

Beyond Protein Engineering: its Applications in Synthetic Biology

Yougen Li*

DuPont Central Research & Development, Biochemical Sciences & Engineering, 200 Powder Mill Road, Wilmington, DE 19803 USA

Introduction

After decades of development, protein engineering has become sufficiently important to industry that its focus has shifted from technology innovation to application exploration. While protein engineering continues to play an important role in established industries like detergent enzymes, it also promises to provide new solutions in emerging industries such as biopharmaceuticals [evolved biologics] and renewable energy [engineered enzymes for biofuel production].

In contrast to protein engineering, synthetic biology is still in its infancy, even though the term, “synthetic biology”, was coined over thirty years ago [1]. Synthetic biology has great potential for enabling sustainable economic growth. Indeed, synthetic pathways can produce many chemicals, including biofuels and drugs, from renewable resources. Currently, synthetic pathways are assembled from biological components culled from nature, and may not be optimal since those biological components were evolved by nature to benefit their native species’ fitness and were not customized for performance in synthetic pathways. Biological components from different species [or even different kingdoms] may not function optimally when simply put together in a pathway. Therefore, to meet industrial standards, synthetic pathways usually require optimization to produce chemicals economically, and this is where protein engineering can play significant roles. In fact, protein engineering has already been applied to optimize a number of synthetic biological components, from pathway enzymes to regulatory elements. It has also been used to balance pathway redox equivalents and control the expression of pathway genes. This editorial highlights applications of protein engineering to synthetic biology in the areas of improving enzyme activity, changing substrate or product specificity, modifying cofactor usage for redox balance, tuning protein expression, and creating scaffold multi-enzyme complexes for metabolite flux control. Due to limited space, only one to two examples are given for each application to demonstrate the concept.

Enzyme Engineering to Improve Enzyme Performance in a Pathway

Creation of a synthetic pathway often involves heterologous expression of multiple enzymes from diverse species. The properties of one or more wild type enzymes, such as activity or specificity, may not meet industrial needs. The problem of low enzyme activity may be addressed by simply increasing the expression level, but this approach may increase the cellular protein burden and may reduce the cellular capacity to produce other necessary pathway enzymes. Changes to other enzymatic properties like substrate specificity require alteration of the enzymes themselves. Protein engineering is an efficient way to change an enzyme’s kinetic parameters to improve its *in vivo* performance, and thus enhance pathway productivity.

One approach to increase pathway productivity is to improve the specific activity of a critical or rate-limiting enzyme through protein engineering. There are many successful examples of this. To increase the production of polyhydroxyalkanoate [PHA] in *E. coli*, Amara et al. [2] engineered a PHA synthase from *Aeromonas punctata* using the mutator strain *E. coli* XL1-Red. They screened about 200,000 mutants

on medium containing Nile red stain and identified five variants with enhanced fluorescence. Four of them exhibited both elevated *in vitro* and *in vivo* PHA synthase activity, with as much as 5-fold improvement over the wild type. This improvement in enzymatic activity resulted in over 126% PHA accumulation in cells, and all five mutants increased PHA weight-average molar mass yield [2].

The biosynthesis of levopimaradiene, the diterpenoid precursor of pharmaceutically important ginkgolides, starts from precursors, isopentenyl diphosphate [IPP] and dimethylallyl diphosphate [DMAPP], which are derived from the 2-C-methyl-D-erythritol-4-phosphate [MEP] pathway in *E. coli*. The condensation of IPP and DMAPP to geranylgeranyl diphosphate [GGPP] is catalyzed by GGPP synthase [GGPPS] while the conversion of GGPP to levopimaradiene [LP] is catalyzed by LP synthase [LPS]. Leonard et al. [3] systematically elevated metabolic flux toward IPP and DMAPP but failed to increase LP titers. However, subsequent combinatorial mutagenesis [site directed mutagenesis and site-saturation mutagenesis based on homology modeling and phylogenetic analysis] of GGPPS and LPS followed by library screening identified improved variants of both enzymes. Combined with the improvements in precursor flux, LP productivity was improved by 2,600-fold [3].

