Review

Advancement of Metabolic Engineering Assisted by Synthetic Biology

Hyang-Mi Lee, Phuong N. L. Vo and Dokyun Na *

School of Integrative Engineering, Chung-Ang University, Seoul 06974, Korea; myhys84@cau.ac.kr (H.-M.L.); lamphuong2895@gmail.com (P.N.L.V.)

* Correspondence: blisszen@cau.ac.kr; Tel.: +82-2-820-5690

Received: 14 November 2018; Accepted: 27 November 2018; Published: 4 December 2018

Abstract: Synthetic biology has undergone dramatic advancements for over a decade, during which it has expanded our understanding on the systems of life and opened new avenues for microbial engineering. Many biotechnological and computational methods have been developed for the construction of synthetic systems. Achievements in synthetic biology have been widely adopted in metabolic engineering, a field aimed at engineering micro-organisms to produce substances of interest. However, the engineering of metabolic systems requires dynamic redistribution of cellular resources, the creation of novel metabolic pathways, and optimal regulation of the pathways to achieve higher production titers. Thus, the design principles and tools developed in synthetic biology have been employed to create novel and flexible metabolic pathways and to optimize metabolic fluxes to increase the cells’ capability to act as production factories. In this review, we introduce synthetic biology tools and their applications to microbial cell factory constructions.

Keywords: synthetic biology; metabolic engineering; microbial cell factory; synthetic metabolic pathways

1. Introduction

Synthetic biology adopts the principles of electrical engineering to rationally design and engineer biological systems. In the early era of synthetic biology, various genetic circuits were constructed, including a genetic toggle switch and a repressilator, which began the rational design of man-made functional biological networks [1,2]. Recently, synthetic biologists have developed various biotechnological and computational methods and tools to manipulate the host genome and create various synthetic systems using various genetic components, such as DNA [3], RNA [4,5], and proteins [6], as well as complex synthetic circuits, such as bio-oscillators [7], toggle switches [1,8], and logic gates [9].

Synthetic systems have been widely applied to metabolic engineering for the biosynthesis of biofuels, commodity chemicals, and pharmaceutical molecules [10–13]. The integration of synthetic biology and metabolic engineering has provided a solution to the current increasing demand for sustainable and renewable resources in response to global concerns, such as fossil fuel depletion and rising healthcare expenses [14,15]. In particular, synthetic biology has provided advanced toolsets and novel systems, which significantly facilitate the development of metabolic engineering [10,16].

For example, recently, the metabolic production of artemisinin (widely used as an antimalarial drug) and its precursor, artemisinic acid, was successfully achieved. Since the supply of artemisinic acid from the plant *Artemisia annua* or via chemical synthesis was insufficient to meet the needs of the market, it was recently produced from the metabolically engineered microbial hosts *Escherichia coli* and *Saccharomyces cerevisiae* with higher productivity and an economically feasible titer and yield [13,17,18]. Micro-organisms have also been engineered to produce fuels for use as alternative and renewable...
sources of energy [19,20]. E. coli was metabolically engineered to produce isopropanol by introducing a synthetic metabolic pathway that converts acetyl-CoA to acetone and finally, to isopropanol. The pathway comprises enzymes originating from other bacterial species, namely, acetyl-coenzyme A [CoA] acetyltransferase (C. acetobutylicum thl), acetoacetate decarboxylase (C. acetobutylicum adc), and secondary alcohol dehydrogenase (C. beijerinckii adh) [21].

Advancements in metabolic engineering assisted by synthetic biology have yielded unprecedented outputs in industrial and pharmaceutical biotechnology [22,23]. Herein, we review the recent advances in various synthetic biology tools and methods for constructing, manipulating, and optimizing metabolic synthetic pathways to achieve microbial cell factory constructions (Figure 1).

**Figure 1.** Synthetic biology tools adopted to metabolic engineering. Pathway construction: A synthetic metabolic pathway can be constructed by assembling heterologous genetic parts (e.g., promoters and coding sequences). An exemplary synthetic pathway is depicted, which consists of heterologous enzymes from three different organisms to convert a substrate (acetyl-CoA) into a desired product (n-butanol): blue, R. eutrophus; red, C. acetobutylicum; green, S. collinus. Pathway regulation and optimization: Synthetic metabolic pathways and their flux optimizations can be achieved by creating synthetic promoter libraries and by constructing a dynamic regulatory pathway. In addition, synthetic regulation factors, such as CRISPRi- and TALE-based transcription factors, and synthetic sRNAs for rational control over translation processes are also used. Spatial organization: Proteins, DNAs, and RNAs can be used as a spatial organization scaffold, where pathway enzymes are physically bound to the scaffold in a designable manner.

### 2. Synthetic Metabolic Pathway Construction and Flux Optimization

Desired chemical compounds produced from micro-organisms are often non-natural, and thus, novel metabolic pathways should be rationally constructed. Synthetic metabolic pathway construction is characterized by the assembly of multiple genes involved in a desired pathway from many different species [24–27]. Unlike natural metabolic pathways that have been optimized through evolution, synthetic pathways composed of enzymes from various species are not optimized in their fluxes. Therefore, toxic intermediates may form and accumulate, thus causing cell growth retardation and production titer reduction. Therefore, the designed synthetic pathways should be further optimized to maximize production of the target substance. Because of the complexity of gene expressions and enzyme kinetics, mathematical and computational methods have also been developed to predict the behavior of the designed pathways and circuits to ensure robustness [25,28–30]. The glossary explains the meaning of specialist terms, which are important for an understanding of the text (Box 1).
Box 1. Glossaries.

