

Overview of Regulatory Strategies and Molecular Elements in Metabolic Engineering of Bacteria

Tianwen Wang · Xingyuan Ma · Guocheng Du · Jian Chen

© Springer Science+Business Media, LLC 2012

Abstract From a viewpoint of biotechnology, metabolic engineering mainly aims to change the natural status of a pathway in a microorganism towards the overproduction of certain bioproducts. The biochemical nature of a pathway implies us that changed pathway is often the collective results of altered behavior of the metabolic enzymes encoded by corresponding genes. By finely modulating the expression of these genes or the properties of the enzyme, we can gain efficient control on the pathway. In this article, we reviewed the typical methods that have been applied to regulate the expression of genes in metabolic engineering. These methods are grouped according to the operation targets in a typical gene. The transcription of a gene is controlled by an indispensable promoter. By utilizing promoters with different strengths, expected levels of expression can be easily achieved, and screening a promoter library may find suitable mutant promoters that can provide tunable expression of a gene. Auto-responsive promoter (quorum sensing (QS)-based or oxygen-inducible) simplifies the induction process by driving the expression of a gene in an automated manner. Light responsive promoter enables reversible and noninvasive control on gene activity, providing a promising method in controlling gene expression with time and space resolution in metabolic engineering involving complicated genetic

circuits. Through directed evolution and/or rational design, the encoding sequences of a gene can be altered, leading to the possibly most profound changes in properties of a metabolic enzyme. Introducing an engineered riboswitch in mRNA can make it a regulatory molecule at the same time; ribosomal binding site is commonly engineered to be more attractive for a ribosome through design. Terminator of a gene will affect the stability of an mRNA, and intergenic region will influence the expression of many related genes. Improving the performance of these elements are generally the main activities in metabolic engineering.

Keywords Metabolic pathway · Metabolic engineering · Riboswitch · Intergenic region · Rational design · Directed evolution

Introduction

In the common sense of living organisms, there are sets of chemical reactions, which are usually catalyzed by various enzymes. In most cases, these reactions are sequentially organized into metabolic pathways. A metabolic pathway responsible for the degradation of a nutrient compound or the biosynthesis of a given biomolecule, is a functionally and structurally complete cellular unit. Instead of being isolated, they interact each other. The collection of the interweaving pathways in a cell makes a metabolic network. The maintenance of life requires a finely concerted balance among these pathways in an integrated metabolic network.

A metabolic network is not static, but dynamic and changing. At a long time scale, the metabolic network evolves to be more and more optimized with the natural selection as the permanent pressure; in a short time range,

T. Wang · G. Du · J. Chen
State Key Laboratory of Food Science and Technology, Jiangnan University, Wuxi 214122, Jiangsu, China

X. Ma (✉)
School of Biotechnology, State Key Laboratory of Bioreactor Engineering, East China University of Science and Technology, Shanghai 214122, China
e-mail: maxy@ecust.edu.cn

the alteration in metabolic network can enable a cell to efficiently and promptly response to the environmental change(s). For any living organism itself, an optimized and ever-changing metabolic network can avoid any unnecessary biosynthesis of biomacromolecules from limited resources or enzymes metabolizing currently unavailable nutrients. As a result, under the natural conditions, in a bacterial cell of a wild-type strain, any cellular substance is synthesized at a reasonably appropriate amount that can support survival of the cell.

Exploring microbial cells as the workhorse for the human being, we must modify the wild-type metabolic pathway of a cell: maximize the flux towards our desired products, and minimize the flux in other metabolic pathways that might lead to loss of precursor metabolites or energy (Fig. 1). These operations are based on the knowledge on the pathway of interest. Advances in biochemistry, molecular biology, and molecular genetics have provided us a broader horizon of metabolic pathways, and their complicated regulation, offering us not only more choices in tool and methodology in metabolic engineering researches, but also the components of a given metabolic pathway that can be used as candidate targets to be engineered. There are many good reviews on discussing the tools and methodology in metabolic engineering [1–3]. In this review, we will provide a relatively comprehensive introduction of the components in a pathway and the main strategies, with which these components can be modified from the viewpoint of metabolic engineering.

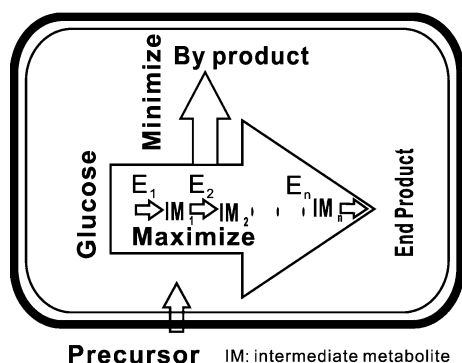


Fig. 1 The principle of metabolic engineering is to maximize the flux towards the formation of the end product(s), and minimize the flux towards by product(s). Metabolizing glucose, the most common and important material that supports cell growth and provides the building blocks for end products, into different intermediate metabolites (IM_1, IM_2, \dots, IM_n) through variable numbers of steps predominantly catalyzed by individual enzyme (E_1, E_2, \dots, E_n) in a cell, is the domain of operations in metabolic engineering. Enhancing the efficiency of end product formation by engineering the steps (including the addition of certain precursor) are the key activities in metabolic engineering

The Targets for Metabolic Engineering

Metabolic pathway is a sequence of biochemical reactions. Enzymes encoded by different genes are the catalysts of these reactions. Therefore, protein-encoding genes are reasonably the key players of a metabolic pathway. In a broad sense, a protein-encoding gene includes the encoding sequence (structural gene) and all regulatory elements. The encoding sequence of a gene primarily determines the biochemical and biophysical properties of the polypeptide. Regulatory elements discussed here refer to other sequences defined by the gene except for the encoding region, including promoter, ribosome binding site sequence, terminator, etc. (Fig. 2).