In certain scenarios, an enzyme employed in a pathway need catalyze the conversion of a non-native substrate. Promiscuous enzymes are suitable for this purpose, but low substrate specificity may reduce desired product yield due to inefficient substrate conversion and byproduct formation. Altering and/or improving substrate specificity can lead to higher yields of desired pathway products. To reduce the formation of byproducts, Zhang et al. increased the substrate specificity of 2-ketoisovalerate decarboxylase from *Lactococcus lactis* on a desired substrate, 2-keto-4-methylhexanoate, by modifying its substrate-binding sites based on a homology model. The authors also altered the substrate specificity of another pathway enzyme, 2-isopropylmalate synthase [LeuA] from *E. coli*. Based on its crystal structure, the authors introduced mutations that enlarged the substrate-binding pocket to accommodate the larger unnatural substrate, 2-keto-3-methylvalerate. This further increased the pathway’s productivity of nonnatural alcohols up to C₈ [4]. In another attempt to produce non-natural chemicals using a synthetic pathway, Zhang et al. [5] changed the substrate preference of *E. coli* glutamate dehydrogenase [GDH] from 2-ketoglutarate to 2-ketobutyrate to produce L-homoalanine, a valuable drug precursor. Together with other pathway optimization

*Corresponding author: Yougen Li, DuPont Central Research & Development, Biochemical Sciences & Engineering, 200 Powder Mill Road, Wilmington, DE 19803 USA, Tel: 302-695-4431; Fax: 302-695-1351; E-mail: yougen.li@usa.dupont.com

Received August 02, 2012; Accepted August 04, 2012; Published August 09, 2012

Citation: Li Y (2012) Beyond Protein Engineering: its Applications in Synthetic Biology. *Enz Eng* 1:e103. doi:10.4172/2329-6674.1000e103

Copyright: © 2012 Li Y. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

steps, the authors managed to produce 5.4 g/L of L-homoalanine from a feedstock of 30 g/L glucose [5]. The work of Zhang et al. [5] has demonstrated that synthetic pathways can be utilized to produce many non-natural chemicals. Moreover, these efforts have shown that the productivity of synthetic pathways can be dramatically enhanced through the alteration of pathway enzyme specificity.

Some enzymes produce multiple products from the same substrate. If such enzymes are to be employed in synthetic pathways, their product specificity may require narrowing to improve the productivity of the desired pathway product. For example, γ -humulene synthase [HUM] from *Abies grandis* can produce as many as 52 different sesquiterpenes from a sole substrate, farnesyl diphosphate [FPP]. To alter the product specificity of HUM, Yoshikuni et al. [6] first identified 19 amino acid residues as targets for mutagenesis based on a homology model. After performing saturation mutagenesis on these residues, they found that only four of them had a significant effect on catalysis. The authors then developed an algorithm to systematically recombine the mutations found to alter product specificity. This resulted in the creation of seven HUM variants, each of which specifically produces a different product [6]. The same team later changed the product specificity of the (+)- δ -cadinene synthase from *Gossypium arboreum* to create an entirely novel sesquiterpene synthase that produces germacrene D-4-ol from FPP, in order to meet the requirement of a synthetic pathway [7].

Engineering cofactor utilization of enzymes in metabolic pathways

Redox [reduction-oxidation] reactions are chemical reactions involving changes in the oxidation state of molecules. In a typical pathway, redox is reflected by the status of pyridine nucleotides, NAD[H] and NADP[H]. Redox balance means that the production and consumption of reduced cofactors [NADPH and/or NADH] are approximately equal. Unbalanced redox can damage cells, waste energy and/or carbon, and even lead to metabolic arrest [8]. Fortunately, redox-imbalanced pathways can sometimes be re-balanced through cofactor engineering. Two approaches to cofactor engineering are (1) removal of excess cofactors via enzymatic conversion [e.g., by expression of a cofactor oxidase or transhydrogenase], and, (2) switching the cofactor-utilization preference [specificity] of pathway enzymes.

