h]	semi-rational (computationally assisted design strategy and proteolytic selection)	mutants stable at >65 °C
geranylgeranyl pyrophosphate synthase and levopimaradiene synthase (terpenoid)[i]	semi-rational (site saturation and epPCR)	18× increase in product titer
3-hydroxy-3-methylglutaryl-CoA reductase and humulene synthase (terpenoid)[j]	semi-rational (*in silico* phylogenetic analysis and residue-specific alanine scanning	43× increase in product titer

[a]Sybyl: molecular visualization software from Tripos Inc. FelxX: software to flexibly dock ligands in a binding site.[84] [b]epPCR: error-prone polymerase chain reaction.[81] [c]GOLD: program for calculating docking modes of small molecules in a binding site.[68] [d]Reference 78. [e]Reference 90. [f]Reference 108. [g]Reference 109. [h]Reference 110. [i]Reference 31. [j]Reference 53.

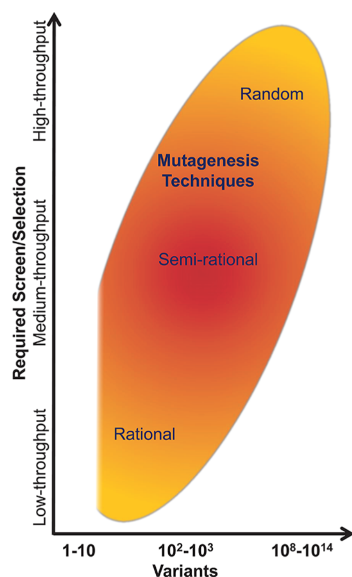


Figure 4. Selecting an enzyme engineering technique. Rational, semi-rational, and random mutagenesis techniques span a spectrum of library sizes whose screening/sorting requires throughput dependent on that size. Protein engineering is limited generally by deficient structure−function understanding of proteins (unshaded fraction of oval − difficult to predict successful single mutations) and experimentally where most low/medium-throughput techniques are insufficient to screen libraries with 10^8-10^{14} variants, typical of random mutagenesis approaches. In contrast, high-throughput techniques are unnecessary when screening small libraries. When selecting which approach to use for engineering enzymes, especially those of secondary metabolism, high-throughput screens are often not available and structure−function knowledge is even more limited, thus the most common approaches are semi-rational. The density gradient of the shaded area suggests the relative proportion of enzymes capable of being engineered by each technique.

confer an ability to build such relationships from which more effective design experiments can be conducted. One advantage of rational design is that it reduces the number of experimental variants necessary to screen to identify useful mutations.[35] However, this approach is limited by the necessity of structural information that is not always available, especially for proteins with no known homologues or that are difficult to crystallize or not amenable to other structural analysis techniques, such as membrane-associated proteins.

Structural information alone does not enable rational design; it depends on computational tools as well. These computational tools must (a) interpret the structural information, (b) use it to make calculations or abstractions of interatomic and/or intermolecular interactions, and (c) translate them into actionable predictions for mutagenesis. Tools range from those related to construction of secondary and tertiary structures from primary sequences using homology modeling such as SWISS-MODEL or MODELER[61−64] to those used for the characterization of structural dynamics and mutational effects such as NAMD, ROSETTA, or YASARA.[65−67] These tools can help take basic amino acid sequence information and translate it into structural models that can then be used as the basis for dynamics or energetics calculations on enzyme, enzyme−substrate, or enzyme−transition state models in the formation of product. Such docking studies can in turn be used to identify useful positions for site-specific mutagenesis.[68] However, even the most advanced energetics tools have difficulty in quantitatively predicting results of mutagenesis.[69] More often, structural and dynamics information can be used to qualitatively predict productive mutations (i.e., a polar residue stabilizing a charged intermediate or a smaller residue opening up space for a larger substrate in an active site). More recently, computational tools have been adapted to enable crowd-sourced conformational and mutational testing, resulting in significantly improved biochemical characteristics of synthetic enzymes including improvements in enzymatic activity through otherwise unpredictable findings.[70]

Semi-rational design is an intermediate approach that involves identification of important residues in an enzyme, whether in the active site, a cofactor site, or elsewhere, and diversifying at a limited set of positions.[31,53,60] Diversity can be focused by restricting mutations to phylogenetic consensus/preference or to conservative residue mutations.[71,72] From this information combinatorial and evolutionary approaches can be applied to increase diversity for medium-throughput screening.[73,74] Here, the number of variants used for screening can be on the order of tens or hundreds, which is more realistic for traditional analytic techniques. In this case structural information is useful but not required, and a high-throughput

screening assay often is not necessary, nor is an ability to predict precisely which mutations will be advantageous.

In the absence of structural or ancestral information, random mutagenesis methods can be useful alternatives. Fully random approaches or deeply combinatorial focused libraries require the generation of millions or billions of variants through mutagenesis, degenerate primer design, or fragment shuffling.[75,76] A high-throughput selection technique must be available to identify desirable mutants from a large library: screening approaches vary from cell sorting methods to panning to automated analytical tools.[77] Traditionally, these have involved colorimetric or fluorescent products, dye conjugates, or fluorescence resonance energy transfer substrates.[78−80] A somewhat underappreciated component of evolutionary approaches for either activity or selectivity is an assay-dependent ability to select for mutations that also increase an enzyme's overall stability and solubility. Random mutagenesis and variant sorting is a valuable protein engineering tool,[78,81] but for most enzyme engineering applications in the context of a metabolic system identifying a high-throughput selection assay is a hurdle since most products of interest are not readily quantified by such assays.