The Catalyst Itself: Genetic Modification of the Encoding Sequences

In general, a gene functions when it is transcribed into RNA, and the protein-encoding part is transcribed into mRNA. Despite the new discovery, that mRNA can function as a regulator through competitively binding to microRNA in tumor biology [4], the canonical role of mRNA is to deliver the information to protein synthesis by a ribosome. An optimized metabolic pathway frequently requires more robust enzymes (higher catalytic efficiency and insensitivity to feedback inhibition of the final product) catalyzing given steps. As a commonsense in biochemistry, the property of a metabolic enzyme is determined by its

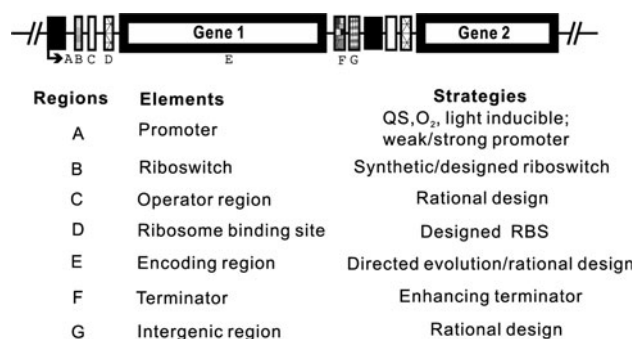


Fig. 2 Elements in a typical protein-encoding gene are the possible targets for metabolic engineering. For a typical protein-encoding gene, it encompasses a promoter (A), an operator region for expression control (C), a ribosome binding site (D) to attract ribosome and to initiate the translation, an encoding region (E) which will be transcribed into the template for translation, and a terminator (F) dictating the stop of transcription of a region. Different genes are arranged to be relatively independent by intergenic regions (G). Possible existence of riboswitch (B) in mRNA functions as a sensor to probe the abundance of the product of a gene or a pathway, directly regulating the translation of the mRNA. When the systematic regulatory network is insufficiently understood, enhancing the efficiency of steps in a pathway by modifying the elements defined by a gene is the most straightforward and realistic practices

structure and ultimately by its amino acid sequences. Currently, there is no available method to change the amino acid sequence of a protein once it is produced through translation by a ribosome. Therefore, any expectable alteration in the protein sequences can only be realized by changing the encoding sequence of a protein, and thus the function of the polypeptide (Not all change in the encoding sequence will produce a mutant protein due to the degeneracy in genetic codon).

In general, there are two different strategies to create improved enzymes that can be adopted in the optimization of a pathway in metabolic engineering studies: directed evolution [5, 6] and rational design [7–9]. In most cases, directed evolution and rational design are believed in the scope of protein engineering, which has been utilized to modify natural proteins (including enzymes) to live up to the demands of different industrial applications. Similarly, protein engineering can also play an imminent role in improvement of enzyme activity, alteration of substrate and product specificity, and modifying regulatory elements [10].

Directed Evolution

Directed evolution (or directed molecular evolution) is artificially designed operations essentially imitating the process taking place in the nature, with which an accelerated speed of evolution can be obtained [11]. It is of particular importance for preparation of industrial enzymes or therapeutic proteins with improved desirable properties in a laboratory. By employing sophisticatedly designed protocols, library comprising diverse mutant sequences from one or more starting sequences can be created. The supposal by Smith provided a theoretically possible route for evolution “functional proteins must form a continuous network which can be traversed by unit mutational steps without passing through nonfunctional intermediates” [12, 13], but in directed molecular evolution, exploring the wider encoding sequence space will inevitably produce a large percentage of inactive mutants. Therefore, the library created in directed molecular evolution must be subjected to activity-based screening. The mutant sequences that encode proteins with improved properties can be used in the metabolic engineering to replace the wild-type one. In the overproduction of levopimaradiene, Leonard et al. [14] metabolically engineered an *Escherichia coli* by introducing the mutant versions of two important metabolic enzymes functioning in the synthesis pathway of levopimaradiene: geranylgeranyl diphosphate synthase from *Taxus Canadensis* and levopimaradiene synthase from *Ginkgo biloba*. The activity of levopimaradiene synthase was improved by site-directed and -saturated mutagenesis with the results from some bioinformatics analysis as

guidance. Another rate-limiting enzyme geranylgeranyl diphosphate synthase was enhanced with error-prone polymerase chain reaction (PCR). As a result, a significantly enhanced overproducer was obtained (2,600 times higher).

Rational Design

Different from directed molecular evolution, rational design of protein largely depends on the availability of structure–function relationship of the protein of interest. With the information, we can create any desirable mutant sequence by PCR-based site-directed mutagenesis, or chemical synthesis of the whole gene. When chemical synthesis of a gene is adopted, decision on what codon should be used for a given amino acid residue of the target protein must be made according to the preference of codon usage of specific organisms. Although this kind of optimization is more frequently done when efficient expression of a gene for better protein yield is the most desirable result, and a higher-level expression of the metabolic enzymes is not the main objective of metabolic engineering (in some cases, higher-level expression is unexpectedly deleterious), optimization according to species-specific codon usage is also profitable. Moreover, it is especially important when one attempts to construct a heterogeneous pathway from other higher species into a lower model organism [15], and when the pathway to be constructed contains genes of different origins [16].

Rational design of a protein can also be simply base on the sequence analysis of its homologous sequences. It is an effective strategy to improve some properties of a protein which usually requires the spontaneous mutations at multiple sites (e.g., solubility and thermostability) [17]. When unsatisfactory activity of an overexpressed metabolic enzyme is the consequence of poor expression level or low percentage of soluble fraction because of its heterogeneity to many widely used hosts, mutating the sites identified by multiple alignments amongst homologous sequences with the “back-to-consensus” strategy might be the first choice. By taking this route, the γ -humulene synthase from *Abies grandis* was engineered by redistributing the glycine and proline residues. A comprehensive mutant containing these beneficial mutations exhibited enhanced solubility, leading to an increased yield by 80-fold [18].

