Perspective

Metabolic Engineering: Perspective of a Chemical Engineer

Gregory Stephanopoulos
Dept. of Chemical Engineering, Massachusetts Institute of Technology, Cambridge MA 02139

Introduction

As hardly a day passes recently without yet another biological breakthrough, you may have asked your microbiologist friend whether he or she knows of a microbe capable of producing the molecule of your choice. In the unlikely event that one is not readily available, he or she may suggest some other microbe that makes a similar product convertible to the desired one, or, better yet, your microbiologist friend will formulate some screen to select microbes and their mutants that can produce a whole family of similar products with potentially better properties than the one initially sought.

The problem with most such candidate organisms is that they only make traces of the desired molecule and under conditions that may be difficult to implement on an industrial scale. These microbes must be improved before their potential can be realized. If you are similarly concerned about a particular disease, you may want to know how a newly discovered gene, or some other gene(s) buried in the sequenced genome, can help discover a drug for the disease or define a strategy for gene therapy. The answer to these questions depends critically on how well we can characterize the physiological state of cells and tissues, and use this information to prescribe the necessary genetic changes and/or environmental controls to improve such cells. This is also the goal and essence of metabolic engineering.

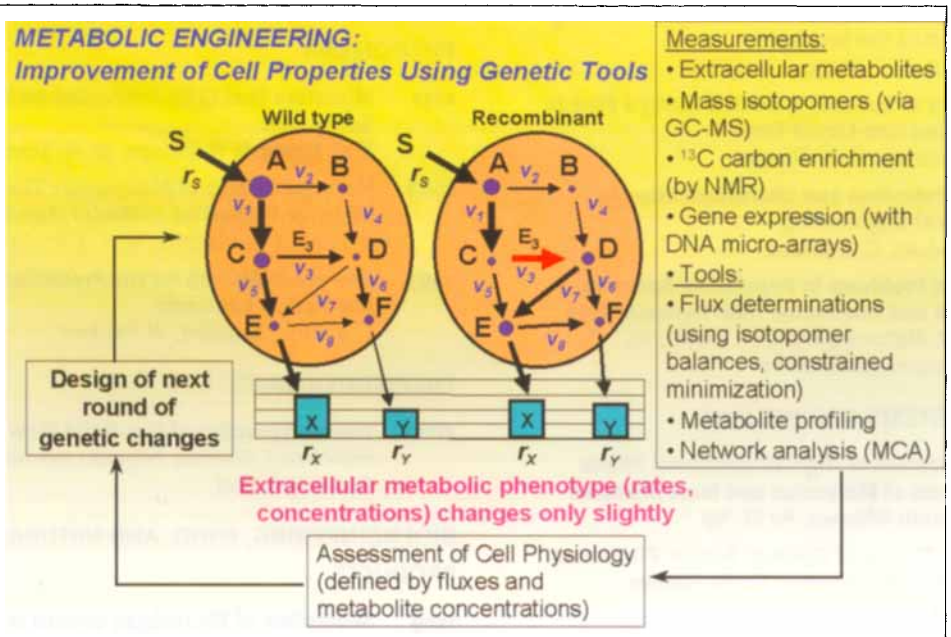

Figure 1. Unanticipated outcomes following single gene overexpression in metabolic engineering.
To increase the low yield of product Y in wild type cells, enzyme E3 is overexpressed. However, due to network interactions, this has minimal effect on the rates of product X and Y accumulation. Purple circles indicate pool size of metabolites in the network. Arrow thickness depicts relative flux magnitude of the corresponding reactions. Cell improvement is achieved by an iterative scheme that involves assessment of cell physiology at each step.

Metabolic engineering was developed in the previous decade to improve industrial strains using modern genetic tools. To be sure, microorganisms were being systematically improved by random mutagenesis and strain screening for at least 50 years before yielding overproducing strains of antibiotics, amino acids, citric acid, acetone-butanol solvents, and many other fermentation products. The development of recombinant technologies in the early 1980s gave rise to the notion of directed strain modification through the introduction of specific genes conferring desirable properties to cells for industrial, medical and environmental applications.