As one approaches protein engineering of secondary metabolism enzymes, it is important to identify what information is available or can be derived computationally. This knowledge, along with a survey of the available analytical techniques, should help guide the adoption of a particular approach. Nevertheless, in principle, any one of these engineering approaches can be applied to enzymes to address the characteristics that need to be optimized prior to designing high-productivity chemical biosynthesis microbial strains.

APPLICATIONS IN ENZYME ENGINEERING

Construction of a novel pathway in an organism begins with identification and selection of the necessary enzymes. Depending on the enzyme, multiple isotypes may exist from evolutionarily divergent organisms (bacteria, plant, or animal). When an endogenous version of the enzyme exists, it is often the best choice because it has evolved within the context of the host environment (i.e., an *Escherichia coli* enzyme will work well in the *E. coli* cytoplasm). On the other hand, endogenous enzymes are usually tightly regulated; thus, heterologous enzymes may avoid such control mechanisms and promote biosynthesis. Efficacy of heterologous enzymes in microbial systems is highly unpredictable; therefore, it is advisable to examine multiple orthologs, to identify a best possible starting point for optimization through protein engineering within the framework of the biosynthetic pathway.

Rapid and detailed identification of active sites is now possible thanks to high-resolution structure determination methods and bioinformatic homology studies,[82] where previously extensive biochemical studies were necessary.[83] Understanding active-site structure allows modulation of key residues capable of altering an enzyme's catalytic efficiency.[31,84] Once the shape and composition of an active site is elucidated, the choice of mutations can be informed by phylogenetic comparison, accessibility, or hydrophobic/electrostatic transition state stabilization. In the absence of structural information, random mutagenesis approaches can also be used to increase enzyme kinetics, so long as an appropriate screening assay can be developed.[85−87] If a target product requires synthetic chemistry not found in nature, it is possible in some cases to design *de novo* enzyme functionalities by restructuring the active site of a high-stability scaffold to coordinate the new chemistry.[88] Such designed enzymes display relatively low initial activities that can be enhanced further by directed evolution. However, endless optimization of enzymatic activity, natural or designed, for chemical biosynthesis is productive only if the reaction being catalyzed is specific.

When optimizing productivity, product specificity is critical. It is important that each enzyme in the constructed metabolic pathway generates only the intermediates that will serve as substrates for the subsequent enzymes and eventually the appropriate final product. Fortunately, both rational design and directed evolution have been shown to be capable of redirecting product specificity.[26,68,78,79,89−92] In rational studies where the active site architecture is known, mutation positions altering specificity can be found in both the first or second contact shell and selection of new residues can be informed by phylogeny where available. If phylogenic information is unavailable, engineering is limited to what can be derived directly from structural analysis and predictions of the particular structure−function relationship. The key to using directed evolution with large libraries to increase catalytic specificity is, as always, a powerful selection assay.

Stability and solubility are important determinants of steady-state enzyme levels for a metabolic pathway.[33,53] Often correlated,[93,94] these two protein characteristics can contribute to the degradation rate of an enzyme. An enzyme with limited solubility or short-lived stability can result in a low enzyme concentration or require greater turnover that is energetically wasteful. Generally, proteins maintain secondary and tertiary structures that bury hydrophobic residues in a core shielded from their aqueous environment by more hydrophilic amino acids. When this equilibrium is disrupted, a protein can unfold or even aggregate. Therefore, amino acid charge at the surface and in the core of an engineered enzyme is important to consider. It can be especially important in cases where the native pathway or enzyme has evolved for activity at temperatures suboptimal for fermentation (e.g., 20 vs 37 °C).

For reasons related to structural flexibility and metabolic control, enzymes tend to evolve thermostability at the edge of their native environment's temperature.[94] Tools to predict the behavior of native or engineered enzymes upon overexpression *in vivo* are available and can help inform initial selection from orthologs (Table 2).[95] Rational design methods to increase enzyme stability and solubility interrogate the fundamental thermodynamics that drive protein folding and use the information to test or predict various point mutations *in silico*.[96] Mutations can range from simple incorporation of disulfide bonds[97] to helix end-capping[98] to surface entropic stabilization or rigidification.[99] While it is clear that the depth of

Table 2. SOLpro[a] Predicted Solubility of Farnesyl Pyrophosphate Synthase Homologs

species	% aa sequence identity	prediction	probability
Escherichia coli	100	insoluble	0.599
Bacillus subtilis	45.1	soluble	0.735
Ginkgo biloba	23.5	insoluble	0.813
Saccharomyces cerevisiae	22.1	insoluble	0.821
Arabidopsis thaliana	19.9	insoluble	0.704
Synechocystis sp. PCC6803	19.0	insoluble	0.839

[a]Reference 78.

understanding in the field between structure or sequence and stability or solubility is sufficient to enable industrious design, directed evolution studies aimed at the same goals suggest that they can be accomplished through a variety of sometimes unpredictable mutations. Random mutagenesis combined with well-designed reporter systems[100] and semi-rational approaches have allowed smaller libraries to be screened.[101,102]