The concentration of NADH in a cell is generally about an order of magnitude higher than that of NADPH. [For example, the reported value for NAD[H] is 1.3 μ M and that for NADP[H] is 0.39 μ M for exponentially-growing *Escherichia coli* [9]. In addition, NADH is more stable than NADPH, and is also roughly an order of magnitude less expensive. Therefore, many efforts to balance redox aim for pathway utilization of NADH.

To optimize the vitamin C biosynthetic pathway, Banta et al. [9] switched the cofactor preference of *Corynebacterium* 2,5-Diketo-D-gluconic acid [2,5-DKG] reductase from NADPH to NADH. The authors first subjected the binding site of the 2'-phosphate group of NADPH to site-directed mutagenesis [3]. Several beneficial mutations were identified and were combined with other substitutions suggested by the sequences of wild type NADH-dependent aldo-keto reductases. The resultant quadruple mutant, F22Y/K232G/R238H/A272G, has NADH-dependent activity over 2 orders of magnitude greater than that of wild type 2,5-DKG reductase [3,5].

The status of the reduced forms of NAD[H] and NADP[H] may greatly affect product yield or the formation of byproducts [8]. To test this idea, Heux et al. [8] expressed an NADH oxidase from *Lactococcus lactis* in *S. cerevisiae* strain V5 to reduce the intracellular NADH

pool [8]. Under controlled microaerobic conditions, the intracellular NADH concentration was reduced five-fold and the NADH/NAD⁺ ratio was decreased six-fold. Several metabolic fluxes were found to be redistributed in response to the imposed NADH consumption, presumably to maintain redox balance. The production levels of ethanol, glycerol, succinate and hydroxyglutarate were significantly decreased, whereas the levels of many oxidized metabolites such as acetaldehyde, acetate and acetoin were increased [8]. At a glucose concentration of 10% [w/v], the biomass yield and glucose consumption were both lower than in the control strain lacking heterologous NADH oxidase [8]. This work demonstrated that when cellular redox balance is changed, cells may be able to make their own adjustments, and this may affect pathway productivity.

Another way to balance redox is to remove excess cofactor by converting it into a more useful form [e.g. NADH to NADPH]. Pyridine nucleotide transhydrogenase catalyzes the reversible transfer of a hydride ion between NADH and NADP⁺. By co-expressing NADP⁺-dependent alcohol dehydrogenase and NAD⁺-dependent formate dehydrogenase, Weckbecker and Hummel managed to improve the production of chiral alcohols in *E. coli* [10]. This strategy has limited applicability, however. For example, the conversion of NADH to NADPH requires a proton motive force and one proton per hydride. This consumes energy and can substantially decrease a pathway's productivity [11].

Engineering of genetic regulatory elements to tune protein expression

Unlike native cellular pathways [e.g., many secondary metabolic pathways], which are usually tightly-regulated, synthetic pathways may not be as well-tuned to respond to dynamic internal and external environments. Up- or down- regulation of the expression of some pathway genes is generally needed to achieve balanced expression of the various pathway enzymes. Genes are regulated at several levels: transcription, post-transcription, translation, and post-translation. In response to changes in the concentration of effector molecules [e.g., activators or inhibitors], transcriptional regulatory proteins can regulate mRNA expression and allosteric regulatory domains in enzymes can modulate enzymatic activity [12]. Regulatory proteins can thus control protein expression or metabolic flux. Protein engineering concepts can also be applied to engineer non-protein regulatory elements, like promoters or other DNA or RNA regulatory regions to affect transcription and/or translation and thus control protein expression.

Pathway productivity can be increased by reducing the inhibition of pathway enzymes by products, intermediates, or other inhibitory small molecules. Such inhibition can be modulated through an allosteric regulatory domain. Deng et al. created error-prone PCR libraries of glucosamine synthase [GlmS] and screened for mutants that were less inhibited by the intermediate, glucosamine-6-phosphate. They discovered a GlmS variant that increased glucosamine production by 15-fold [13]. Baez-Viveros et al. [14] engineered chorismate mutase-prephenate dehydratase [CM-PDT] for resisting feedback inhibition and managed to enhance the yield of L-phenylalanine by 1.6-fold using the evolved CM-PDT in concert with other optimization strategies [14].