Codon usage influences the translation efficiency because translation of different codons on mRNA generally needs corresponding transfer RNAs (tRNA). The absence of tRNA that can recognize some “rare codon” will lead to lowered translation efficacy or total failure of translation. Meanwhile, the codon usage can also control ribosome speed, exerting a fine-tuning translation control on the efficiency of protein synthesis [19–21]. In translating an

mRNA, ribosome moves towards the 3'-end at a varying speed: in the first 50 codons, ribosome proceeds at a low speed (ramp stage), then at a high speed. These new findings are consistent with other discoveries that the sequence at the beginning of the gene strongly influenced translation, and that the expression level was inversely correlated with predicted mRNA secondary structure [22, 23]. The beginning region of a gene controls the initiation and/or early elongation, depending on the particular gene. At the same time, in the sequence of a gene, the tendency to using the same codon for its second appearance as the first one (a phenomenon termed as "autocorrelation") can be beneficial for translation in some way. These understandings in the mechanism of translation may provide an effective way to speed up translation, although there has not been a successful example to apply these findings to optimization of pathway enzymes in metabolic engineering until now.

The overproduction of L-tyrosine by *E. coli* involves many steps catalyzed by different enzymes. In Juminaga's effort for constructing an overproducer of L-tyrosine, they identified the bottlenecks limiting the yield through targeted proteomics and metabolite profiling. One of the bottlenecks is the production of intermediate shikimate catalyzed by dehydroquinate synthase (AroB). This enzyme expressed constitutively at a low level. Optimization of the first 15 codons of the gene alleviated the constraint exerted by this enzyme due to its poor expression [24]. This example supported the theory proposed by Cannarozzi et al. [19]. Probably, it also implies that the codon optimization of a gene for improved expression level should be concentrated on the 5'-terminal of the encoding gene (about 50 amino acid residues).

The Regulatory Components

The regulatory components control when and to what extent of a gene is expressed. Strategic utilization of regulatory components can realize the fine tuning in gene expression with minimal intervention to the host's metabolism.

Engineering Ribosome Binding Site

A necessary step for protein synthesis by a ribosome is its binding to a sequence of three to nine nucleotide region upstream of the starting codon, the ribosome binding site (RBS), on an mRNA [25, 26]. In bacteria, this RBS interacts with a sequence near the 3' end of the 16S rRNA through complementary pairing. The number of base-pairs involved in the interaction and the location of these base-pairs in relation to that of the starting codon dictated the translation efficiency [27]. Further studies formed the idea that the enhanced stability of the RBS region through the

formation of secondary structure can negatively affect the efficiency of translation. By quantitative analysis on how the secondary structure of the RBS determined translational efficiency, researchers found that exposure of only the RBS region or the starting codon was insufficient for preferential recognition, and the translation efficiency was strictly correlated with the fraction of mRNA molecules with an unfolded RBS. The reason was that a ribosome could not recognize nucleotides outside the RBS region and the initiation codon, and structured RBS region [28].

Biologists have developed methods to control protein expression through automated design of synthetic RBSs. With the model based on free energy changes in relevant biological events in the process of translation initiation, Salis et al. [29] developed a predictive method to design synthetic RBSs for rational control over the protein expression level. According to their experimental validation of the model with >100 predictions in *E. coli*, they proposed that the accurate method was very useful to connect a genetic sensor to a synthetic circuit by rationally optimizing protein expression. Different from Salis's methodology, Na et al. [30] paid their attentions to the dynamics of mRNA folding and ribosome binding to estimate translational efficiency with mRNA sequence as the sole input. This model was successful with the manifestation of a high correlation factor ($R^2 = 0.87$) observed between the experimental data from expression and estimated ones when *luxR* mRNA derivatives were taken as an example.

Although the workability of the two above mentioned models were proven by the expression data of some proteins, successful implementations of these models to other designs towards expression control were scarcely reported. Library-based approach involved the generation of libraries of RBS sequences and screening for more robust sequences. Wang et al. [31] invented the method of multiplex automated genome engineering (MAGE). They applied their invention to the engineering of a recombinant *E. coli* strain (EchW2) for lycopene production. Targeting the 20 genes responsible for lycopene synthesis, the RBS regions were modified through allelic replacements using oligos containing degenerate RBS sequences (DDRRRRRDDDD; D = A, G, T; R = A, G). The higher similarity between the replaced RBS region and the canonical Shine-Dalgarno sequence (TAAGGAGGT) gave rise to enhanced translation efficiency.

Promoter

Control on Expression Level by Adopting Strong or Weak Inducible Promoters Promoter is a DNA fragment that can be recognized by a group of proteins responsible for initiation of transcription of a gene. Very often, more than

one related gene (an operon or gene cluster) shares the same promoter. In this case, the expressions of these genes are under the control of the same promoter. Transcription is the first step of gene expression. Therefore, promoter can exert an efficient regulation on the expression level of a gene. At the same time, promoter provides the most economical control on gene activity: if the expression of a gene is unnecessary, the synthesis of its mRNA should also be avoided (despite the recent discovery that mRNA can also play a role in regulating gene expression, besides its basic function as the template for protein synthesis [4]). As a result, it has been a routine practice to obtain different expression levels of the target gene by adopting different promoters [32, 33], especially in metabolic engineering [34, 35]. Some frequently utilized promoters are *lac*, *tac*, *ara*, and so on. These promoters are strong promoter and are inducible.

Beyond these well-known promoters with different strengths, achievement of different expression levels of a gene frequently involves the screening a promoter library [36–39] if the promoter is selected as the control point. A further improvement of the library method should be the ability to select the ideal mutant promoter according to strength comparison between the promoters in the library and the native one with the promoter strength predictive model [38]. De Mey et al. [38] introduced a simpler and more efficient methods to insert an artificial or replace a native promoter in *E. coli*, making it more practical to control gene expression with promoter.