<table>
<thead>
<tr>
<th>Term</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>TAL effectors</td>
<td>Transcription activator-like effectors secreted by <em>Xanthomonas</em> bacteria, which recognize plant DNA sequences through a central repeat domain of ~34 amino acid repeats, and activate the expression of individual genes to aid bacterial infection</td>
</tr>
<tr>
<td>Bio-oscillators</td>
<td>Synthetic genetic circuits that mimic the natural genetic clocks of organisms, which induces the periodical behavior of a system</td>
</tr>
<tr>
<td>Toggle switch</td>
<td>Synthetic genetic circuits that display a stepwise function: ON or OFF</td>
</tr>
<tr>
<td>Logic gates</td>
<td>Synthetic genetic circuits consisting of multiple cellular sensors and actuators to perform precise logical operation (e.g., AND-, OR-, NOT-, and NAND-gates)</td>
</tr>
<tr>
<td>Phosphoketolase pathway</td>
<td>The pathway that contributes to carbon metabolism, in which a key enzyme, phosphoketolase, cleaves pentose phosphate into glyceraldehyde-3-phosphate and acetyl phosphate</td>
</tr>
<tr>
<td>Riboswitches</td>
<td>RNA sequences that resize in 5' UTR and regulate transcription and/or translation by conformational changes upon ligand binding</td>
</tr>
<tr>
<td>Small regulatory RNAs (sRNAs)</td>
<td>Small non-coding RNAs that regulate translation by base-pairing with target mRNAs</td>
</tr>
<tr>
<td>Protospacer-adjacent motif (PAM)</td>
<td>5'−NGG−3' sequence immediately following the target DNA sequence in the CRISPR-Cas system</td>
</tr>
</tbody>
</table>

2.1. Pathway Construction

There have been several reports on microbial cell factories that were metabolically engineered to produce bio-substances with enhanced metabolic pathway performance. For example, high-titer production of n-butanol from *E. coli* was achieved with a chimeric pathway assembled from three different organisms [31]. n-Butanol is obtained from acetyl-CoA by condensation of two acetyl-CoA units and subsequent conversions. A major challenge in metabolic pathways is reversibility. Bond-Watts et al. [31] constructed a synthetic pathway comprising enzymes from three different bacterial species: phaA (acetoacetyl-CoA thiolase/synthase) and phaB (3-hydroxybutyryl-CoA dehydrogenase) from *Ralstonia eutrophus*, crt (crotonase), adhE2 (bifunctional butyraldehyde and butanol dehydrogenase) from *C. acetobutylicum*, and ccr (crotonyl-CoA reductase) from *Streptomyces collinus*. The forward reactions of phaA and phaB are favored; thus, they are capable of providing sufficient metabolic flux toward n-butanol. To convert the product of phaB to n-butanol, crt, ccr, and adhE2 genes were also introduced into the synthetic pathway. To shift the overall equilibrium further toward n-butanol, ccr and adhE2 were produced from a strong T7 promoter, while phaA, phaB, and crr were produced from a weak arabinose promoter.

There have been reports of the construction of CoA-dependent 1-butanol synthetic pathways in *E. coli* [32,33]. A novel 1-butanol synthetic pathway was designed with the modification of the (R)-1,3-butanediol pathway in order to supply more acetyl-CoA and NAD(P)H. The designed synthetic pathway consisted of heterologous enzymes from different species: butyraldehyde dehydrogenase (phaA, phaB, bld) from *Clostridium saccharoperbutylacetonicum*, inherent alcohol dehydrogenases (adhs) from the host *E. coli*, trans-enzyme-CoA reductase (ter) from *Treponema denticola*, and (R)-specific enoyl-CoA hydratase (phaJ) from *Aeromonas caviae*. Since inherent lactate dehydrogenase (ldhA) catalyzes the conversion of glucose to lactic acid under oxygen-depleted conditions, which disturbs 1-butanol production, the deletion of the ldhA gene eliminated the effect of the oxygen level on 1-butanol production and consequently, further increased the titer of 1-butanol.

In addition, biodiesel has gained great interest recently—specifically, fatty-acid-derived fuel molecules [34–36]—as they have lower toxicity to microbial hosts and higher energy density as compared with diesel fuel. Certain fatty acid-derived fuels, such as fatty acid methyl esters (FAMEs) and fatty acid ethyl esters (FAEEs), are usually formed via transesterification of fatty acyl-CoAs with
alcohols. This reaction is primarily catalyzed by a wax ester synthase. Fatty acid derivatives have been well-demonstrated and produced in micro-organisms, such as *S. cerevisiae* [34] and *E. coli* [37].

De Jong et al. constructed synthetic pathways in *S. cerevisiae* for enhanced production of FAEEs (fatty acid ethyl esters) [34]. The production of acyl-CoA requires acetyl-CoA and NADPH. To increase the supply of acetyl-CoA and NADPH, the inherent ethanol degradation pathway was up-regulated, and a synthetic phosphoketolase pathway was additionally introduced into the host cell. The former pathway directed the carbon flux toward the biosynthesis of acetyl-CoA, resulting in 408 ± 270 µg of FAEE gCDW⁻¹. This was achieved by overexpressing ADH2 (alcohol dehydrogenase 2) and ALD6 (acetaldehyde dehydrogenase), which respectively catalyze the conversion of ethanol to acetaldehyde and acetaldehyde to acetate, along with a heterologous gene, aceSE<sup>ΔacETP</sup>, from Salmonella enterica that encodes acetyl-CoA synthetase. The latter engineered pathway was the heterologous phosphoketolase pathway, which improved the production of both acetyl-CoA and NADPH. The engineered phosphoketolase pathway was constructed with two heterogeneous genes, *xpkA* (xylulose-5-phosphate phosphoketolase) and *ack* (acetate kinase), from *Aspergillus nidulans*. *xpkA* catalyzes the reaction where xylulose-5-phosphate, the precursor of the phosphoketolase pathway, is converted to acetyl-phosphate and glyceraldehyde-3-phosphate. Acetyl-phosphate is then converted into acetate as an intermediate by acetate kinase. Alternatively, acetyl-phosphate can be directly converted into acetyl-CoA by replacing *ack* with *pta* (phosphotransacetylase), which originates from Bacillus subtilis. The two phosphoketolase pathways were shown to result in increases in FAEE production to 5100 ± 509 and 4670 ± 379 µg FAEE gCDW⁻¹, respectively.