Metabolic engineering thus emerged as the scientific discipline occupied with the improvement of cellular properties through introduction to cells of specific transport, enzymatic or regulatory reactions using primarily recombinant technologies (Bailey, 1991). One should note the broader definition of metabolic engineering aiming at the improvement of *cellular properties* in general instead of simply the improvement of the yield or productivity of an industrial product. This broadening reflected recognition of the fact that important microbial traits, such as resistance to product inhibition,

viability and growth of aerobic organisms under hypoxic conditions, resistance to toxicity or high osmotic pressure, depend on specific genes and, as such, they can be manipulated just as well as product yield.

An impressive number of metabolic engineering applications have appeared in less than a decade along with a new conference series, a new journal, and several books on this subject. Research activity has yielded important results in diverse areas ranging from amino acid fermentations (Koffas, 2000) to polyketide and novel antibiotic synthesis (Pfeifer et al., 2001), indigo and aromatic amino acid synthesis in *Escherichia coli* (Murdock et al., 1993), golden rice (Ye et al., 2000), ethanologenic *E. coli* (Ingram et al., 1999), lycopene pathway for β-carotene synthesis (Farmer and Liao, 2000), indene biocatalysis for the synthesis of chiral pharmaceuticals (Stafford et al., 2002), tricistronic gene expression in Chinese Hamster Ovary cells for foreign protein overproduction under no growth and high viability conditions (Schlatter et al., 2001), manipulation of the glycosylation pathway in mammalian cells (Weikert et al., 1999), and many others (Ostergaard et al., 2000). In addition, a number of applications have appeared in the medical (Yarmush et al., 1999) and environmental (Keasling et al., 2000) areas.

A dominant element in the above applications is the extensive use of applied molecular biological methods for the introduction of genetic modifications and controls of cellular functions at the genetic level. A natural question then is whether metabolic engineering is nothing more than an industrial variant of genetic engineering and, if so, what is its particular relevance to chemical engineers. I will present in the sequel a broader vision of metabolic engineering that extends beyond the simple transfer of genes and manipulation of genetic controls in cells.

First, however, I would argue that even the transfer of genes, no matter how foreign to some, should not be excluded from the portfolio of research activities of chemical engineers. For one, gene transfer and manipulation are just steps in the construction of a better biocatalyst, typically carried out in simple operations using reagents available from commercial vendors. Additionally, the degree of sophistication of methods employed in molecular biological research is comparable to those required for making new materials with detailed molecular structures for use as catalysts or other applications, all areas where chemical engineers excel.

As chemical engineering research becomes increasingly molecular, I think that molecular biological applications should be encouraged to make this critically important area an important component of the portfolio of legitimate chemical engineering research. Manipulating genes, after all, is just another kind of chemistry and chemical engineers have demonstrated repeatedly that they can deploy these tools profitably in areas where no one ventured before them. As with all cutting-edge interdisciplinary research, rigorous training is a necessary element in metabolic engineering and most successful in an interdisciplinary environment where specialized expertise is available at all times.

Engineering Content of Metabolic Engineering

Rigorous Determination of the Metabolic Phenotype. Beside the synthetic part of genetic engineering discussed above and being enabled by molecular biology, metabolic engineering also comprises an equally important analytical part that is very rich in engineering content. This part deals with the rigorous evaluation of the physiology of cells or *metabolic phenotype*, as a means of assessing the impact of a genetic change in a recombinant strain relatively to the wild-type control. This evaluation was limited in the past to the measurement of simple macroscopic variables such as cell growth and extracellular metabolite exchange rates.