Auto-responsive Promoter Auto-responsive promoters are virtually inducible. Different from other generally recognized inducible promoters, the induction of these promoters does not require the addition of certain inducible chemicals. These promoters response to molecules which is either one of the integral components of the environment (e.g., oxygen [40]) or the metabolites produced by the microorganisms accompanying their growth (e.g., signal molecules in the quorum sensing system (QS) of a bacterium [41]). The unique response profile of these promoters can be strategically adopted to design system to realize automatic induction when the expression of target genes is necessary.

QS-based Expression System. QS is a process that coordinates the population behavior of bacteria to act as an integrity by sensing a strain-specific small molecule [42, 43]. With QS, bacteria can specifically probe the presence of other individuals and the density of the population [44]. Bacteria can dynamically turn on (up)/off (down) the expression of many genes by response to the auto-inducers (AI) whose concentration is proportional to the density of bacterial cells [45].

By employing this mechanism, authors have succeeded in producing recombinant proteins by autonomous

induction through minimally rewiring native quorum sensing regulon of different strains, *E. coli* W3110 and BL21 [41]. The author tested their novel expression system with the overexpression of some model proteins: green fluorescent protein (GFP), chloramphenicol acetyl transferase (CAT), and β -galactosidase (*LacZ*). In this system, the expression of heterogeneous gene is under the control of autoinducer-2 (AI-2) (a small molecule whose concentration is closely correlated to the density of the cultured cells), enabling the automatic initiation of expression of target gene when the cell density is high enough. No manual operation (such as inducer addition and temperature shifting) is needed. Meanwhile, since the expression of target gene will and can only occur when the cell density reaches a certain value, the optimization of induction initiation will be avoided. It also eliminates the laborious continual monitoring of the bacterial growth to start the induction in every batch.

Oxygen-induced Expression of *vhb*. In most cases of bioindustrial process, cell growth is markedly influenced by dissolved oxygen (DO). Insufficient supply of DO will frequently constitute a big problem, especially when cell density is very high in aerobic fermentation. Discovery of the bacterial hemoglobin from *Vitreoscilla* sp. (VHb) [46] and its strong ability to binding DO provides an effective solution to alleviate the DO deficiency. Besides the oxygen-binding capability of VHb, the expression of *vhb* in *Vitreoscilla* sp. is also regulated by oxygen. Constructing genetic unit by putting the *vhb* gene under the control of its native promoter from *Vitreoscilla* sp. can realize the “intelligent” and autonomous induced expression of *vhb* when DO is limited, sophisticatedly avoiding the waste of cellular materials due to unnecessary expression when DO is adequate. This conditional expression is quite different from and advantageous than other cases in which the expression of genetically introduced genes is almost “constitutive” if the expression starts. The *vhb* promoter (P_{vhb}) has been successfully utilized in the expression of GFP protein and toluene dioxygenase with high copy-number plasmid by Liu et al. [47], and the combined application of P_{vhb} with VHb gives rise to a better productivity under low aeration. This energy-saving production process of protein is of special significance in industrial-scale operations. Integration of VHb gene and the P_{vhb} promoter into metabolic engineered strain as an ancillary component will bring it more industrial merits. The *vgb* gene has been used in improved production of recombinant proteins [48, 49], chemicals [50, 51], and antibiotics [52]. It could also enhance the performance of microbes in bioremediation [53], and improve the physiological state [54]. A comprehensive introduction to the *vhb*'s application in metabolic engineering can be found in a review [55].

Light Responsive Promoter The activation of a conventional inducible promoter depends on the existence of a chemical molecule (inducer) added into the medium or produced by the cells themselves, or heat shock. Once the expression of a gene is initiated, it cannot stop until depletion of the inducer: these promoters are not reversible. They are very simple and successful (although they carry inherent problems [56]); at same time, they are very suitable for expression that requires no time and space control. In 2002, Shimizu-Sato et al. [56] introduced the light induced expression of some reporter genes based on the phenomenon that the binding of the plant photoreceptor phytochrome to the protein PIF3 could be induced by red light, and dissociated by far-red light. In 2011, a green light inducible transcription system in *E. coli* based on green/red photoswitchable two-component system from cyanobacteria was reported by Tabor [57]. Light inducible expression system could provide rapid, noninvasive switchable control on gene by exposure to the light; it allows achieving multichromatic control of gene expression, realizing multistate control in the scale of the time and space [58]. This powerful light-controlled promoter enriches the choices of promoter in metabolic engineering and synthetic biology in which the coordinated expression of several genes are extremely expected, see reviews [58, 59].

Riboswitch Control

Riboswitch is a part of an mRNA molecule that can directly bind a small target molecule, and the binding can affect the gene's activity [60–62]. Therefore, mRNA containing a riboswitch can directly regulate its own activity by responding to the concentrations of its target molecule. The existence of riboswitches in all domains of life therefore adds some supports to the RNA world hypothesis. With the increasing know-how in detailed mechanism of riboswitch, some scientists have employed riboswitch to design finely regulated biosystems. In a recent interesting article, the author reported that they successfully modified bacterium that could seek and destroy controversial herbicide atrazine by controlling the translation of flagellar motor complex with a synthetic riboswitch [63].

The first use of a synthetic riboswitch to control bacterial mobility was reported by Topps and Gallivan [64]. They engineered an *E. coli* to follow a small molecule theophylline by taking advantage of the favorable feature of riboswitch that it can control gene expression in ligand-dependent and protein-coeffector-free fashion [65, 66]. The independence of accessory protein in riboswitch regulating gene expression makes it possible to construct regulatory riboswitch system capable of controlling gene expression in diverse bacterial species. Topp et al. [67] succeeded in

developing five synthetic riboswitches that efficiently induced gene expression in eight different bacterial species (Gram-negative and positive). Notably, with these riboswitches, they realized the induced gene expression in conditional human pathogen *Acinetobacter baumannii*, in which successful inducible expression of foreign gene has never been reported. More detailed information on biotechnological relevance of riboswitch can be found in two good reviews [68, 69].