Recently, Sherkhanov et al. demonstrated a strategy to produce FAMEs in *E. coli* and reported a 35-fold improved FAME titer compared with that previously reported [37,38]. They constructed a synthetic metabolic pathway wherein FAMEs were produced through the direct methylation of fatty acids by a broader-range fatty acid methyl transferase [Drosophila melanogaster Juvenile Hormone Acid O-Methyltransferase]. This pathway was implemented in a β-oxidation-deficient and phospholipid synthesis-deficient *E. coli* strain that can tolerate high levels of medium-chain fatty acids expressed by acyl-ACP thioesterase (BTE) originating from *Umbellularia californica*. The resulting engineered strain produced the highest FAME titer ever observed in *E. coli*.

In addition to fueling biosynthesis, synthetic pathways have been built for the biosynthesis of various chemical compounds, such as succinic acid [39], 2,4-dihydroxybutyric acid [40], and maleate [41], and the engineered bio-factories have already been commercialized for the food industry.

However, in most cases, synthetic pathway construction requires laborious and tedious cloning work. Synthetic biology has developed several efficient multi-gene assembly strategies, such as Randomized BioBrick [30], which has been applied to construct an efficient lycopene biosynthesis pathway by random combinations of genetic components and has been shown to achieve a 30% increased titer in *E. coli*. In this method, various genetic parts, including promoters, RBSs, and terminators, are randomly assembled from PCR-amplified BioBricks simultaneously to identify the most efficient combination. The method successfully optimized a metabolic pathway that converts farnesyl diphosphate (FPP) to lycopene, which consists of three enzymes: geranylgeranyl diphosphate synthase (*crtE*), phytoene synthase (*crtB*), and phytoene desaturase (*crtI*). Other common methods are Golden Gate cloning [42], Golden Braid [43], MoClo [44], and MODAL [45]. However, such strategies have some disadvantages, such as the generation of scar sequences and the risk of PCR errors. Recently, Hochrein et al. introduced a user-friendly toolkit, AssemblX, to quickly create DNA structures with up to 25 functional units based on overlap-based cloning [46].

Furthermore, plasmid-based systems have drawbacks, such as the waste of cellular energy resources due to excessive metabolism [47,48] and plasmid maintenance. Recently, instead of using plasmids, chromosomal integration of genes has been favored for stable enzyme expression. Based on the clustered, regularly interspaced, short palindromic repeat (CRISPR) and CRISPR-associated systems (Cas), recently developed genomic editing methods have promoted the integration of metabolic pathways into the host genome. For example, a modified method of CRISPR-Cas, the delta integration CRISPR-Cas platform,
can be used to increase the homologous recombination efficiency by designing guided RNA sequences that specifically cleaves the target sequence at multiple delta sites in the genome of *S. cerevisiae* [49]. In this delta integration CRISPR-Cas platform, the plasmid expressing a Cas9 protein and a delta-targeting guide RNA is co-transformed with the 8-to-24-kb linear donor DNA fragments flanked between two homology arms of the delta sequences. The method has been used successfully to integrate a synthetic pathway composed of three enzymes in the (R,R)-2,3-butanediol production pathway: acetolactate synthase (alsS) and (R,R)-butanediol dehydrogenase (budA) from *B. subtilis*, and acetolactate decarboxylase (bdhA) from *Enterobacter aerogenes*. There have also been many other reports based on the use of CRISPR-Cas for chromosomal integration of heterologous pathways [50,51].

Novel synthetic pathways are widely used for metabolic engineering; thus, there are computational methods to design de novo metabolic pathways [52–54]. The in silico pathway design methods firstly enumerate possible metabolic pathways by connecting reactions obtained from various databases. As enzymes can often recognize molecules that are similar to their natural counterparts, potential and hypothetical reactions are also used during the pathway enumeration to extend the search space of metabolic pathways. Second, the enumerated pathways are pruned because the first step generates all combinations of pathways, which may include unnatural and unrealistic reactions and network structures. For example, generated pathways may include reactions that may not actually occur in nature, the lengths of generated pathways may be too long, or the pathways may form a cyclic structure wherein the target metabolite is used as an intermediate. Therefore, the enumerated pathways are evaluated to determine whether they comprise feasible enzymatic reactions and whether they can be tested and constructed in the laboratory. Third, the performance of the designed pathways is qualitatively evaluated in a given organism. In this step, the thermodynamics of enzyme reactions is considered, and several approaches have been developed, including the group contribution method and thermodynamics-based flux balance analysis. These methods finally suggest several feasible synthetic pathways that are predicted to perform the best in a given organism. The advances in synthetic metabolic pathway design methods facilitate the construction of metabolically engineered bio-factories that are capable of producing desired substances at high concentrations and can even produce non-natural chemical substances.

2.2. Enzyme Gene Expression Optimization

One of the challenges in synthetic metabolic pathway construction is to balance the kinetics of enzymes in the pathway. If a certain enzyme has lower expression or poor activity, it becomes a rate-limiting factor and induces abnormal accumulation of its substrates. Such intermediate metabolites are often toxic when accumulated, and thus, this accumulation may lead to growth retardation and eventually, to reduced production titer of the desired metabolite [48]. Furthermore, when a synthetic metabolic pathway is constructed, the metabolic flux within the pathway should be optimized by tuning the levels of enzyme gene expressions. There are an increasing number of studies working on understanding metabolic fluxes to address this obstacle.