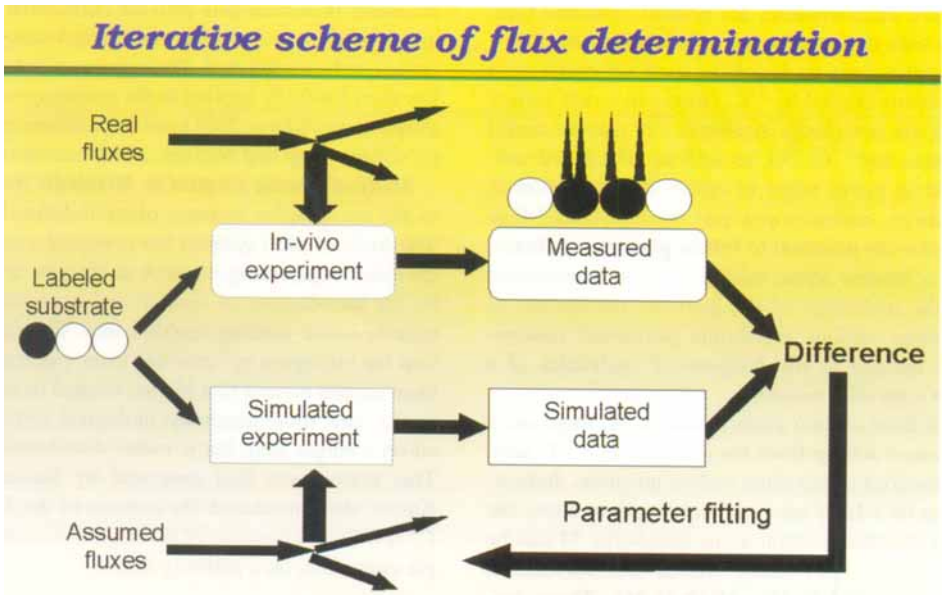

Figure 2. Iterative scheme of flux determination using labeled substrates and isotopomer measurements by GC-MS and ^{13}C NMR.

Such measurements, however, are often inadequate for assessing the true physiological state of a cell and directing the next round of genetic manipulations. Figure 1 shows the pathway in a wild type and a recombinant cell resulting from the amplification of enzyme E_3 deemed important for increasing the yield of extracellular metabolite product Y. However, unforeseen changes in the pools of intracellular metabolites (depicted by the diameter of corresponding circles) give rise to flux redistributions that minimize the overall impact of the increase in the activity of E_3. As a result, many-fold increases in enzyme activity often have only marginal effect on the rates of metabolite production. Furthermore, measurement of the extracellular products X and Y alone contain inadequate information for revealing the new intracellular flux configuration shown in the recombinant cell.

It has been suggested (Simpson et al., 1999) that a most accurate representation of the cellular physiological state is provided by the

pathway fluxes defined as the actual *in vivo* rates of pathway metabolite conversion. For a linear pathway at steady state, the pathway flux equals the reaction rates of the intermediate steps. The new fluxes in the recombinant cell, especially the *flux changes* relative to the wild type control, contain information about the distribution of kinetic control in the bioreaction network, which could be used to direct successive rounds of genetic modifications.

Fluxes are important determinants of cell physiology. The actual physiological state of a culture would be represented most accurately by variables describing what the cells *do* under a particular set of conditions rather than how the cells *look*, as is sometimes the practice. This information is provided most comprehensively by the vector of intracellular metabolic fluxes, ($[v_1, v_2, ..., v_8]$ for the network of Figure 1). These fluxes can be determined to varying extent depending on what measurements are available. For example, measurement of the extracellular rates r_S, r_X, and r_Y, along with the assumption of pseudo-steady state for the pools of intracellular metabolites A-F of Figure 1, allows the determination of some combination of the fluxes, but not each of them individually. For this purpose, more measurements are needed, obtained typically in combination with the introduction of a labeled substrate.