Operator Region

The function of an operator region in an operon or a gene has been well recognized by the example of *lac* operon [70]. In the *lac* operon, the operator region can control the transcription of downstream genes by binding the repressor protein *lacI* to block the proceeding of a RNA polymerase. Similarly, in other operons/genes that contain such an element, operator region also plays an important role in regulating gene expression. However, operator region as an operating target is not popular in literatures related metabolic engineering. It is more frequently adopted to design an inducible gene expression system (e.g., the pET and derivative expression system). Stronger affinity between the operator region and repressor protein will minimize the background expression level of a gene by giving a more stringent control on its expression when the corresponding repressor is absent. It can be regarded as an alternative strategy for the widely adopted method for minimized background expression level of target gene by increasing the amount of repressor protein.

Intergenic Region

The human genome is larger than the sum of all the encoding sequences (including the generally recognized necessary regulatory elements) [71]. It is also the case for other lower living organisms (bacteria, fungi, plants, and other animals), despite the extensive existence of overlapped genes [72, 73]. This phenomenon implies there are many intergenic sequences whose functions are far from being thoroughly understood. Fortunately, the regulatory functions of these intergenic regions in gene expression have been discovered, to some extent [74–76]. Pflieger et al. [77] invented a method to realize the balanced expression of multiple genes as operons by the combinational utilization of various post-transcriptional control elements and tunable intergenic region (TIGR) screened from a library. They first tested the effect of an intergenic region on expression of two reporter genes and a change of 100-fold was observed. Due to the balanced expression of the genes in the heterologous biosynthetic pathway, a sevenfold productivity of mevalonate was achieved. Park et al. [78]

paid their attentions to 5'-untranslational region (5'-UTR) to tune the gene expression level, because the 5'-UTR contained sequences that influenced protein synthesis in a more direct way. In order to lower the background level of expression controlled by broad-host-range promoter *Pm*, Lale et al. [79] constructed and screened a library of the 5'-UTR, they observed that mutations in the UTR DNA region flanking the Shine–Dalgarno sequence strongly reduced the background expression level from *Pm* by reducing translational efficiency, while the response to induction remained unchanged (the allowed minimal concentration of inducer was 1 μ M).

Transcriptional Terminator

Compared with the attentions and efforts paid to promoter in metabolic engineering, the transcriptional terminator is apparently neglected. Transcriptional terminator is a segment of DNA sequence whose function is to define where the transcription of a gene or an operon should be stopped through different mechanisms [80].

Different transcriptional terminations will lead to the different version of transcripts, offering a powerful regulatory mechanism on expression of gene. In addition to its primary function to terminate a transcription, transcriptional terminator was found to stabilize its own mRNA [81]. Engineering transcriptional terminator might also be a target that can be used to regulate gene expression in bacterial metabolic engineering studies, because a similar result was confirmed in yeast [82]: a higher-level mRNA and protein of the cloned gene in a plasmid was achieved when TPS1 terminator (TPS1t) was selected from the four studied terminators CYC1t, TDH3t, and PGK1t.

Concluding Remarks

It is a quite straightforward idea to alter the metabolic pathway of an organism to a direction that is expected by the human being. To fulfill this task, at an earlier time, random mutations were introduced with the aid of chemomutagen or physical treatment (e.g., the most routinely used UV irradiation); combined with subsequent screening, expected mutants might be obtained. In most cases, except for the production of recombinant protein, a better productivity is frequently the results of a changed pathway. Therefore, such kind of operation is essentially metabolic engineering, at least “irrational” one. Now, we can do the same job in a quite different and more rational way due to the increased understanding in metabolic pathways in many microorganisms. The functional elucidation of the individual part of a gene (the dictator of enzymes that defines a metabolic pathway) enables us to

find the suitable targets that can be changed to alter the metabolic pathway. We can change the property of a metabolic enzyme by changing its encoding sequence through directed evolution or rational design, and control the time scale and strength of expression of the gene by a different promoter (strong or weak inducible promoter, auto-responsive promoter). The sophisticated mechanisms in quorum sensing and VHB functioning can be used to control gene expression in an automated way. Newly discovered and universally occurring riboswitch provides the most efficient and economical control strategies in gene expression at the transcriptional level across different species. In an operon of more than one gene, the intergenic region can also play an important role in regulation of the expression of all the member genes in it. These offer more targets that can be used to metabolically engineering microbes into a better biotechnology workhorse.

Acknowledgments The research was financially supported by grants from the National Science Foundation of China (No. 20625619, 20836003, 31000054); the Key Project of Chinese National Programs for Fundamental Research and Development (973 program, No. 2007CB71403, 2010CB535014); the Program for Changjiang Scholars and Innovative Research Team in University (No. IRT0532); the Fundamental Research Funds for the Central Universities (No. JUSRP10917).