Metabolic optimization through rational control over translation has been gathering great attention [55,56]. Enzyme genes should be expressed at a desired level to balance metabolic flux within a synthetic metabolic pathway to avoid the accumulation of intermediates and to maximize their overall performance. The gene expression process comprises two steps: transcription and translation. Since promoters are independent from downstream sequences, it is relatively easy to optimize the level of transcription. Conversely, translation is mediated by ribosome-binding sites (RBSs), and RBS sequences are structurally affected by downstream sequences, such as coding sequences. Since the structural effects on RBS cannot be predicted, so far, studies have focused on modeling translation processes and thereby, predicting accurate protein production levels.

Synthetic biologists have attempted to redesign RBS sequences for controlled enzyme expression [10,57]. Several synthetic tools have been developed that can predict the translation initiation rate and design RBS sequences to achieve balanced and robust gene expression. Existing tools
include the RBS Calculator [58], RBSDesigner [59], and other innovative models for the automated design of synthetic RBSs [60–62]. RedLibs [63] was recently developed; it is a tool that generates a library of RBS sequences for pathway flux optimization to reduce experimental trials and errors. Such models have been used in many studies to facilitate the construction and optimization of protein expression levels in synthetic pathways [64–66].

Similarly, randomized RBS sequences have been used to generate various expression levels of enzymes (idi, crtE, crtB, and ctrl) involved in lycopene biosynthetic pathway [67]. When 1080 colonies were examined out of a possible $3.8 \times 10^6$ combinations, the highest lycopene production found was an increase of up to 15.17 mg/g DCW.

2.3. Dynamic Metabolic Pathway Regulation

Another challenge in bio-factory construction is the optimization of the overall cellular metabolic fluxes toward the substance of interest. Certain metabolites often display toxicity when accumulated, and thereby, reduce cell growth or interfere with cellular functions. Introduced pathways may compete with cellular processes for cellular resources. These phenomena increase the burden on the host cell and eventually decrease the production titer [68]. Thus, cells use their own regulation and control of metabolic fluxes to balance the fitness of cellular machinery [69]. This suggests that the dynamic regulation of metabolic pathways becomes important for the efficient production of a desired substance by balancing metabolic fluxes without decreasing the cell growth rate.

Recently, dynamic regulatory systems have been developed for bio-factory construction [70,71]. The systems assist the engineered cells to adapt metabolic flux to fluctuations and changes within the host in real-time. The regulatory systems contain a sensor that detects the level of metabolite and an actuator that controls the metabolic flux (e.g., enzyme expressions or activities) [72]. For example, Zhang et al. developed a regulatory system that dynamically controls the expression of enzyme genes involved in bio-diesel production in response to acyl-CoA [73]. Overproduction of fatty acids would increase the production of bio-diesel but slow down the host cell growth. Bio-diesel is produced from fatty acyl-CoA that is produced from the natural fatty acid biosynthesis pathway. Thus, acyl-CoA is formed as an intermediate in the bio-diesel biosynthetic pathway. They used the FadR transcriptional regulator, which turns on a promoter in the absence of fatty acids/acyl-CoA and turns off in the presence of acyl-CoA. Thus, FadR was used to detect the intermediate concentration (fatty acids/acyl-CoA) and to dynamically control the expression of the enzyme genes involved in the diesel biosynthetic pathway. This regulatory system improved the stability of bio-diesel-producing strains and consequently, increased the bio-diesel titer. Liu et al. developed a similar strategy to over-produce fatty acids by regulating the expression of acetyl-CoA carboxylase in response to malonyl-CoA, since the overexpression of acetyl-CoA carboxylase induces cellular toxicity and slows down cell growth [74]. Several other sensors have also been used for bio-factory construction, such as aldehydes [75], butanol [76], and alkanes [77].

Although these models rely on a sensor for gene expression regulation, there are metabolites that do not have any appropriate sensor regulators. For instance, the heterologous mevalonate-based isoprenoid pathway produced two intermediates: 3-hydroxy-3-methylglutaryl-coenzyme A (HMG-CoA) and farnesyl pyrophosphate (FPP). These intermediates inhibit cell growth and affect fatty acid biosynthesis in S. cerevisiae and E. coli [78,79]. Thus, there have been reports recommending the use of synthetic promoters that are responsive to the two intermediates for the optimal regulation of the pathway. Dahl et al. developed stress response promoters to regulate the concentration of cellular HMG-CoA and FPP intermediates in the isoprenoid biosynthetic pathway [80].

Xu et al. constructed two synthetic promoters that are responsive to malonyl-CoA in opposite ways [81]. They introduced the FadR transcriptional regulator into E. coli, which binds to the fapO operator site when bound with malonyl-CoA. Interestingly, T7-fapO promoter was repressed by FapR, and malonyl-CoA relieved the repression. Conversely, pGAP-fapO promoter was activated by FapR, and this activation was removed by malonyl-CoA. Using these two promoters, they constructed
a pathway switch that is toggled by malonyl-CoA. They constructed a pathway (accADBC genes under the control of the pGAP-fapO promoter) to supply malonyl-CoA and another pathway (fabADGI and tesA' genes under the control of the T7-fapO promoter) to convert the malonyl-CoA into fatty acids, and the expressions of the pathways were toggled by the malonyl-CoA-responsive synthetic promoters. This strategy to balance cell growth and product formation increased fatty acid synthesis in *E. coli*.

### 2.4. Spatial Organization

A recent strategy for maximizing the performance of synthetic pathways is the spatial organization of pathway enzymes. The spatial organization of enzymes in close proximity creates channels where molecules can move quickly through enzymes without diffusion and accumulation. This channeling effect greatly increases the overall pathway efficiency. Protein, DNA, and RNA molecules are used as a spatial organization scaffold, where pathway enzymes are physically bound to the scaffold in a designable manner [82].

