Available online at www.sciencedirect.com

ScienceDirect

Current Opinion in Biotechnology

Designing artificial metabolic pathways, construction of target enzymes, and analysis of their function

Yutaro Mori and Tomokazu Shirai

Artificial design of metabolic pathways is essential for the production of useful compounds using microbes. Based on this design, heterogeneous genes are introduced into the host, and then various analysis and evaluation methods are conducted to ensure that the target enzyme reactions are functionalized within the cell. In this chapter, we list successful examples of useful compounds produced by designing artificial metabolic pathways, and describe the methods involved in analyzing, evaluating, and optimizing the target enzyme reaction.

Address
Biomass Engineering Research Division, Center for Sustainable Resource Science, RIKEN, 1-7-22, Suehiro-cho, Tsurumi-ku, Yokohama 230-0045, Japan

Corresponding author: Shirai, Tomokazu (tomokazu.shirai@riken.jp)

Current Opinion in Biotechnology 2018, **54**:41–44

This review comes from a themed issue on **Analytical biotechnology**

Edited by **Hiroshi Shimizu** and **Fumio Matsuda**

https://doi.org/10.1016/j.copbio.2018.01.021

Introduction to building artificial metabolic pathways

One of the technologies for utilizing non-fossil raw materials includes the production of useful compounds by microbial 'fermentation'. When utilizing microbes to produce useful compounds, it is vital to design the metabolic reactions optimally at the genomic scale, considering the carbon flow within the cell as well as the balance of production and consumption of energy substances such as ATP, and oxidation–reduction substances [1–3]. Many artificial metabolic pathways have already been designed and their effectiveness has been verified. There are several reports of artificial designed metabolic pathways actually being introduced into hosts, and subsequent production of both natural and unnatural common compounds from renewable resources [4–6]. Examples include an extremely wide range of products including biofuels (bioethanol and biodiesel) [7,8], polymer materials (muconic acid, maleic acid, adipic acid, 3-hydroxypropionic acid, terephthalic acid) [9–11], food additives (amino acids, GABA, vitamins) [12–14], bulk chemicals (ethylene glycol, butanediol, cadaverine) [15–17], and active pharmaceutical ingredients (artemisin, thebaine) [18,19]. Recently, Luo and Lee successfully produced terephthalic acid using *Escherichia coli* with *p*-xylene as the substrate [20•]. *Comamonas testosteroni* and *Pseudomonas putida*-derived genes were introduced into *E. coli*, and the yield of terephthalic acid was improved by adjusting the expression level of each of these genes. Noda *et al.* also introduced four types of foreign genes into *E. coli* and successfully produced maleic acid from glucose for the first time using fermentative production [21]. These compounds were previously synthesized chemically from fossil raw materials, but these results demonstrate the possibility of producing these compounds from microbes using fermentative production. Undoubtedly, many more useful compounds will be produced by microbial cells in the future.

Recently, isobutanol has attracted much attention as a substitute for jet fuel. Two *Lactococcus lactis*-derived genes were introduced into *E. coli* to successfully construct a pathway producing isobutanol from 2-ketoisovalerate [22]. The above reaction was also introduced into *Saccharomyces cerevisiae*, one of the yeast models, and the isobutanol yield was successfully improved after designing the intracellular metabolism by taking the cytosolic and mitochondrial compartments into consideration [23]. When using isobutanol as jet fuel, it must first go through a chemical process of dehydration and conversion to isobutene; thus, a large amount of active research is currently underway to investigate the direct production of isobutene by microbes [24,25]. Useful compounds have been produced by combining known *in vivo* metabolic reactions that have been compiled in a database. However, it is difficult to achieve microbial production of many useful compounds with this information alone. As databases of enzymatic reactions and compounds have increased in size, simulation tools were developed to construct possible reaction pathways for biosynthesis of the target compounds [26–29]. Based on enzymatic reactions and metabolic pathways on databases such as KEGG (http://www.genome.jp/kegg/) or BRENDA (https://www.brenda-enzymes.org/), information to find heterologous key enzymes are provided. Reaction patters are derived from the chemical structures of precursors and products described with such as SMILES strings in known enzymatic reactions. These are applied to find possible metabolic pathways for synthetic design of target compounds, including putative chemicals and enzymatic

reactions. Using these tools, researchers can cyclopaedically search for possible novel metabolic reactions *in silico*.