References

1. Goldman, S. (2010). Genetic chemistry: production of non-native compounds in yeast. *Current Opinion in Chemical Biology*, *14*, 390–395.
2. Jarboe, L. R., Zhang, X., Wang, X., Moore, J. C., Shanmugam, K. T., & Ingram, L. O. (2010). Metabolic engineering for production of biorenewable fuels and chemicals: Contributions of synthetic biology. *Journal of Biomedicine and Biotechnology*, *2010*, 761042.
3. Keasling, J. D. (1999). Gene-expression tools for the metabolic engineering of bacteria. *Trends in Biotechnology*, *17*, 452–460.
4. Polisenio, L., Salmena, L., Zhang, J., Carver, B., Haveman, W. J., & Pandolfi, P. P. (2010). A coding-independent function of gene and pseudogene mRNAs regulates tumour biology. *Nature*, *465*, 1033–1038.
5. Jackel, C., Kast, P., & Hilvert, D. (2008). Protein design by directed evolution. *Annual Review of Biophysics*, *37*, 153–173.
6. Tracewell, C. A., & Arnold, F. H. (2009). Directed enzyme evolution: Climbing fitness peaks one amino acid at a time. *Current Opinion in Chemical Biology*, *13*, 3–9.
7. Butterfoss, G. L., & Kuhlman, B. (2006). Computer-based design of novel protein structures. *Annual Review of Biophysics and Biomolecular Structure*, *35*, 49–65.
8. Kang, S. G., & Saven, J. G. (2007). Computational protein design: Structure, function and combinatorial diversity. *Current Opinion in Chemical Biology*, *11*, 329–334.
9. Poole, A. M., & Ranganathan, R. (2006). Knowledge-based potentials in protein design. *Current Opinion in Structural Biology*, *16*, 508–513.
10. Foo, J. L., Ching, C. B., Chang, M. W., & Leong, S. S. (2011). The imminent role of protein engineering in synthetic biology.

- Biotechnol Advances*, in press. doi: [10.1016/j.biotechadv.2011.09.008](https://doi.org/10.1016/j.biotechadv.2011.09.008).
11. Wang, T. W., Zhu, H., Ma, X. Y., Zhang, T., Ma, Y. S., & Wei, D. Z. (2006). Mutant library construction in directed molecular evolution: Casting a wider net. *Molecular Biotechnology*, *34*, 55–68.
 12. Smith, J. M. (1970). Natural selection and the concept of a protein space. *Nature*, *225*, 563–564.
 13. Povolotskaya, I. S., & Kondrashov, F. A. (2010). Sequence space and the ongoing expansion of the protein universe. *Nature*, *465*, 922–926.
 14. Leonard, E., Ajikumar, P. K., Thayer, K., Xiao, W. H., Mo, J. D., et al. (2010). Combining metabolic and protein engineering of a terpenoid biosynthetic pathway for overproduction and selectivity control. *Proceedings of the National Academy of Science United States*, *107*, 13654–13659.
 15. Pfeifer, B. A., Admiraal, S. J., Gramajo, H., Cane, D. E., & Khosla, C. (2001). Biosynthesis of complex polyketides in a metabolically engineered strain of *E. coli*. *Science*, *291*, 1790–1792.
 16. Dueber, J. E., Wu, G. C., Malmirchegini, G. R., Moon, T. S., Petzold, C. J., et al. (2009). Synthetic protein scaffolds provide modular control over metabolic flux. *Nature Biotechnology*, *27*, 753–759.
 17. Davidson, A. R. (2006). Multiple sequence alignment as a guideline for protein engineering strategies. *Methods in Molecular Biology*, *340*, 171–181.
 18. Yoshikuni, Y., Dietrich, J. A., Nowroozi, F. F., Babbitt, P. C., & Keasling, J. D. (2008). Redesigning enzymes based on adaptive evolution for optimal function in synthetic metabolic pathways. *Chemistry & Biology*, *15*, 607–618.
 19. Cannarozzi, G., Schraudolph, N. N., Faty, M., von Rohr, P., Friberg, M. T., et al. (2010). A role for codon order in translation dynamics. *Cell*, *141*, 355–367.
 20. Fredrick, K., & Ibba, M. (2010). How the sequence of a gene can tune its translation. *Cell*, *141*, 227–279.
 21. Tuller, T., Carmi, A., Vestsigian, K., Navon, S., Dorfan, Y., et al. (2010). An evolutionarily conserved mechanism for controlling the efficiency of protein translation. *Cell*, *141*, 344–354.
 22. de Smit, M. H., & van Duin, J. (2003). Translational standby sites: How ribosomes may deal with the rapid folding kinetics of mRNA. *Journal of Molecular Biology*, *331*, 7737–7743.
 23. Studer, S. M., & Joseph, S. (2006). Unfolding of mRNA secondary structure by the bacterial translation initiation complex. *Molecular Cell*, *22*, 105–115.
 24. Juminaga, D., Baidoo, E. E., Redding-Johanson, A. M., Batth, T. S., Burd, H., et al. (2012). Modular engineering of L-tyrosine production in *Escherichia coli*. *Applied and Environment Microbiology*, *78*, 89–98.
 25. Kozak, M. (1999). Initiation of translation in prokaryotes and eukaryotes. *Gene*, *234*, 187–208.