DNA regulatory regions, like promoters, operators, and enhancers, affect genetic transcription and consequently, protein expression. Many synthetic promoters have been engineered to fine-tune pathway enzyme expression [15]. A library of mutated promoters is usually

cloned upstream of a reporter gene [e.g., a fluorescent protein]. After screening and identification of desirable promoter variants, the reporter genes are replaced with pathway genes. Alper et al. [1] created an error-prone PCR library of $P_{L-\lambda}$ promoter, which was cloned into a reporter plasmid upstream of a green fluorescent protein [GFP] gene [1]. After screening in *E. coli*, nearly 200 promoter variants with a wide range of strengths were generated. The authors applied the same concept to engineer the TEF1 promoter with yellow fluorescent protein [YFP] as the reporter gene in *S. cerevisiae*. The resulting mutated promoters also covered a wide range of strengths, and can in principle serve as a “catalog” of choices to fine-tune the expression of a multitude of pathway genes [1].

RNA molecules are well-known to form secondary structural elements. These folds can have a great impact on gene expression by modulating mRNA stability or controlling the rate of translation initiation. Pflieger et al. [16] generated libraries of tunable intergenic regions [TIGRs] with various mRNA secondary structures and RNase cleavage sites to coordinate the expression of multiple genes in a heterologous mevalonate biosynthetic pathway. By this approach, they improved mevalonate production by 7-fold [16]. Smolke et al. used specially-designed RNA sequences to control gene expression both temporally and spatially [17], and exploited this technology to increase secondary metabolite biosynthesis in yeast [18].

Many other genetic regulatory elements, like ribosome binding sites [19], the ribosome itself [20], and operon structures [21], have also been engineered to alter pathway gene expression and affect pathway productivity.

Enzyme scaffold engineering to control metabolite flux

Pathway intermediates can be toxic to host cells, consumed by endogenous cellular activities, or reduced through secretion [22]. The issues could be mitigated by spatially arranging pathway enzymes so the substrate moves efficiently from one step to the next. The increased local concentration of pathway intermediates, reduction of opportunities for diversion of intermediates to other cellular enzymes, and decreased chance of secretion are all potential advantages afforded by this approach. Scaffold proteins are known in nature that spatially organize specific proteins in the cell, and are very important for coordinating certain cellular activities [23]. Recently, researchers have exploited such scaffolds to co-localize synthetic pathway enzymes into complexes with optimal enzyme stoichiometries.

Inspired by natural substrate-channeling systems, Dueber et al. [24] engineered protein scaffolds using protein-protein interaction domains and ligands from metazoan cells [mouse SH3 and PDZ domains, and rat GBD domain] for pathway enzymes in an effort to improve pathway productivity [24]. The three-enzyme pathway [acetoacetyl-CoA thiolase [AtoB], hydroxyl-methylglutaryl-CoA synthase [HMGS] and hydroxyl-methylglutaryl-CoA reductase [HMGR]] for producing mevalonate from acetyl-CoA was chosen as the model system to demonstrate the idea. The scaffolds consist of the three domains, each with a variable copy number [e.g. $[GBD]_x[SH3]_y[PDZ]_z$]. The three pathway enzymes attach themselves to the scaffold via cognate ligand tags [AtoB with a GBD tag, HMGS with an SH3 tag, and HMGR with a PDZ tag]. Through varying x, y, and z, the pathway was optimized and the product titer was increased by 77-fold [24]. Because the pathway enzymes are co-localized in the same scaffold and enzyme stoichiometry can be independently adjusted, the optimal balance of enzyme activities within the complex can be achieved, improving pathway productivity. Using a similar scaffold design, Moon et al. [25]

increased glucaric acid titer by 5-fold compared to the non-scaffolded control [25].

Conclusion and Perspectives

In the past several decades, particularly the past two, many robust and efficient protein engineering technologies have been developed. These technologies are now starting to be employed in a very exciting and nascent field – synthetic biology, as reviewed in this editorial. Some concepts from protein engineering have also been borrowed by synthetic biologists. For example, rational design and directed evolution, two general strategies for protein engineering [28], have been deployed in synthetic biology [26 27]. More techniques of protein engineering are expected to be used in synthetic biology in the future.