CONCLUSIONS

As metabolic engineering ambitiously broadens the scope of renewable chemicals production, the use of secondary metabolites, which are as diverse in their chemical composition as they are in their potential applications, is inevitable. Design of commercially viable microbial production strains for secondary metabolites is dependent upon high-efficiency conversion of feedstock to product through several intermediate metabolic steps, which is related to the efficacy of the enzymes involved. In primary metabolism, biosynthetic enzymes evolve under highly selective conditions, which results in improved biochemical and biophysical characteristics. In contrast, enzymes involved in secondary metabolism actually evolve to increase product diversity: they exhibit substrate and product promiscuity through flexibility that correlates with decreased stability and activity. As a result, metabolic engineers must begin to incorporate protein engineering into strain design when utilizing such enzymes, but there are relatively few examples where this has been done.[31,53] As synthetic biology further enables advanced control of metabolic systems, the entire gamut of tools for enzyme manipulation should be employed to optimize in the context of the system.

Rational design, semi-rational approaches, and directed evolution have each been previously applied to engineer enzymes with altered specificity, stability, and activity (Table 1). To date, the majority of enzyme engineering has focused on catalysts for biotransformation processes with operating conditions often quite different than microbial fermentation. To identify enzyme variants ideal for *in vivo* biosynthesis of secondary metabolites, we advise a semi-rational approach. Any available structure−function knowledge should be leveraged to inform specific alterations. In the absence of such information, diversity libraries covering focused segments of the enzyme with a limited set of possible substitutions can be screened with mid-to-high throughput assays for desirable characteristics.[103] Where high-throughput screens are available for the reaction product or a downstream metabolite, random mutagenesis can be important to identify unanticipated scaffold mutations. In the future, protein engineering efforts must be integrated with the metabolic engineering and pathway balancing required to develop a commercially viable microbe and will be fundamental to obtaining high-productivity strains.

Unlike primary metabolites, many secondary metabolites are used in smaller specialty chemical applications such as consumer products, food additives, and pharmaceuticals. The markets for individual metabolites of this class are frequently not large enough to compel investment of the tens of millions of dollars required for developing microbial strains and fermentation-based production for these chemicals. As demand for such microbes becomes commonplace, applying protein engineering with metabolic engineering will become a necessity for faster and cheaper microbial strain engineering. Soon, industry demand may extend beyond the repertoire of natural chemicals toward new chemical entities accessible only through the catalytic powers of engineered enzymes. Secondary metabolism itself exemplifies the capability of existing enzyme scaffolds to evolve new functionality to create new chemical diversity. This chemical diversity can be extended further by modifying the metabolite scaffolds using traditional chemistry techniques. Clearly, pathway integrated enzyme optimization must be drawn into the metabolic engineering toolkit to facilitate secondary metabolite biosynthesis and will play a growing role in the future as novel biochemistries are explored. Finally, taking bold steps to integrate and utilize tools from traditional biology/biochemistry research will be necessary for rapid increases in the ease and speed of strain engineering for industrial renewable chemicals production.

AUTHOR INFORMATION

Corresponding Author

*E-mail: pkaji@manusbio.com.

Notes

The authors declare the following competing financial interest(s): C.M.P., M.D.M., K.L.J.P, and P.K.A. have financial interests in Manus Biosynthesis, Inc.

ACKNOWLEDGMENTS

The authors acknowledge constructive discussion and suggestions on the manuscript from K. Tyo (Chemical and Biological Engineering, Northwestern University) and M. Ackerman (Thayer School of Engineering, Dartmouth College). M.D.M. acknowledges the support of the Multidisciplinary Research Partnership Ghent Bio-Economy.

REFERENCES

(1) Dien, B. S., Cotta, M. A., and Jeffries, T. W. (2003) Bacteria engineered for fuel ethanol production: current status. *Appl. Microbiol. Biotechnol.* 63, 258−266.

(2) Lee, S. K., Chou, H., Ham, T. S., Lee, T. S., and Keasling, J. D. (2008) Metabolic engineering of microorganisms for biofuels production: from bugs to synthetic biology to fuels. *Curr. Opin. Biotechnol.* 19, 556−563.

(3) Steen, E. J., Chan, R., Prasad, N., Myers, S., Petzold, C. J., Redding, A., Ouellet, M., and Keasling, J. D. (2008) Metabolic engineering of Saccharomyces cerevisiae for the production of n-butanol. *Microb. Cell. Fact.* 7, 36.

(4) Nakamura, C. E., and Whited, G. M. (2003) Metabolic engineering for the microbial production of 1,3-propanediol. *Curr. Opin. Biotechnol.* 14, 454−459.

(5) Patel, M. A., Ou, M. S., Harbrucker, R., Aldrich, H. C., Buszko, M. L., Ingram, L. O., and Shanmugam, K. T. (2006) Isolation and characterization of acid-tolerant, thermophilic bacteria for effective fermentation of biomass-derived sugars to lactic acid. *Appl. Environ. Microbiol.* 72, 3228−3235.

(6) Wendisch, V. F., Bott, M., and Eikmanns, B. J. (2006) Metabolic engineering of Escherichia coli and Corynebacterium glutamicum for biotechnological production of organic acids and amino acids. *Curr. Opin. Microbiol.* 9, 268−274.