The majority of applications to date have used substrates with one or more carbon atoms labeled by ^{13}C. However, a rich variety of isotopic substrates are potentially available. ^{15}N may be useful to track nitrogen fluxes and ^{17}O ^{18}O, as well as ^{2}H labeled substrates, may be used in novel ways to reveal fluxes (Kelleher, 2001). In some situations, radioisotopes may be more useful than stable isotopes and offer the potential to follow phosphorus fluxes. The introduction of a labeled atom, such as ^{13}C, in a metabolic pathway will label the molecules of the pathway metabolites to varying extents creating various metabolite *positional isotopomers*. The latter are defined as the fractions of molecules of a metabolite labeled in a specific manner.

For example, for a three-carbon atom molecule M, there are 8 (2^3) types of isotopomers arising from the presence of a ^{13}C label at all possible combinations of the three carbon positions. Indicating by * the presence of a label on a particular carbon atom, the eight isotopomers of the three-carbon atom metabolite M can be represented as: M(1,2,3), M(1*,2,3), M(1,2*,3), M(1,2,3*), M(1*,2*,3), M(1*,2,3*), M(1,2*,3*), M(1*,2*,3*). These isotopomers can be observed by two general measurement methods: NMR spectroscopy and gas chromatography-mass spectrometry (GC-MS). The former allows the measurement of total isotopic carbon enrichment of each of the metabolite carbons. For example, the measurement by NMR of the fraction of metabolite molecules labeled on carbon 1, C(1), of the above metabolite M is related to the isotopomer fractions as follows:

$$C(1) = M(1*,2,3) + M(1*,2*,3) + M(1*,2,3*) + M(1*,2*,3*) \quad (1a)$$

And similarly for other carbon atoms,

$$C(2) = M(1,2*,3) + M(1*,2*,3) + M(1,2*,3*) + M(1*,2*,3*) \quad (1b)$$
$$C(3) = M(1,2,3*) + M(1*,2,3*) + M(1,2*,3*) + M(1*,2*,3*) \quad (1c)$$

GC-MS, on the other hand, provides measurements of *mass isotopomers*, i.e., the fractions of M molecules with the same mass. Indicating by M the fraction of unlabeled molecules of the metabolite, M+1 the fraction of molecules labeled on only one carbon, etc., mass isotopomer measurements are related to the positional isotopomers as follows:

$$M = M(1,2,3) \quad (2a)$$
$$M+1 = M(1*,2,3)+M(1,2*,3)+M(1,2,3*) \quad (2b)$$
$$M+2 = M(1*,2*,3)+M(1*,2,3*)+M(1,2*,3*) \quad (2c)$$
$$M+3 = M(1*,2*,3*) \quad (2d)$$

For a particular type of labeling substrate, intracellular fluxes determine the distribution of all positional isotopomers which, in turn, are mapped into the carbon enrichment and mass isotopomer measurements through Eqs. 1 and 2. The problem of flux determination is then one of inverting the above relationship subject to the stoichiometric constraints of each isotopomer at a metabolic and isotopic steady state. For a typical metabolic network there may be of the order of several hundred to a few thousand such constraints combined with a few hundred measurements. The problem is one of nonlinear constraint minimization type with interesting issues of *observability* and solution *multiplicity*. This is a challenging class of problems whose solution, however, yields *holistic* results about the performance of metabolic networks in their entirety as opposed to specific results about individual reactions. It is noted that the available measurements provide substantial redundancy yielding an *overdetermined system* whose consistency and statistical properties can be established. This approach is depicted in Figure 2 and has been fruitfully applied in the investigation of *Corynebacterium glutamicum* (Klapa, 2001) and *Penicillium chrysogenum* physiology (Christensen and Nielsen, 2000) among others.