Several protein scaffolds have been developed thus far and used to co-localize enzymes. For instance, a protein scaffold containing GBD, SH3, and PDZ domains was constructed, and three mevalonate biosynthesis enzymes (AtoB, HMGS, and HMGR) were engineered to contain ligands to bind to the domains. When the enzymes and the scaffold were co-expressed, the enzymes were co-localized to the scaffold protein by ligand–domain interactions. Owing to the artificial channeling, the scaffold system achieved a 77-fold increase in mevalonate production yield as compared with the non-scaffold pathway [6]. In addition, other attempts to efficiently improve the production titer using scaffold proteins have also been demonstrated to produce glucaric acid (5.0-fold increase) [83], butyrate (3.0-fold increase) [84], and resveratrol (5.0-fold increase) [85].

In addition to protein scaffolds, DNA- and RNA-based scaffolds have also been developed to enhance pathway performance. For example, the three enzymes (homoserine dehydrogenase, homoserine kinase, and threonine synthase) involved in the threonine biosynthetic pathway were engineered to fuse with DNA-binding domains (ZFPs); thus, the enzymes were able to bind to particular DNA sequences. Corresponding DNA sequences (ZFP-binding sites) were introduced into a plasmid, and these ZFP-binding sites were used to provide a DNA scaffold to arrange the three enzymes in a particular order. The improved pathway performance resulted in a 50% reduction in the threonine production time [84]. Similarly, transcription activator-like effectors (TALEs) have been used to bind DNA scaffolds. TALE domains and their binding DNA scaffold were used to colocalize tryptophan-2-mono-oxygenase (IAAM) and indole-3-acetamide hydrolase (IAAH), and the scaffold increased the production of indole-3-acetic acid (IAA) in *E. coli* [86]. In addition, Sachdeva et al. identified eight RNA binding domains and their corresponding RNA motifs [87]. RNA sequences can be designed using the RNA motifs to provide a geometric scaffold for enzyme arrangement [88,89]. When RNA scaffolds were used for synthetic pathways in *E. coli*, the RNA scaffolds improved the production titers of succinate and pentadecane [87,88].

Furthermore, in 2013, José et al. designed an enzyme compartmentalization technique for high production of three advanced biofuels, i.e., isobutanol, isopentanol, and 2-methyl-1-butanol, in engineered *S. cerevisiae* [90]. The natural metabolic pathway for isobutanol production consists of an upstream mitochondrial pathway and a downstream cytoplasmic pathway. The upstream pathway contains enzymes associated with the valine biosynthetic pathway, such as acetolactate synthase, ketolacid reductoisomerase, and dehydroxyacid dehydratase. The downstream pathway is the valine Ehrlich degradation pathway that contains α-ketoacid decarboxylase and alcohol dehydrogenase. The compartmentalization of the designed plasmid with both upstream and downstream pathways into mitochondria was achieved by tagging the N-terminal mitochondrial localization signal from subunit IV of the yeast cytochrome c oxidase to the enzymes in the pathways, which resulted in a substantial surge of isobutanol by 260% as compared with only a 10% yield increase when overexpressing the same pathway in the cytoplasm.

Recent successful applications of synthetic biology approaches to metabolic engineering are listed and summarized in Table 1.
Table 1. Successes of synthetic biology methods for metabolic engineering.

<table>
<thead>
<tr>
<th>Methods</th>
<th>Description</th>
<th>Applications</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Pathway construction</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Synthetic engineered pathway</td>
<td>Rapid construction and optimization of multi-gene pathways for higher-titer production</td>
<td>- n-Butanol&lt;br&gt;- FAEEs, FAMEs&lt;br&gt;- Succinic acid, 2,4-dihydroxybutyric acid, maleate</td>
<td>[31–34,37,39–41,52–54]</td>
</tr>
<tr>
<td>In silico pathway design</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Randomized BioBrick AssembIX</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Chromosomal integration</td>
<td>Efficient and stable enzyme expression and higher production yield</td>
<td>- (R,R)-2,3-butanediol&lt;br&gt;- Mevalonate and bisabolene titers increased by 41-fold and 5-fold, respectively</td>
<td>[49–51]</td>
</tr>
<tr>
<td><strong>Enzyme expression optimization</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RBS Calculator</td>
<td>Control over protein expression levels through ribosome-binding sites</td>
<td>- Optimization of isoprenoid production&lt;br&gt;- L-tyrosine production&lt;br&gt;- N-acetylneuraminate biosynthesis&lt;br&gt;- Lycopene synthesis</td>
<td>[64–67]</td>
</tr>
<tr>
<td>RBSDesigner</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RedLibs</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Pathway regulation</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dynamic regulation system</td>
<td>Dynamic regulation and balancing of metabolic systems as well as redirection of fluxes to achieve high production of desired protein</td>
<td>- A fatty acyl-CoA biosensor FadR&lt;br&gt;- Fatty acids, aldehydes, butanol, and alkanes&lt;br&gt;- Isoprenoid biosynthetic</td>
<td>[73–77,80,81]</td>
</tr>
<tr>
<td><strong>Spatial organization</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Protein scaffolding</td>
<td>Concentrating intermediates and rapidly directing them through the metabolic pathway</td>
<td>- 77-fold increase in mevalonate production&lt;br&gt;- 5-fold increase in glucaric acid synthesis&lt;br&gt;- 3-fold increase in butyrate production&lt;br&gt;- 5-fold increase in resveratrol synthesis</td>
<td>[6,83–85]</td>
</tr>
<tr>
<td>DNA and RNA scaffolds</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>- 50% reduction in production time of L-threonine&lt;br&gt;- Production of IAA&lt;br&gt;- Succinate&lt;br&gt;- Pentadecane</td>
<td>[84,86–88]</td>
</tr>
</tbody>
</table>
### Table 1. Cont.