(7) Song, H., and Lee, S. Y. (2006) Production of succinic acid by bacterial fermentation. *Enzyme Microb. Technol.* 39, 352−361.

(8) Wargacki, A. J., Leonard, E., Win, M. N., Regitsky, D. D., Santos, C. N. S., Kim, P. B., Cooper, S. R., Raisner, R. M., Herman, A., Sivitz, A. B., Lakshmanaswamy, A., Kashiyama, Y., Baker, D., and Yoshikuni, Y. (2012) An engineered microbial platform for direct biofuel production from brown macroalgae. *Science* 335, 308−313.

(9) Newman, D. J., and Cragg, G. M. (2007) Natural products as sources of new drugs over the last 25 years. *J. Nat. Prod.* 70, 461−477.

(10) Weisshaar, B., and Jenkins, G. I. (1998) Phenylpropanoid biosynthesis and its regulation. *Curr. Opin. Plant Biol.* 1, 251−257.

(11) Fischbach, M. A., and Clardy, J. (2007) One pathway, many products. *Nat. Chem. Biol.* 3, 353−355.

(12) Gershenzon, J., and Dudareva, N. (2007) The function of terpene natural products in the natural world. *Nat. Chem. Biol. 3*, 408−414.

(13) Schäfer, H., and Wink, M. (2009) Medicinally important secondary metabolites in recombinant microorganisms or plants: progress in alkaloid biosynthesis. *Biotechnol. J. 4*, 1684−1703.

(14) Salomon, C. E., Magarvey, N. A., and Sherman, D. H. (2004) Merging the potential of microbial genetics with biological and chemical diversity: an even brighter future for marine natural product drug discovery. *Nat. Prod. Rep. 21*, 105−121.

(15) Lenihan, J. R., Tsuruta, H., Diola, D., Renninger, N. S., and Regentin, R. (2008) Developing an industrial artemisinic acid fermentation process to support the cost-effective production of antimalarial artemisinin-based combination therapies. *Biotechnol. Prog. 24*, 1026−1032.

(16) Zhao, Y., Yang, J., Qin, B., Li, Y., Sun, Y., Su, S., and Xian, M. (2011) Biosynthesis of isoprene in Escherichia coli via methylerythritol phosphate (MEP) pathway. *Appl. Microbiol. Biotechnol. 90*, 1915−1922.

(17) Keasling, J. D. (2010) Manufacturing molecules through metabolic engineering. *Science 330*, 1355−1358.

(18) Celińska, E. (2010) Debottlenecking the 1,3-propanediol pathway by metabolic engineering. *Biotechnol. Adv. 28*, 519−530.

(19) Santos, C. N. S., and Stephanopoulos, G. (2008) Combinatorial engineering of microbes for optimizing cellular phenotype. *Curr. Opin. Chem. Biol. 12*, 168−176.

(20) Bond-Watts, B. B., Bellerose, R. J., and Chang, M. C. Y. (2011) Enzyme mechanism as a kinetic control element for designing synthetic biofuel pathways. *Nat. Chem. Biol. 7*, 222−227.

(21) Tyo, K. E. J., Fischer, C. R., Simeon, F., and Stephanopoulos, G. (2010) Analysis of polyhydroxybutyrate flux limitations by systematic genetic and metabolic perturbations. *Metab. Eng. 12*, 187−195.

(22) 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., Chang, M. C. Y., Withers, S. T., Shiba, Y., Sarpong, R., and Keasling, J. D. (2006) Production of the antimalarial drug precursor artemisinic acid in engineered yeast. *Nature 440*, 940−943.

(23) Yim, H., Haselbeck, R., Niu, W., Pujol-Baxley, C., Burgard, A., Boldt, J., Khandurina, J., Trawick, J. D., Osterhout, R. E., Stephen, R., Estadilla, J., Teisan, S., Schreyer, H. B., Andrae, S., Yang, T. H., Lee, S. Y., Burk, M. J., and Dien, S. V. (2011) Metabolic engineering of Escherichia coli for direct production of 1,4-butanediol. *Nat. Chem. Biol. 7*, 445−452.

(24) Group, Bio FAB, Baker, D., Church, G., Collins, J., Endy, D., Jacobson, J., Keasling, J., Modrich, P., Smolke, C., and Weiss, R. (2006) Engineering life: building a fab for biology. *Sci. Am. 294*, 44−51.

(25) Boyle, P. M., and Silver, P. A. (2012) Parts plus pipes: Synthetic biology approaches to metabolic engineering. *Metab. Eng 14*, 223−232.

(26) Dietrich, J. A., Yoshikuni, Y., Fisher, K. J., Woolard, F. X., Ockey, D., McPhee, D. J., Renninger, N. S., Chang, M. C. Y., Baker, D., and Keasling, J. D. (2009) A novel semi-biosynthetic route for artemisinin production using engineered substrate-promiscuous P450(BM3). *ACS Chem. Biol. 4*, 261−267.

(27) Lee, P. C., Holtzapple, E., and Schmidt-Dannert, C. (2010) Novel activity of Rhodobacter sphaeroides spheroidene monooxygenase CrtA expressed in Escherichia coli. *Appl. Environ. Microbiol. 76*, 7328−7331.