 26. Kozak, M. (2005). Regulation of translation via mRNA structure in prokaryotes and eukaryotes. *Gene*, *361*, 13–37.
 27. Chen, H., Bjerknes, M., Kumar, R., & Jay, E. (1994). Determination of the optimal aligned spacing between the Shine-Dalgarno sequence and the translation initiation codon of *Escherichia coli* mRNAs. *Nucleic Acids Research*, *22*, 4953–4957.
 28. de Smit, M. H., & van Duin, J. (1990). Secondary structure of the ribosome binding site determines translational efficiency: A quantitative analysis. *Proceedings of the National Academy of Science United States*, *87*, 7668–7672.
 29. Salis, H. M., Mirsky, E. A., & Voigt, C. A. (2009). Automated design of synthetic ribosome binding sites to control protein expression. *Nature Biotechnology*, *27*, 946–950.
 30. Na, D., Lee, S., & Lee, D. (2010). Mathematical modeling of translation initiation for the estimation of its efficiency to computationally design mRNA sequences with desired expression levels in prokaryotes. *BMC Systems Biology*, *4*, 71.
 31. Wang, H. H., Isaacs, F. J., Carr, P. A., Sun, Z. Z., Xu, G., et al. (2009). Programming cells by multiplex genome engineering and accelerated evolution. *Nature*, *460*, 894–898.
 32. Alper, H., Fischer, C., Nevoigt, E., & Stephanopoulos, G. (2005). Tuning genetic control through promoter engineering. *Proceedings of the National Academy of Science United States*, *102*, 12678–12683.
 33. De Mey, M., Maertens, J., Boogmans, S., Soetaert, W. K., Vandamme, E. J., et al. (2010). Promoter knock-in: A novel rational method for the fine tuning of genes. *BMC Biotechnology*, *10*, 26.
 34. Lee, K. H., Park, J. H., Kim, T. Y., Kim, H. U., & Lee, S. Y. (2007). Systems metabolic engineering of *Escherichia coli* for L-threonine production. *Molecular Systems Biology*, *3*, 149.
 35. Anthony, J. R., Anthony, L. C., Nowroozi, F., Kwon, G., Newman, J. D., & Keasling, J. D. (2009). Optimization of the mevalonate-based isoprenoid biosynthetic pathway in *Escherichia coli* for production of the anti-malarial drug precursor amorpha-4,11-diene. *Metabolic Engineering*, *11*, 13–19.
 36. Qin, X., Qian, J., Yao, G., Zhuang, Y., Zhang, S., & Chu, J. (2011). GAP promoter library for fine-tuning of gene expression in *Pichia pastoris*. *Applied and Environment Microbiology*, *77*, 3600–3608.
 37. Qin, X., Qian, J., Xiao, C., Zhuang, Y., Zhang, S., & Chu, J. (2011). Reliable high-throughput approach for screening of engineered constitutive promoters in the yeast *Pichia pastoris*. *Letters in Applied Microbiology*, *52*, 634–641.
 38. De Mey, M., Maertens, J., Lequeux, G. J., Soetaert, W. K., & Vandamme, E. J. (2007). Construction and model-based analysis of a promoter library for *E. coli*: An indispensable tool for metabolic engineering. *BMC Biotechnology*, *7*, 34.
 39. Bakke, I., Berg, L., Aune, T. E., Brautaset, T., Sletta, H., et al. (2009). Random mutagenesis of the PM promoter as a powerful strategy for improvement of recombinant-gene expression. *Applied and Environment Microbiology*, *75*, 2002–2011.
 40. Wei, X. X., & Chen, G. Q. (2008). Applications of the VHB gene *vgb* for improved microbial fermentation processes. *Methods in Enzymology*, *436*, 273–287.
 41. Tsao, C. Y., Hooshangi, S., Wu, H. C., Valdes, J. J., & Bentley, W. E. (2010). Autonomous induction of recombinant proteins by minimally rewiring native quorum sensing regulon of *E. coli*. *Metabolic Engineering*, *12*, 291–297.
 42. Rosenberg, S. M. (2009). Life, death, differentiation, and the multicellularity of bacteria. *PLoS Genetics*, *5*, e1000418.
 43. Ng, W. L., & Bassler, B. L. (2009). Bacterial quorum-sensing network architectures. *Annual Review of Genetics*, *43*, 197–222.
 44. Waters, C. M., & Bassler, B. L. (2005). Quorum sensing: Cell-to-cell communication in bacteria. *Annual Review of Cell and Developmental Biology*, *21*, 319–346.
 45. Fuqua, C., Parsek, M. R., & Greenberg, E. P. (2001). Regulation of gene expression by cell-to-cell communication: Acyl-homoserine lactone quorum sensing. *Annual Review of Genetics*, *35*, 439–468.
 46. Khosla, C., & Bailey, J. E. (1988). The Vitreoscilla hemoglobin gene: Molecular cloning, nucleotide sequence and genetic expression in *Escherichia coli*. *Molecular and General Genetics*, *214*, 158–161.
 47. Liu, T., Chen, J. Y., Zheng, Z., Wang, T. H., & Chen, G. Q. (2005). Construction of highly efficient *E. coli* expression systems containing low oxygen induced promoter and partition region. *Applied Microbiology and Biotechnology*, *68*, 346–354.