Study of Kinetic Control in Metabolic Networks. Reaction networks are complex systems, often ill-defined and poorly observed. The study of such systems has occupied a considerable fraction of chemical engineering research in the past and has been facilitated by the introduction of simplifying assumptions such as the presence of a rate-limiting reaction step. The validity of this assumption for biological systems has been vigorously questioned. It has been argued instead that kinetic control in enzymatic reaction networks, and, more generally, biological systems, is not concentrated on a single step, but is rather distributed among several steps. This notion was first proposed by Kacser and Burns (1973). Kacser also introduced the concept of the *flux control coefficient* C_i^J (FCC) as a measure of the degree of control exercised by a single enzyme E_i on a pathway flux J

$$C_i^J = (dJ/dE_i)(E_i/J) \quad (3)$$

Thus, the (dimensionless) FCC can be thought of as the fractional change of the pathway flux divided by the fractional change in the amount or activity of a particular enzyme. In essence, this framework, known as metabolic control analysis (MCA), is a sensitivity analysis of system-wide properties such as the flux of a pathway or network of reactions, with respect to the various parameters affecting the pathway flux, such as enzymatic activity. Of the several theorems accompanying the MCA framework (see Fell (1997)), perhaps the most influential in terms of shaping a new mind frame of pathway understanding is the summation theorem stating that the sum of all FCCs of a flux with respect to all pathway enzymes equals unity

$$\Sigma_i C_i^J = 1 \quad (4)$$

For a linear pathway, for example, all FCCs are positive and less than unity. Unless its FCC is very close to unity, an enzyme does not qualify as a rate-limiting step. Consequently, manyfold increases in the amount or activity of this enzyme will bring about only

marginal changes in the pathway flux. This simple concept is at odds with persistent efforts to identify the "rate limiting enzyme" of a pathway, but it can explain the often-disappointing results obtained by the amplification of such enzymes. Furthermore, it questions the validity of the *silver bullet*-based approaches to many biotechnological applications ranging from drug development to industrial production.

MCA was developed initially in the field of biochemistry where it is presently influencing the scientific thinking about enzymatic kinetics and their systemic effects on cellular function. It was embraced quickly by engineers seeking a better understanding of the behavior of entire bioreaction networks, who also further extended its concepts and applications. For larger systems, for example, reaction grouping is a very useful concept in identifying the relative importance of *groups* of reactions for a flux of interest. This is expressed by *group flux control coefficients, gFCC,* measuring the impact of introducing changes in the reactions of one group on the flux of another group. For the lysine biosynthetic network, for example, reactions can be organized in three groups: Group A comprising all glycolytic and pentose phosphate pathway reactions, group B comprising the citric acid cycle reactions, and group C of all reactions participating in the pathway of biosynthesis of aspartic acid family of aminoacids.

Following the introduction of several perturbations (Simpson et al., 1998), whose effect was confined in the corresponding groups, the gFCCs were determined (Table 1). Simple inspection of the table reveals that most of the kinetic control for lysine biosynthesis resides in the lysine pathway. This is a powerful result as it now directs attention to a specific area of a large network of reactions for further investigation. I should note that the results of Table 1 were obtained from flux measurements alone in a set of experiments that lasted approximately six months. Of particular interest is the fact that an important enzyme in the lysine pathway (pyruvate carboxylase) had not been identified at the time that the above table was compiled, and also that the predictions of Table 1 were recently validated experimentally in a qualitative sense (Koffas, 2000).

Table 1. Perturbed Reaction Group*

Affected Group				
	A	0.07	0.51	0.42
	B	0.09	1.22	–0.31
	C	0.02	–0.34	1.32

*gFCC for three reaction groups formed around the branch point of Phosphoenolpyruvate and Pyruvate for the lysine biosynthetic network. Group A comprises all glycolytic and pentose phosphate pathway reactions, group B reactions in the citric acid cycle and group C anaplerotic and other reactions in the pathway of aspartic acid family aminoacids biosynthesis. Table entries describe the relative change in the flux of the affected group per unit of relative change in the enzymatic activities of the perturbed reaction group.