(28) Sawayama, A. M., Chen, M. M. Y., Kulanthaivel, P., Kuo, M.-S., Hemmerle, H., and Arnold, F. H. (2009) A panel of cytochrome P450 BM3 variants to produce drug metabolites and diversify lead compounds. *Chem.—Eur. J. 15*, 11723−11729.

(29) Firn, R. D., and Jones, C. G. (2003) Natural products—a simple model to explain chemical diversity. *Nat. Prod. Rep. 20*, 382−391.

(30) Bar-Even, A., Noor, E., Savir, Y., Liebermeister, W., Davidi, D., Tawfik, D. S., and Milo, R. (2011) The moderately efficient enzyme: evolutionary and physicochemical trends shaping enzyme parameters. *Biochemistry 50*, 4402−4410.

(31) Leonard, E., Ajikumar, P. K., Thayer, K., Xiao, W.-H., Mo, J. D., Tidor, B., Stephanopoulos, G., and Prather, K. L. J. (2010) Combining metabolic and protein engineering of a terpenoid biosynthetic pathway for overproduction and selectivity control. *Proc. Natl. Acad. Sci. U.S.A. 107*, 13654−13659.

(32) Foo, J. L., Ching, C. B., Chang, M. W., and Leong, S. S. J. (2012) The imminent role of protein engineering in synthetic biology. *Biotechnol. Adv. 30*, 541−549.

(33) Pleiss, J. (2011) Protein design in metabolic engineering and synthetic biology. *Curr. Opin. Biotechnol. 22*, 611−617.

(34) Zabala, A. O., Cacho, R. A., and Tang, Y. (2012) Protein engineering towards natural product synthesis and diversification. *J. Ind. Microbiol. Biotechnol. 39*, 227−241.

(35) Lutz, S., and Bornscheuer, U. T. (2009) *Protein Engineering Handbook*, Wiley-VCH, Weinheim.

(36) Kai, K., Mizutani, M., Kawamura, N., Yamamoto, R., Tamai, M., Yamaguchi, H., Sakata, K., and Shimizu, B. (2008) Scopoletin is biosynthesized via ortho-hydroxylation of feruloyl CoA by a 2-oxoglutarate-dependent dioxygenase in Arabidopsis thaliana. *Plant J. 55*, 989−999.

(37) Vogt, T. (2010) Phenylpropanoid Biosynthesis. *Mol. Plant 3*, 2−20.

(38) Ikeda, M., and Katsumata, R. (1999) Hyperproduction of tryptophan by Corynebacterium glutamicum with the modified pentose phosphate pathway. *Appl. Environ. Microbiol. 65*, 2497−2502.

(39) Lütke-Eversloh, T., and Stephanopoulos, G. (2007) L-tyrosine production by deregulated strains of Escherichia coli. *Appl. Microbiol. Biotechnol. 75*, 103−110.

(40) Ajikumar, P. K., Tyo, K., Carlsen, S., Mucha, O., Phon, T. H., and Stephanopoulos, G. (2008) Terpenoids: opportunities for biosynthesis of natural product drugs using engineered microorganisms. *Mol. Pharm. 5*, 167−190.

(41) Chen, F., Tholl, D., Bohlmann, J., and Pichersky, E. (2011) The family of terpene synthases in plants: a mid-size family of genes for specialized metabolism that is highly diversified throughout the kingdom. *Plant J. 66*, 212−229.

(42) Noel, J. P., Dellas, N., Faraldos, J. A., Zhao, M., Hess, B. A., Jr, Smentek, L., Coates, R. M., and O'Maille, P. E. (2010) Structural elucidation of cisoid and transoid cyclization pathways of a sesquiterpene synthase using 2-fluorofarnesyl diphosphates. *ACS Chem. Biol. 5*, 377−392.

(43) Jung, S. T., Lauchli, R., and Arnold, F. H. (2011) Cytochrome P450: taming a wild type enzyme. *Curr. Opin. Biotechnol. 22*, 809−817.

(44) Treutter, D. (2005) Significance of flavonoids in plant resistance and enhancement of their biosynthesis. *Plant Biol. (Stuttgart) 7*, 581−591.

(45) Ververidis, F., Trantas, E., Douglas, C., Vollmer, G., Kretzschmar, G., and Panopoulos, N. (2007) Biotechnology of flavonoids and other phenylpropanoid-derived natural products. Part I: Chemical diversity, impacts on plant biology and human health. *Biotechnol. J. 2*, 1214−1234.

(46) Korkina, L., Kostyuk, V., De Luca, C., and Pastore, S. (2011) Plant phenylpropanoids as emerging anti-inflammatory agents. *Mini-Rev. Med. Chem. 11*, 823−835.

(47) Jang, S., Kelley, K. W., and Johnson, R. W. (2008) Luteolin reduces IL-6 production in microglia by inhibiting JNK phosphorylation and activation of AP-1. *Proc. Natl. Acad. Sci. U.S.A. 105*, 7534−7539.

(48) Murakami, A., Ashida, H., and Terao, J. (2008) Multitargeted cancer prevention by quercetin. *Cancer Lett. 269*, 315−325.

(49) Ferrer, J.-L., Austin, M. B., Stewart, C., and Noel, J. P. (2008) Structure and function of enzymes involved in the biosynthesis of phenylpropanoids. *Plant Physiol. Biochem. 46*, 356−370.