48. Suthar, D. H., & Chattoo, B. B. (2006). Expression of Vitreoscilla hemoglobin enhances growth and levels of alpha-amylase in *Schwanniomyces occidentalis*. *Applied Microbiology and Biotechnology*, *72*, 94–102.
49. Chien, L. J., & Lee, C. K. (2005). Expression of bacterial hemoglobin in the yeast, *Pichia pastoris*, with a low O₂-induced promoter. *Biotechnology Letters*, *27*, 1491–1497.
50. Dogan, I., Pagilla, K. R., Webster, D. A., & Stark, B. C. (2006). Expression of Vitreoscilla hemoglobin in *Gordonia anarae* enhances biosurfactant production. *Journal of Industrial Microbiology and Biotechnology*, *33*, 693–700.
51. Geckil, H., Barak, Z., Chipman, D. M., Erenler, S. O., Webster, D. A., & Stark, B. C. (2004). Enhanced production of acetoin and butanediol in recombinant *Enterobacter aerogenes* carrying Vitreoscilla hemoglobin gene. *Bioprocess and Biosystems Engineering*, *26*, 325–330.
52. Brunker, P., Minas, W., Kallio, P. T., & Bailey, J. E. (1998). Genetic engineering of an industrial strain of *Saccharopolyspora erythraea* for stable expression of the Vitreoscilla haemoglobin gene (vhb). *Microbiology*, *144*(Pt 9), 2441–2448.
53. Kim, Y., Webster, D. A., & Stark, B. C. (2005). Improvement of bioremediation by *Pseudomonas* and *Burkholderia* by mutants of the Vitreoscilla hemoglobin gene (vgb) integrated into their chromosomes. *Journal of Industrial Microbiology and Biotechnology*, *32*, 148–154.
54. Zhu, H., Wang, T. W., Sun, S. J., Shen, Y. L., & Wei, D. Z. (2006). Chromosomal integration of the Vitreoscilla hemoglobin gene and its physiological actions in *Tremella fuciformis*. *Applied Microbiology and Biotechnology*, *72*, 770–776.
55. Zhang, L., Li, Y., Wang, Z., Xia, Y., Chen, W., & Tang, K. (2007). Recent developments and future prospects of Vitreoscilla hemoglobin application in metabolic engineering. *Biotechnology Advances*, *25*, 123–136.
56. Shimizu-Sato, S., Huq, E., Tepperman, J. M., & Quail, P. H. (2002). A light-switchable gene promoter system. *Nature Biotechnology*, *20*, 1041–1044.
57. Tabor, J. J., Levskaya, A., & Voigt, C. A. (2011). Multichromatic control of gene expression in *Escherichia coli*. *Journal of Molecular Biology*, *405*, 315–324.
58. Camsund, D., Lindblad, P., & Jaramillo, A. (2011). Genetically engineered light sensors for control of bacterial gene expression. *Biotechnology Journal*, *6*, 826–836.
59. Drepper, T., Krauss, U., Meyer zu Berstenhorst, S., Pietruszka, J., & Jaeger, K. E. (2011). Lights on and action! Controlling microbial gene expression by light. *Applied Microbiology and Biotechnology*, *90*, 23–40.
60. Vitreschak, A. G., Rodionov, D. A., Mironov, A. A., & Gelfand, M. S. (2004). Riboswitches: The oldest mechanism for the regulation of gene expression? *Trends in Genetics*, *20*, 44–50.
61. Tucker, B. J., & Breaker, R. R. (2005). Riboswitches as versatile gene control elements. *Current Opinion in Structural Biology*, *15*, 342–348.
62. Batey, R. T. (2006). Structures of regulatory elements in mRNAs. *Current Opinion in Structural Biology*, *16*, 299–306.
63. Sinha, J., Reyes, S. J., & Gallivan, J. P. (2010). Reprogramming bacteria to seek and destroy an herbicide. *Nature Chemical Biology*, *6*, 464–470.
64. Topp, S., & Gallivan, J. P. (2007). Guiding bacteria with small molecules and RNA. *Journal of the American Chemical Society*, *129*, 6807–6811.
65. Roth, A., & Breaker, R. R. (2009). The structural and functional diversity of metabolite-binding riboswitches. *Annual Review of Biochemistry*, *78*, 305–334.
66. Montange, R. K., & Batey, R. T. (2008). Riboswitches: Emerging themes in RNA structure and function. *Annual Review of Biophysics*, *37*, 117–133.
67. Topp, S., Reynoso, C. M., Seeliger, J. C., Goldlust, I. S., Desai, S. K., et al. (2010). Synthetic riboswitches that induce gene expression in diverse bacterial species. *Applied and Environment Microbiology*, *76*, 7881–7884.
68. Wieland, M., & Fussenegger, M. (2010). Ligand-dependent regulatory RNA parts for Synthetic Biology in eukaryotes. *Current Opinion in Biotechnology*, *21*, 760–765.
69. Mishler, D. M., Topp, S., Reynoso, C. M., & Gallivan, J. P. (2010). Engineering bacteria to recognize and follow small molecules. *Current Opinion in Biotechnology*, *21*, 653–656.
70. Jacob, F., & Monod, J. (1961). Genetic regulatory mechanisms in the synthesis of proteins. *Journal of Molecular Biology*, *3*, 318–356.
71. Claverie, J. M. (2001). Gene number. What if there are only 30,000 human genes? *Science*, *291*, 1255–1257.
72. Rocha, E. P. (2008). The organization of the bacterial genome. *Annual Review of Genetics*, *42*, 211–233.
73. Normark, S., Bergstrom, S., Edlund, T., Grundstrom, T., Jaurin, B., et al. (1983). Overlapping genes. *Annual Review of Genetics*, *17*, 499–525.
74. Smolke, C. D., & Keasling, J. D. (2002). Effect of gene location, mRNA secondary structures, and RNase sites on expression of two genes in an engineered operon. *Biotechnology and Bioengineering*, *80*, 762–776.
75. Smolke, C. D., Carrier, T. A., & Keasling, J. D. (2000). Coordinated, differential expression of two genes through directed mRNA cleavage and stabilization by secondary structures. *Applied and Environment Microbiology*, *66*, 5399–5405.
76. Berg, L., Lale, R., Bakke, I., Burroughs, N., & Valla, S. (2009). The expression of recombinant genes in *Escherichia coli* can be strongly stimulated at the transcript production level by mutating the DNA-region corresponding to the 5′-untranslated part of mRNA. *Microbial Biotechnology*, *2*, 379–389.
77. Pflieger, B. F., Pitera, D. J., Smolke, C. D., & Keasling, J. D. (2006). Combinatorial engineering of intergenic regions in operons tunes expression of multiple genes. *Nature Biotechnology*, *24*, 1027–1032.
78. Park, Y. S., Seo, S. W., Hwang, S., Chu, H. S., Ahn, J. H., et al. (2007). Design of 5′-untranslated region variants for tunable expression in *Escherichia coli*. *Biochemical and Biophysical Research Communications*, *356*, 136–141.
79. Lale, R., Berg, L., Stuttgart, F., Netzer, R., Stafsnes, M., et al. (2011). Continuous control of the flow in biochemical pathways through 5′ untranslated region sequence modifications in mRNA expressed from the broad-host-range promoter Pm. *Applied and Environment Microbiology*, *77*, 2648–2655.
80. Henkin, T. M. (1996). Control of transcription termination in prokaryotes. *Annual Review of Genetics*, *30*, 35–57.
81. Aiba, H., Hanamura, A., & Yamano, H. (1991). Transcriptional terminator is a positive regulatory element in the expression of the *Escherichia coli* crp gene. *The Journal of biological chemistry*, *266*, 1721–1727.
82. Yamanishi, M., Katahira, S., & Matsuyama, T. (2011). TPS1 terminator increases mRNA and protein yield in a *Saccharomyces cerevisiae* expression system. *Bioscience, Biotechnology, and Biochemistry*, *75*, 2234–2236.