<table>
<thead>
<tr>
<th>Methods</th>
<th>Description</th>
<th>Applications</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>Synthetic components</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>for expression regulation</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Promoter tuning</td>
<td>Optimal regulation of the gene expression strength in a dynamic range up to hundreds-of-fold activation</td>
<td>- Production of PepN protein at about 10–15% of the total cellular protein</td>
<td>[91]</td>
</tr>
<tr>
<td>Promoter regulation</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Synthetic transcription factors</td>
<td>CRISPR-based regulation</td>
<td>- Production of 2.0 g/L β-carotene (fed-batch fermentation)</td>
<td>[92,93]</td>
</tr>
<tr>
<td></td>
<td>Regulation of gene expression using a nuclease-deficient Cas9 (dCas9)</td>
<td>- Production control of P (3HB-co-4HB) in <em>E. coli</em></td>
<td></td>
</tr>
<tr>
<td>Engineering transcription factors</td>
<td>Engineering AraC protein through altering effector specificity using a method of saturation mutagenesis</td>
<td>- 20-fold increased production of triacetic acid lactone in <em>E. coli</em></td>
<td>[94]</td>
</tr>
<tr>
<td>Chimeric transcription factor</td>
<td>Fusion of the target metabolite recognition domain and the promoter regulatory domain</td>
<td>- Tyrosine</td>
<td>[95]</td>
</tr>
<tr>
<td>Synthetic RNAs</td>
<td>Regulation of gene expression based on physicochemical models of RNA</td>
<td>- Development of high-yield VB12 production strains</td>
<td>[96,97]</td>
</tr>
</tbody>
</table>
3. Synthetic Components for Gene Expression

From DNA to protein: The regulation of the transcription and translation processes is a fundamental theme for the development of practical tools for metabolic engineering. The engineering of microbial cell factories requires the introduction of heterologous enzyme genes and/or the modification of endogenous enzyme genes. For optimal performance, the expression of enzyme genes should be finely regulated [98]. Gene expression can be regulated at transcription and/or translation. Thus, the various techniques introduced in this section represent the most recent and influential synthetic tools for controlling the transcription and translation processes that can be applied to the metabolic engineering of various micro-organisms. There have been attempts to design synthetic promoters and their regulators using the TAL effector and CRISPR for transcriptional regulation. For translational regulation, riboswitches and small regulatory RNAs (sRNAs) have been developed and used for metabolic engineering.

3.1. Transcription Regulation

3.1.1. Promoters for Tuning

Various synthetic promoters have been developed to optimally tune the enzyme expression levels for the efficient production of a desired substance. Conventional inducible promoters have been frequently used to acquire different levels of gene expression by controlling the concentrations of inducers; however, there are only few promoters available, and their strengths are not diverse. Thus, to fine-tune enzyme expression for metabolic flux optimization, synthetic promoter libraries have been developed to identify the best combination of promoters for enzyme genes in a synthetic metabolic pathway. As mentioned in Section 2.2 of this paper, randomly generated libraries of synthetic promoters have also been used to fine-tune the metabolic flux within a synthetic pathway.

A wide range of synthetic promoter activities was created by the randomization of spacer sequences between −35 to −10 consensus sequences [99,100]. Such randomized libraries have been used for the fine-tuning of enzyme expression and accordingly, to achieve high-level production of compounds. For example, Rud et al. developed a synthetic promoter for the gram-positive bacterium *Lactobacillus plantarum* by randomizing the non-consensus sequence of the rRNA promoter. This randomization generated 33 different constitutive promoter sequences, and this promoter library showed various relative expression levels ranging from 1 to 160. Using the library, they identified an optimal promoter that showed stable expression of PepN and GusA [91]. In another study, a library of synthetic promoters was randomly generated in actinomycetes. The non-consensus nucleotides of the ermEp1 promoter were randomized, and accordingly, 56 synthetic promoters were generated. The library of synthetic promoters showed 2% to 319% activity compared with the native ermEp1 promoter, and their strengths were confirmed by RNA-seq analysis. As a proof of concept, the strongest promoter was used to express the *rppA* gene encoding for a type III polyketide synthase, which converts five malonyl-CoA compounds to 1,3,6,8-tetrahydroxynaphthalene, which is spontaneously oxidized to flavolin. Consequently, a 3.3-fold increase in the production of flavolin was achieved with the synthetic promoter [101]. Synthetic promoter libraries have also been commonly used to fine-tune enzyme expressions and thus obtain higher titers of substances, including 2,3-butanediol in *E. coli*, actinorhodin in *Streptomyces coelicolor A3(2)* [102], and endoxylanase in *Corynebacterium glutamicum* [103].

Furthermore, there have also been attempts to introduce synthetic promoters into the genome. Braatsch et al. used Red/ET recombination to replace a genomic native promoter with synthetic promoters. Synthetic promoter sequences were PCR-amplified with a kanamycin resistance gene flanked with Flp recognition targets. These amplified sequences were then replaced with the genomic sequence by homologous recombination. After appropriate selection, the kanamycin resistance gene was excised by Flp recombinase [104]. They replaced the native pgi promoter with synthetic promoters exhibiting a wide spectrum of strengths ranging from 25% to 570% of the native pgi promoter.
Thus far, various computational methods have been developed to decipher the effects of DNA sequences on promoter strength and to design de novo design synthetic promoter sequences. Position weight matrix models were used to predict the strength of *E. coli* core promoter sequences recognized by the sigma factor $\sigma^E$ [105]. The models were then improved by incorporating the effect of upstream elements into promoter strength [106]. However, since position weight matrix models are highly reliant on experimental data obtained from promoter-strength studies, they are only applicable to well-studied micro-organisms, such as *E. coli*. Other models have also been developed to predict promoter strengths based on the partial least squares [107] and artificial neural network [108] methods among other methods. Computation models have been used for the de novo design of synthetic promoters to fine-tune the enzyme genes involved in the deoxxyxylulose phosphate pathway in *E. coli* [108].