(50) Fowler, Z. L., and Koffas, M. A. G. (2009) Biosynthesis and biotechnological production of flavanones: current state and perspectives. *Appl. Microbiol. Biotechnol. 83*, 799−808.

(51) Leonard, E., Yan, Y., and Koffas, M. A. G. (2006) Functional expression of a P450 flavonoid hydroxylase for the biosynthesis of

plant-specific hydroxylated flavonols in Escherichia coli. *Metab. Eng. 8*, 172−181.

(52) Routaboul, J.-M., Kerhoas, L., Debeaujon, I., Pourcel, L., Caboche, M., Einhorn, J., and Lepiniec, L. (2006) Flavonoid diversity and biosynthesis in seed of Arabidopsis thaliana. *Planta 224*, 96−107.

(53) Yoshikuni, Y., Dietrich, J. A., Nowroozi, F. F., Babbitt, P. C., and Keasling, J. D. (2008) Redesigning enzymes based on adaptive evolution for optimal function in synthetic metabolic pathways. *Chem. Biol. 15*, 607−618.

(54) Ajikumar, P. K., Xiao, W.-H., Tyo, K. E. J., Wang, Y., Simeon, F., Leonard, E., Mucha, O., Phon, T. H., Pfeifer, B., and Stephanopoulos, G. (2010) Isoprenoid pathway optimization for Taxol precursor overproduction in Escherichia coli. *Science 330*, 70−74.

(55) Albertsen, L., Chen, Y., Bach, L. S., Rattleff, S., Maury, J., Brix, S., Nielsen, J., and Mortensen, U. H. (2011) Diversion of flux toward sesquiterpene production in Saccharomyces cerevisiae by fusion of host and heterologous enzymes. *Appl. Environ. Microbiol. 77*, 1033−1040.

(56) Dueber, J. E., Wu, G. C., Malmirchegini, G. R., Moon, T. S., Petzold, C. J., Ullal, A. V., Prather, K. L. J., and Keasling, J. D. (2009) Synthetic protein scaffolds provide modular control over metabolic flux. *Nat. Biotechnol. 27*, 753−759.

(57) Bastian, S., Liu, X., Meyerowitz, J. T., Snow, C. D., Chen, M. M. Y., and Arnold, F. H. (2011) Engineered ketol-acid reductoisomerase and alcohol dehydrogenase enable anaerobic 2-methylpropan-1-ol production at theoretical yield in Escherichia coli. *Metab. Eng. 13*, 345−352.

(58) Fasan, R., Crook, N. C., Peters, M. W., Meinhold, P., Buelter, T., Landwehr, M., Cirino, P. C., and Arnold, F. H. (2011) Improved product-per-glucose yields in P450-dependent propane biotransformations using engineered Escherichia coli. *Biotechnol. Bioeng. 108*, 500−510.

(59) Böttcher, D., and Bornscheuer, U. T. (2010) Protein engineering of microbial enzymes. *Curr. Opin. Microbiol. 13*, 274−282.

(60) Stefan, L. (2010) Beyond directed evolution—semi-rational protein engineering and design. *Curr. Opin. Biotechnol. 21*, 734−743.

(61) Das, R., and Baker, D. (2008) Macromolecular modeling with rosetta. *Annu. Rev. Biochem. 77*, 363−382.

(62) Kiefer, F., Arnold, K., Künzli, M., Bordoli, L., and Schwede, T. (2009) The SWISS-MODEL Repository and associated resources. *Nucleic Acids Res. 37*, D387−392.

(63) Joo, K., Lee, J., Seo, J., Lee, K., Kim, B., and Lee, J. (2009) All-atom chain-building by optimizing MODELLER energy function using conformational space annealing. *Proteins: Struct., Funct., Bioinf. 75*, 1010−1023.

(64) Fiser, A., and Šali, A. (2003) Modeller: Generation and Refinement of Homology-Based Protein Structure Models. In *Macromolecular Crystallography, Part D*, pp 461−491, Academic Press, New York.

(65) Phillips, J. C., Braun, R., Wang, W., Gumbart, J., Tajkhorshid, E., Villa, E., Chipot, C., Skeel, R. D., Kalé, L., and Schulten, K. (2005) Scalable molecular dynamics with NAMD. *J. Comput. Chem. 26*, 1781−1802.

(66) Das, R., André, I., Shen, Y., Wu, Y., Lemak, A., Bansal, S., Arrowsmith, C. H., Szyperski, T., and Baker, D. (2009) Simultaneous prediction of protein folding and docking at high resolution. *Proc. Natl. Acad. Sci. U.S.A. 106*, 18978−18983.

(67) Krieger, E., Koraimann, G., and Vriend, G. (2002) Increasing the precision of comparative models with YASARA NOVA—a self-parameterizing force field. *Proteins 47*, 393−402.

(68) Greenhagen, B. T., O'Maille, P. E., Noel, J. P., and Chappell, J. (2006) Identifying and manipulating structural determinates linking catalytic specificities in terpene synthases. *Proc. Natl. Acad. Sci. U.S.A. 103*, 9826−9831.

(69) Baugh, L., Le Trong, I., Cerutti, D. S., Mehta, N., Gülich, S., Stayton, P. S., Stenkamp, R. E., and Lybrand, T. P. (2012) Second-contact shell mutation diminishes streptavidin-biotin binding affinity through transmitted effects on equilibrium dynamics. *Biochemistry 51*, 597−607.