Constructing artificial metabolic pathways with expression of heterologous proteins

Synthetic bioengineering involves a repeating cycle of design, build, test and learn, and there are large number of considerations to produce the target substance [30,31]. However, for the progress of the key reaction using heterologous genes, an essential thing to be considered is stable expression of heterologous proteins with enzymatic activity. Standard methods for heterologous protein expression are avoiding stable secondary structure formation in mRNA at the 5′ end region and optimizing codon usage to match the host cell [32]. Recently, a new index of codon selection that affects protein expression has been advocated [33•]. When genes were redesigned based on this index, the mRNA lifetime and concentration were improved, resulting in a notable increase in protein concentration compared to that before redesigning. The codon optimization based on this codon-influence metric could be a new standard in heterologous expression. When heterologous proteins form inclusion bodies, its formation can be prevented by promoting folding through introduction of solubility-tag and co-expression of chaperones [34]. Furthermore, extracting hot spots that cause protein aggregation using developed software has been reported [35]. By applying two rules in this study (the α-helix rule and the hydropathy contradiction rule) with multiple mutations, solubilization of inclusion body formed-proteins was achieved. This method expands the availability of heterologous proteins that were previously difficult to use because of their insolubility. Although the use of heterologous proteins containing disulfide bonds is hard because the reducing environment in the cytoplasm prevent the correct disulfide bond formation in *E. coli*, the soluble expression of these proteins was reported with the co-expression of sulfhydryl oxidase and disulfide bond isomerase [36]. In addition, using heterogeneous enzymes containing coenzymes, it is notable that supplying coenzymes increase the proportion of holoenzymes and the entire enzymatic activity. Lastly, prolonging the lifetime of the protein itself *in vivo* improves the substance production with artificial metabolic pathways. The lifetime of a protein is depending on the N-terminal amino acid residues and changed drastically from a few minutes to ∼100 h (the N-end rule) [37]. The protein that its N-terminal residues are recognized by N-recognin is rapidly degraded by the 26S proteasome. This N-end rule differs depending upon the host cell; hence, it is possible to violate this rule with the heterologous protein expression. In these instances, the N-terminal amino acid sequence should be modified to be lived long. With bacteria, the N-terminal of proteins is normally N-formylmethionine, which is converted to methionine by deformylase, then methionyl aminopeptidase (MAP) removes the methionine and the second amino acid residue becomes the new N-terminal residue. MAP activity is dependent on the second amino acid, therefore, to extend the lifetime of the heterologous proteins in bacteria, the second amino acid residue should be selected not to have the activity of MAP and N-recognin. With the optimization as described above, constructing artificial metabolic pathways with heterologous proteins is achieved.

Enzyme function analysis and evaluation

Following the construction of artificial pathways, the analysis and evaluation of heterologous enzymes is necessary to improve the production of the target compound. Particularly important kinetic parameters include k_{cat}, which indicate the maximum number of substrate molecules that each enzyme converts to the product per unit of time, K_m, the affinity of the enzyme to the substrate, and k_{cat}/K_m the enzymatic efficiency, respectively. Generally, these values calculated experimentally *in vitro* have been used. Although there was skepticism whether these were applicable as the values *in vivo* where a different environment surrounds the enzyme, recently, the correlation of k_{cat} and K_m *in vitro* and *in vivo* has been clarified [38,39•]. *In vivo*, due to the crowded environment, diffusion of the substrate is rate-limited step hence the apparent substrate concentration is decreased compared to that in *in vitro* experiments. Due to the k_{cat} value is independent of substrate concentration but the K_m value is not, the K_m *in vivo* apparently increases, which then reduces the k_{cat}/K_m value with increasing enzyme concentration. In addition, there are various compounds present in the *in vivo* environment; thus, depending on the substrate specificity of the introduced heterologous enzyme, the enzymatic efficiency is further reduced through competitive inhibition. However, it is a complex situation because the inhibitory substance also decreases in apparent concentration, and the effect of competitive inhibition decreases compared to the *in vitro* conditions. Therefore, it should be considered that the enzymatic kinetic parameters *in vitro* cannot be simply applied to the *in vivo* environment. Taking into consideration the differences in the environment surrounding the enzyme both *in vitro* and *in vivo*, investigations into the construction of an enzyme assay with resembling the *in vivo* environment have also been made [40]. Considering that substrate diffusion rate is limited and apparent substrate concentration is decreased in *in vivo* reactions, localizing enzymes involved in the consecutive artificial designed reactions is one solution leading to the highly efficient production of the target compound *via* the cascade reaction. It is known that localizing enzymes is more effective to improve the enzymatic efficiency of the overall reaction with lower substrate concentrations in cascade reaction. In facts, several articles of artificial enzyme complexes to produce target compounds *in vivo* have been reported [41]. Metabolic control analysis (MCA) is a powerful method to analyze the sensitivity of metabolic fluxes to the change in the activities of enzymatic reactions [42]. In MCA, an experiment for perturbation of target enzyme