3.1.2. Insulated Promoters

The expression of enzymes downstream of regulatory promoter could be interfered with by cellular factors; thus, it is often required for synthetic systems to be insulated from cellular systems for reliable operation. The recently developed SELEX-based screening of bacterial genomes identified the complex regulatory targets of the characterized 116 transcription factors via 156 transcription profiles [109]. Because the interactions have not yet been completely elucidated, when constructing a synthetic system, DNA sequences that may interact with cellular factors could be included. Because of this unexpected inclusion of interacting DNA sequences, synthetic systems may behave unexpectedly and lose their robustness. To resolve this concern, Davis et al. devised a simple strategy to insulate a promoter by flanking the promoter sequence with insulating sequences [110] because upstream or downstream sequences of a promoter may contain regulatory sequences that can affect the promoter. Briefly, they used the *E. coli* rrnB P1 promoter as the core and added insulating sequences covering $-105$ to $+55$. Randomization of the space nucleotides of the rrnB P1 promoter generated a library of promoters showing different levels of strength. When evaluated with the GFP gene under the control of the insulated promoter, the promoter’s strength was not affected by upstream or downstream sequences due to the long flanking sequences. Zong et al. also improved the insulation-based engineering strategy by identifying insulated promoter cores and operators and predicting the consequence of their combinations via a biophysical model of synthetic transcription [111]. They demonstrated a modular system that correctly programs synthesis circuits by designing 83 combined promoters that randomly combine 53 different promoter cores and 36 synthesis operators to encode the NOT gate function in *E. coli* DH10B.

3.1.3. Synthetic Transcription Factors

Endogenous enzyme genes often need to be regulated for optimized metabolic flux regulation. However, there are only a limited number of regulator proteins available and they require specific recognition sequences. Thus, conventional regulators cannot be used for the regulation of endogenous genes that have very diverse promoter sequences. There have been many attempts to develop customizable synthetic transcription factors [92,93,112–115]. For example, TALE proteins can be reprogrammed to activate or inhibit transcription by binding at a particular sequence. More recently, clustered, regularly interspaced, short palindromic repeat (CRISPR)-Cas9 systems have been used to regulate gene expression more efficiently than conventional methods. Therefore, these synthetic transcription factors could be the key to synthetic biology and metabolic engineering.

CRISPR-Based Regulation

Recently, there has been growing evidence that the CRISPR/Cas9-based genome editing approach enables efficient, intricately controlled gene expression and versatile cross-species genome editing in various micro-organisms [5,113,114]. Briefly, the action mechanism of the CRISPR/Cas9 system is that the complex of small crRNAs produced by the coordinated action of a small, trans-activating crRNA (tracrRNA), the Cas9 nuclease, and the host RNaseIII directs Cas9 to cleave the target DNA,
which is followed by an adjacent protospacer-adjacent motif (PAM). The target DNA is recognized by its alignment with the 12–15 nucleotides at the 3′ end of the crRNA guide sequence as well as the NGG sequence of PAM. Unlike the effort to manipulate other transcription factors, the CRISPR/Cas9 system can be simply applied to genome editing and metabolic pathway engineering by only altering the target sequence without any protein engineering. For example, Jiang et al. reported the introduction of point mutations and codon changes in the E. coli genome using the CRISPR/Cas9 method [5]. For instance, the β-carotene biosynthetic pathway, the methylerythritol-phosphate pathway, and the central metabolic pathways in E. coli were systematically optimized for β-carotene production by iterative editing of the 2-day editing cycle, and the best production was 2.0 g/L β-carotene (fed-batch fermentation) [92].

With respect to the regulation of gene expression, CRISPRi-based interference (CRISPRi) has also been applied to metabolic engineering using a nuclease-deficient Cas9 (dCas9) that contains two point mutations, one each in the RuvC (D10A) and HNH (H840A) domains [4,113]. Due to the gRNA binding ability of dCas9, it has been used as a programmable transcription repressor by circumventing the binding of RNA polymerase to the promoter sequence or as a transcription terminator by binding to the target gene. Recently, CRISPRi was applied to target the P\text{P7}/LacO\text{1} promoter for preventing the leaky expression of toxic proteins [115]. Conversely, the fusion of RNA polymerase omega subunit (ω) and dCas9 achieved programmable transcriptional activation by enhancing the binding stability of RNA polymerase to the upstream promoter region. Therefore, these RNA-guided transcriptional reprogramming approaches have a high potential for fine-tuning the production of target genes in metabolic engineering. According to Lv et al., the repression of multiple genes in polyhydroxyalkanoate (PHA) biosynthesis using CRISPRi enables the production control of P (3HB-co-4HB) in E. coli [93].

### TALE-Based Regulation

As a new genetic switch tool, TALE proteins, which can bind to specific DNA target sequences, have been engineered to produce TALE Dimers (TALEDs) by replacing the nuclease domain of a TALE with a well-studied FKBP dimerization domain [112]. The engineered TALEDs contain a DNA-binding domain and dimerization domain; therefore, they can bend the DNA structure by binding to two TALED-binding sites and by subsequent dimerization. Based on the DNA looping paradigm, engineered TALEDs have been implemented to switch the transcription initiation of the lac operon system of E. coli.