(70) Eiben, C. B., Siegel, J. B., Bale, J. B., Cooper, S., Khatib, F., Shen, B. W., Players, F., Stoddard, B. L., Popovic, Z., and Baker, D. (2012) Increased Diels-Alderase activity through backbone remodeling guided by Foldit players. *Nat. Biotechnol. 30*, 190−192.

(71) Bershtein, S., Goldin, K., and Tawfik, D. S. (2008) Intense neutral drifts yield robust and evolvable consensus proteins. *J. Mol. Biol. 379*, 1029−1044.

(72) Jonson, P. H., and Petersen, S. B. (2001) A critical view on conservative mutations. *Protein Eng. 14*, 397−402.

(73) Amar, D., Berger, I., Amara, N., Tafa, G., Meijler, M. M., and Aharoni, A. (2012) The transition of human estrogen sulfotransferase from generalist to specialist using directed enzyme evolution. *J. Mol. Biol. 416*, 21−32.

(74) Goldsmith, M., and Tawfik, D. S. (2012) Directed enzyme evolution: beyond the low-hanging fruit. *Curr. Opin. Struct. Biol. 22*, 406−412.

(75) Herman, A., and Tawfik, D. S. (2007) Incorporating Synthetic Oligonucleotides via Gene Reassembly (ISOR): a versatile tool for generating targeted libraries. *Protein Eng., Des. Sel. 20*, 219−226.

(76) Shivange, A. V., Marienhagen, J., Mundhada, H., Schenk, A., and Schwaneberg, U. (2009) Advances in generating functional diversity for directed protein evolution. *Curr. Opin. Chem. Biol. 13*, 19−25.

(77) Dietrich, J. A., McKee, A. E., and Keasling, J. D. (2010) High-throughput metabolic engineering: advances in small-molecule screening and selection. *Annu. Rev. Biochem. 79*, 563−590.

(78) Antipov, E., Cho, A. E., Wittrup, K. D., and Klibanov, A. M. (2008) Highly L and D enantioselective variants of horseradish peroxidase discovered by an ultrahigh-throughput selection method. *Proc. Natl. Acad. Sci. U.S.A. 105*, 17694−17699.

(79) Reetz, M. T. (2009) Directed evolution of enantioselective enzymes: an unconventional approach to asymmetric catalysis in organic chemistry. *J. Org. Chem. 74*, 5767−5778.

(80) Varadarajan, N., Gam, J., Olsen, M. J., Georgiou, G., and Iverson, B. L. (2005) Engineering of protease variants exhibiting high catalytic activity and exquisite substrate selectivity. *Proc. Natl. Acad. Sci. U.S.A. 102*, 6855−6860.

(81) Wang, C., Oh, M. K., and Liao, J. C. (2000) Directed evolution of metabolically engineered Escherichia coli for carotenoid production. *Biotechnol. Prog. 16*, 922−926.

(82) Li, B., Yu, J. P. J., Brunzelle, J. S., Moll, G. N., van der Donk, W. A., and Nair, S. K. (2006) Structure and mechanism of the lantibiotic cyclase involved in nisin biosynthesis. *Science 311*, 1464−1467.

(83) Karsten, W. E., Chooback, L., Liu, D., Hwang, C. C., Lynch, C., and Cook, P. F. (1999) Mapping the active site topography of the NAD-malic enzyme via alanine-scanning site-directed mutagenesis. *Biochemistry 38*, 10527−10532.

(84) Lee, S., Shin, S. Y., Lee, Y., Park, Y., Kim, B. G., Ahn, J.-H., Chong, Y., Lee, Y. H., and Lim, Y. (2011) Rhamnetin production based on the rational design of the poplar O-methyltransferase enzyme and its biological activities. *Bioorg. Med. Chem. Lett. 21*, 3866−3870.

(85) Liu, L., Schmid, R. D., and Urlacher, V. B. (2010) Engineering cytochrome P450 monooxygenase CYP 116B3 for high dealkylation activity. *Biotechnol. Lett. 32*, 841−845.

(86) Chen, I., Dorr, B. M., and Liu, D. R. (2011) A general strategy for the evolution of bond-forming enzymes using yeast display. *Proc. Natl. Acad. Sci. U.S.A. 108*, 11399−11404.

(87) Kabumoto, H., Miyazaki, K., and Arisawa, A. (2009) Directed evolution of the actinomycete cytochrome P450moxA (CYP105) for enhanced activity. *Biosci. Biotechnol. Biochem. 73*, 1922−1927.

(88) Röthlisberger, D., Khersonsky, O., Wollacott, A. M., Jiang, L., DeChancie, J., Betker, J., Gallaher, J. L., Althoff, E. A., Zanghellini, A., Dym, O., Albeck, S., Houk, K. N., Tawfik, D. S., and Baker, D. (2008) Kemp elimination catalysts by computational enzyme design. *Nature 453*, 190−195.

(89) López-Gallego, F., Wawrzyn, G. T., and Schmidt-Dannert, C. (2010) Selectivity of fungal sesquiterpene synthases: role of the active site's H-1 alpha loop in catalysis. *Appl. Environ. Microbiol. 76*, 7723−7733.