One area of protein engineering not mentioned in this editorial involves unnatural amino acids [29]. There are a few examples of the application of unnatural amino acids in synthetic biology [30]. Since unnatural amino acids can be used to functionalize proteins and immobilize pathway enzymes *in vivo*, one potential application is to channel substrates for synthetic pathways to improve pathway productivity, in a similar fashion to the synthetic scaffold structure approach described above [24].

With the rapid development of “-omics” technologies, an overwhelming amount of biological data have become available [31]. In addition, computational power has been dramatically expanding and molecular modeling has become more accurate. All these developments make rational design more attractive, rapid, and powerful. As a result, more rational and semi-rational designs are likely to be seen in protein engineering and its applications to synthetic biology going forward.

While we appreciate the importance of protein engineering to synthetic biology, we should not underestimate the role of synthetic biology in further advancing protein engineering. For example, synthetic biology can be used to develop more efficient high-throughput screening methods to capitalize on protein engineering techniques [32].

It is clear that the marriage of protein engineering and synthetic biology promises to be a fruitful one. Trained to think on a molecular level, protein engineers and synthetic biologists routinely “think big” while working on the small [e.g. microbial cells]. And though these ideas are all implemented first in microscopic single cells, the biggest of them will surely do nothing less than change the world.

Acknowledgement

The author thanks Alex Tobias and Jenafer Evans for valuable editing help.

References

1. Alper H, Fischer C, Nevoigt E, Stephanopoulos G (2005) Tuning genetic control through promoter engineering. *Proc Natl Acad Sci U S A* 102: 12678-12683.
2. Amara AA, Steinbuechel A, Rehm BH (2002) In vivo evolution of the *Aeromonas punctata* polyhydroxyalkanoate (PHA) synthase: isolation and characterization of modified PHA synthases with enhanced activity. *Appl Microbiol Biotechnol* 59: 477-482.
3. Leonard E, Ajikumar PK, Thayer K, Xiao WH, Mo JD, et al. (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.
4. Zhang K, Sawaya MR, Eisenberg DS, Liao JC (2008) Expanding metabolism for biosynthesis of nonnatural alcohols. *Proc Natl Acad Sci U S A* 105: 20653-20658.
5. Zhang K, Li H, Cho KM, Liao JC (2010) Expanding metabolism for total biosynthesis of the nonnatural amino acid L-homoalanine. *Proc Natl Acad Sci U S A* 107: 6234-6239.