### Engineering/Chimeric Transcription Factors

In addition to CRISPR- and TALE-based regulation, there have been other attempts to develop customizable synthetic transcription factors by engineering native transcription factors. For example, by altering effector specificity using a method of inducing saturation mutagenesis, Tang et al. engineered AraC to sense triacetic acid lactone (TAL) and used the engineered AraC in the directed evolution of Gerbera hybrida 2-pyrene synthase to catalyze the synthesis of TAL [94]. As a result, they were able to identify an improved G. hybrida 2-pyrene synthase variant and by using the synthase, they achieved 20-fold increased production of TAL. Frei et al. also developed engineered AraC variants that respond to new inducer compounds, vanillin and salicylic acid, for molecular sensing and reporting [116]. These designed transcription factors may expand the sensing spectrum of native transcription factors to regulate and optimize synthetic pathways. In addition, Chou et al. introduced a new adaptive control system called feedback-regulated evolution of phenotype, in which two modules dynamically control the mutation rate (an actuator module) based on the concentration of a target molecule (a sensor module) [95]. For this system, a chimeric transcription factor has been developed that not only can bind to the target metabolite but also regulate the promoter for tyrosine and isoprenoid production in E. coli.
3.2. Expression Regulation

The diversity, specificity, and kinetics of regulatory RNAs contribute to the regulation of diverse physiological responses, including transcription, translation, and mRNA stability in living cells [96, 97, 117–122]. The structures, functions, and mechanisms of regulatory RNAs have been elucidated in many bacterial species for decades. The well-known bacterial regulatory RNAs called riboswitches, which are located at the 5' untranslated regions of mRNAs, comprise a ligand-binding aptamer region and the expression platform. Upon binding of a metabolite ligand, the riboswitches regulate the expression of a downstream coding sequence through the formation of terminator/anti-terminator structures or masking/unmasking Shine–Dalgarno (SD) sequences in the mRNA [120, 121]. Riboswitches repress gene expression as cis-acting RNA regulators. Engineered riboswitches have been used as a genetic toolkit to obtain desired gene production or toxic product regulation in bacterial species. For example, the application of engineered VB12 riboswitches to develop high-yield VB12 production strains using a flow cytometry high-throughput screening system in Salmonella typhimurium has been reported [96]. Moreover, six different theophylline-responsive riboswitches have been used for intricate transcriptional regulation of the toxic protein SacB in several cyanobacterial species; these riboswitches have shown more efficient performances than isopropyl-D-thiogalactopyranoside (IPTG) induction of the lacIq-Ptrc promoter system [97].

Synthetic regulatory RNAs have also been specifically developed to regulate gene expression during the post-transcription process [117–119]. Isaacs et al. developed two types of artificial riboregulator: cis-repressed mRNA (crRNA), which interferes with ribosome binding through intramolecular interaction with SD, and trans-activating RNA (taRNA), which activates gene expression via an intermolecular interaction that releases the SD sequence [118]. In addition, there are trans-acting sRNAs that regulate the translation and stability of target mRNA through base pair complementation. Like natural sRNAs, synthetic sRNAs mainly comprise two parts: a scaffold sequence and a target-binding sequence [119]. The scaffold sequence is able to recruit the Hfq protein, which not only promotes effective hybridization of the synthetic sRNA and its target mRNA but also facilitates the degradation of the target mRNA. The target binding sequence that is complementary to the translation initiation region of the target mRNA competes with ribosomes to achieve efficient translation inhibition. Using this mechanism, synthetic sRNAs have been applied to metabolic engineering for efficient gene knockdown for tyrosine and phenol production in E. coli [119, 123] and butanol production in C. acetobutylicum [122]. Recently, orchestrated regulation during both transcription and translation has been reported to achieve a high muconic acid titer of up to 1.8 g/L [124].

4. Summary

Metabolic engineering for the microbial production of chemicals and pharmaceuticals has been accelerated with the help of various genetic tools in synthetic biology. In this review, we introduced key technologies that are used in synthetic biology for metabolic pathway engineering, such as the design, construction, and optimization of metabolic processes as well as the engineering of synthetic components, which could make cellular processes predictable and robust to achieve desirable outcomes. Various strategies and tools that are used in synthetic biology to develop promising pathways and modules enable the establishment of intellectual foundations for cell factory construction in metabolic engineering.

One of the future goals of synthetic biology that is applicable to metabolic engineering is the artificial design of a novel micro-organism that possesses the capability to produce desired metabolites most efficiently and in large quantities like a factory. However, the prediction of whole cell physiologies from DNA sequences and thereby, the design of novel organisms from scratch is a great challenge. Even simple protein expression processes are very complex because of interruptions by other molecules. In addition, to our knowledge there are still no methods to accurately predict the kinetics of proteins (e.g., enzymes, interactions among proteins, DNAs, metabolites, etc.) solely from DNA sequences. Thus, synthetic biology is still at the early stages of development. To fulfil the goal, more knowledge
on cellular machineries should be accumulated to allow the prediction of cell behaviors from DNA nucleotides, and new technologies should also be developed to produce designed micro-organisms.

**Author Contributions:** Conceptualization, H.-M.L. and P.N.L.V.; Writing-Original Draft Preparation, H.-M.L. and P.N.L.V.; Writing-Review & Editing, D.N.; Supervision, D.N.

**Funding:** This work was supported by C1 Gas Refinery Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Science, ICT and Future Planning (NRF-2016M3D3A1A01913244) and was also supported by the Chung-Ang University Research Grants in 2017.

**Conflicts of Interest:** The authors declare no conflict of interest.

**References**


47. Keasling, J.D. Synthetic biology for synthetic chemistry. ACS Chem. Biol. 2008, 3, 64–76. [CrossRef]


77. Reed, B.; Blazeck, J.; Alper, H. Evolution of an alkanе-inducible biosensor for increased responsiveness to short-chain alkanes. *J. Biotechnol.* 2012, 158, 75–79. [CrossRef] [PubMed]


83. Moon, T.S.; Dueber, J.E.; Shie, E.; Prather, K.L. Use of modular, synthetic scaffolds for improved production of glucaric acid in engineered *E. coli*. *Metab. Eng.* 2010, 12, 298–305. [CrossRef]


106. Rhodius, V.A.; Mutalik, V.K.; Gross, C.A. Predicting the strength of UP-elements and full-length *E. coli* sigmaE promoters. *Nucleic Acids Res.* **2012**, *40*, 2907–2924. [CrossRef]


© 2018 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (http://creativecommons.org/licenses/by/4.0/).