(90) Seifert, A., Vomund, S., Grohmann, K., Kriening, S., Urlacher, V. B., Laschat, S., and Pleiss, J. (2009) Rational design of a minimal and highly enriched CYP102A1 mutant library with improved regio-, stereo- and chemoselectivity. *ChemBioChem 10*, 853−861.
(91) Reetz, M. T., and Zheng, H. (2011) Manipulating the expression rate and enantioselectivity of an epoxide hydrolase by using directed evolution. *ChembBioChem 12*, 1529−1535.
(92) Otten, L. G., Hollmann, F., and Arends, I. W. C. E. (2010) Enzyme engineering for enantioselectivity: from trial-and-error to rational design? *Trends Biotechnol. 28*, 46−54.
(93) Idicula-Thomas, S., and Balaji, P. V. (2007) Correlation between the structural stability and aggregation propensity of proteins. *In Silico Biol. 7*, 225−237.
(94) Somero, G. N. (1995) Proteins and Temperature. *Annu. Rev. Physiol. 57*, 43−68.
(95) Magnan, C. N., Randall, A., and Baldi, P. (2009) SOLpro: accurate sequence-based prediction of protein solubility. *Bioinformatics 25*, 2200−2207.
(96) Cabrita, L. D., Gilis, D., Robertson, A. L., Dehouck, Y., Rooman, M., and Bottomley, S. P. (2007) Enhancing the stability and solubility of TEV protease using in silico design. *Protein Sci. 16*, 2360−2367.
(97) Mansfeld, J., Vriend, G., Dijkstra, B. W., Veltman, O. R., Van den Burg, B., Venema, G., Ulbrich-Hofmann, R., and Eijsink, V. G. (1997) Extreme stabilization of a thermolysin-like protease by an engineered disulfide bond. *J. Biol. Chem. 272*, 11152−11156.
(98) Marshall, S. A., Morgan, C. S., and Mayo, S. L. (2002) Electrostatics significantly affect the stability of designed homeodomain variants. *J. Mol. Biol. 316*, 189−199.
(99) Eijsink, V. G. H., Bjørk, A., Gåseidnes, S., Sirevåg, R., Synstad, B., van den Burg, B., and Vriend, G. (2004) Rational engineering of enzyme stability. *J. Biotechnol. 113*, 105−120.
(100) Heddle, C., and Mazaleyrat, S. L. (2007) Development of a screening platform for directed evolution using the reef coral fluorescent protein ZsGreen as a solubility reporter. *Protein Eng., Des. Sel. 20*, 327−337.
(101) Jochens, H., Aerts, D., and Bornscheuer, U. T. (2010) Thermostabilization of an esterase by alignment-guided focussed directed evolution. *Protein Eng., Des. Sel. 23*, 903−909.
(102) Asano, Y., Dadashipour, M., Yamazaki, M., Doi, N., and Komeda, H. (2011) Functional expression of a plant hydroxynitrile lyase in Escherichia coli by directed evolution: creation and characterization of highly in vivo soluble mutants. *Protein Eng., Des. Sel. 24*, 607−616.
(103) Fox, R. J., and Huisman, G. W. (2008) Enzyme optimization: moving from blind evolution to statistical exploration of sequence-function space. *Trends Biotechnol. 26*, 132−138.
(104) Caspi, R., Foerster, H., Fulcher, C. A., Kaipa, P., Krummenacker, M., Latendresse, M., Paley, S., Rhee, S. Y., Shearer, A. G., Tissier, C., Walk, T. C., Zhang, P., and Karp, P. D. (2008) The MetaCyc Database of metabolic pathways and enzymes and the BioCyc collection of Pathway/Genome Databases. *Nucleic Acids Res. 36*, D623−631.
(105) Kanehisa, M., Goto, S., Sato, Y., Furumichi, M., and Tanabe, M. (2011) KEGG for integration and interpretation of large-scale molecular data sets. *Nucleic Acids Res. 40*, D109−D114.
(106) Nobeli, I., Favia, A. D., and Thornton, J. M. (2009) Protein promiscuity and its implications for biotechnology. *Nat. Biotechnol. 27*, 157−167.
(107) Yu, X., Xie, X., and Li, S.-M. (2011) Substrate promiscuity of secondary metabolite enzymes: prenylation of hydroxynaphthalenes by fungal indole prenyltransferases. *Appl. Microbiol. Biotechnol. 92*, 737−748.
(108) Janocha, S., Bichet, A., Zöllner, A., and Bernhardt, R. (2011) Substitution of lysine with glutamic acid at position 193 in bovine CYP11A1 significantly affects protein oligomerization and solubility but not enzymatic activity. *Biochim. Biophys. Acta 1814*, 126−131.
(109) Morawski, B., Quan, S., and Arnold, F. H. (2001) Functional expression and stabilization of horseradish peroxidase by directed evolution in Saccharomyces cerevisiae. *Biotechnol. Bioeng. 76*, 99−107.
(110) Diaz, J. E., Lin, C.-S., Kunishiro, K., Feld, B. K., Avrantinis, S. K., Bronson, J., Greaves, J., Saven, J. G., and Weiss, G. A. (2011) Computational design and selections for an engineered, thermostable terpene synthase. *Protein Sci. 20*, 1597−1606.