6. Yoshikuni Y, Ferrin TE, Keasling JD (2006) Designed divergent evolution of enzyme function. *Nature* 440: 1078-1082.
7. Yoshikuni Y, Martin VJ, Ferrin TE, Keasling JD (2006) Engineering cotton (+)-delta-cadinene synthase to an altered function: germacrene D-4-ol synthase. *Chem Biol* 13: 91-98.
8. Heux S, Cachon R, Dequin S (2006) Cofactor engineering in *Saccharomyces cerevisiae*: Expression of a H₂O-forming NADH oxidase and impact on redox metabolism. *Metab Eng* 8: 303-314.
9. Banta S, Swanson BA, Wu S, Jarnagin A, Anderson S, et al. (2002) Alteration of the specificity of the cofactor-binding pocket of *Corynebacterium* 2,5-diketo-D-gluconic acid reductase A. *Protein Eng* 15: 131-140.
10. Weckbecker A, Hummel W (2004) Improved synthesis of chiral alcohols with *Escherichia coli* cells co-expressing pyridine nucleotide transhydrogenase, NADP⁺-dependent alcohol dehydrogenase and NAD⁺-dependent formate dehydrogenase. *Biotechnol Lett* 26: 1739-1744.
11. Chin JW, Khankal R, Monroe CA, Maranas CD, Cirino PC, et al. (2009) Analysis of NADPH supply during xylitol production by engineered *Escherichia coli*. *Biotechnol Bioeng* 102: 209-220
12. Foo JL, Ching CB, Chang MW, Leong SS (2012) The imminent role of protein engineering in synthetic biology. *Biotechnol Adv* 30: 541-549.
13. Deng MD, Severson DK, Grund AD, Wassink SL, Burlingame RP, et al. (2005) Metabolic engineering of *Escherichia coli* for industrial production of glucosamine and N-acetylglucosamine. *Metab Eng* 7: 201-214.
14. Báez-Viveros JL, Osuna J, Hernández-Chávez G, Soberón X, Bolívar F, et al. (2004) Metabolic engineering and protein directed evolution increase the yield of L-phenylalanine synthesized from glucose in *Escherichia coli*. *Biotechnol Bioeng* 87: 516-524.
15. Seo SW, Kim SC, Jung GY (2012) Synthetic regulatory tools for microbial engineering *Biotechnol Bioeng* 17: 1-7.
16. Pfleger BF, Pitera DJ, Smolke CD, Keasling JD (2006) Combinatorial engineering of intergenic regions in operons tunes expression of multiple genes. *Nat Biotechnol* 24: 1027-1032.
17. Chang AL, Wolf JJ, Smolke CD (2012) Synthetic RNA switches as a tool for temporal and spatial control over gene expression. *Curr Opin Biotechnol*.
18. Siddiqui MS, Thodey K, Trenchard I, Smolke CD (2012) Advancing secondary metabolite biosynthesis in yeast with synthetic biology tools. *FEMS Yeast Res* 12: 144-170.
19. Salis HM, Mirsky EA, Voigt CA (2009) Automated design of synthetic ribosome binding sites to control protein expression. *Nat Biotechnol* 27: 946-950.
20. Xia XX, Qian ZG, Ki CS, Park YH, Kaplan DL et al. (2010) Native-sized recombinant spider silk protein produced in metabolically engineered *Escherichia coli* results in a strong fiber. *Proc Natl Acad Sci U S A* 107: 14059-14063.
21. Lim HN, Lee Y, Hussein R (2011) Fundamental relationship between operon organization and gene expression. *Proc Natl Acad Sci U S A* 108: 10626-10631.
22. Lee H, DeLoache WC, Dueber JE (2012) Spatial organization of enzymes for metabolic engineering. *Metab Eng* 14: 242-251.
23. Good MC, Zalatan JG, Lim WA (2011) Scaffold proteins: hubs for controlling the flow of cellular information. *Science* 332: 680-686.
24. Dueber JE, Wu GC, Malmirchegini GR, Moon TS, Petzold CJ et al. (2009) Synthetic protein scaffolds provide modular control over metabolic flux. *Nat Biotechnol* 27: 753-759.
25. Moon TS, Dueber JE, Shiue E, Prather KL (2010) Use of modular, synthetic scaffolds for improved production of glucaric acid in engineered *E. coli*. *Metab Eng* 12: 298-305.
26. Cobb RE, Si T, Zhao H (2012) Directed evolution: an evolving and enabling synthetic biology tool. *Curr Opin Chem Biol* 16: 285-291
27. Cambray G, Mutalik VK, Arkin AP (2011) Toward rational design of bacterial genomes. *Curr Opin Microbiol*.14: 624-630.
28. Yip SH, Foo JL, Schenk G, Gahan LR, Carr PD, et al. (2011) Directed evolution combined with rational design increases activity of GpdQ toward a non-physiological substrate and alters the oligomeric structure of the enzyme. *Protein Eng Des Sel* 24: 861-872.
29. Wang K, Schmied WH, Chin JW (2012) Reprogramming the genetic code: from triplet to quadruplet codes. *Angew Chem Int Ed Engl* 51: 2288-2297.
30. Voloshchuk N, Montclare JK (2010) Incorporation of unnatural amino acids for synthetic biology. *Mol Biosyst* 6: 65-80.
31. Fukushima A, Kusano M, Redestig H, Arita M, Saito K, et al. (2009) Integrated omics approaches in plant systems biology. *Curr Opin Chem Biol* 13: 532-538.
32. Santos CN, Stephanopoulos G (2008) Melanin-based high-throughput screen for L-tyrosine production in *Escherichia coli*. *Appl Environ Microbiol*. 74: 1190-1197.