Framework of Integration for the Study of Cellular Function. Metabolic engineering is concerned with the dual task of changing cells so that they are more effective for the intended industrial or medical application, as well as rigorously evaluating the resulting cellular physiology. Genes and gene regulation are essential for the first task, along with experimental methods from molecular biology. Cell physiology, on the other hand, is a multivariable concept that needs to be defined and articulated in a holistic manner. This was illustrated earlier by the methods aiming at the determination of intracellular fluxes. As more methods are developed for the *parallel* measurement of various classes of important intracellular molecules and increasing amounts of data related to the cellular phenotype accumulate, the need for a framework of integration of this information will become more pronounced.

Integration has been all along a distinguishing characteristic of metabolic engineering. From the very beginning, metabolic engineering has been interested in the properties of cellular systems and metabolic bioreaction networks *in their entirety* rather than individual genes and enzymes (Stephanopoulos and Sinskey, 1993). As such, it has addressed questions of pathway synthesis (Mavrovouniotis et al., 1992), thermodynamic feasibility and properties of pathways, distribution of kinetic control in metabolic networks, metabolic flux, and flux control (see previous subsection), and metabolic reconstruction from individual enzymatic reactions (Stephanopoulos et al., 1998). This is an excellent foundation for furthering the goals of metabolic engineering as it seeks to develop understanding of the overall cellular system from the properties of its constituent elements. This integration will be accomplished through metabolite and isotopomer balances, as described earlier, but also by using pattern discovery methods aiming at establishing associations between classes of genomic and physiological data. Besides advancing its central goal of physiology determination, *pattern discovery* tools from metabolic engineering will have a very significant impact in structuring the intellectual content of the emerging field of *Systems Biology.*

Some Important Problems

The great promise of and initial excitement about genetic engineering was that superior cells could be constructed that would be sustained at high viability under no or minimal growth, utilize an inexpensive substrate and secrete the product of interest at high rates, and make little or no other side product for maximal yield. It was soon realized, however, that to accomplish this goal, more is needed than the simple introduction of product forming genes in a host production cell. The reason is that gene products interact in intricate, nonlinear, and unpredictable ways so that simple genetic transformations are accompanied by unanticipated and often undesirable results.

This is not surprising in light of the distribution of kinetic control discussed earlier and the structural and regulatory complexity of cellular systems. The implication for metabolic engineering is that, to achieve the above goal of a superior cell, one would have to attempt the coordinated expression of several genes. Similar to microbial metabolite overproduction, strategies for restoration of normal human cell function or screening for new drugs will have a higher probability of success if they target multiple reaction steps. Although single gene modification of substrate transport or product forming pathway produced interesting results in early studies, genetic modification of several steps will be key to altering truly systemic cellular properties. *Identifying such genes and controlling them effectively is a core problem of metabolic engineering.* For small (product forming or biocatalytic) pathways, rational approaches can be applied whereby the properties of the pathway are expressed and optimized as functions of the individual enzymes with well-defined and reliable kinetics (Stafford et al., 2002). Fundamental aspects of flux determination and flux control, as presented in the previous section, are instrumental to this end.

On the other hand, complex metabolic networks are initially better handled by combinatorial methods testing large numbers of random sets of genes as to their impact on the targeted cellular properties and industrial figures of merit. Combinatorial approaches will benefit from good molecular biological techniques for the con-

struction of the libraries of random gene sets, and high throughput technologies for the efficient evaluation of the large number of resulting mutants. This mind-frame of research is exemplified in recent developments of single enzyme and pathway optimization using methods like directed evolution (Kuchner and Arnold, 1997) and gene shuffling (Stemmer, 2000). Once a satisfactory biocatalyst has been obtained by such combinatorial approaches, it can be further optimized by methods involving rational pathway flux analysis and elucidation of flux control.

The engineering part of metabolic engineering emphasizes physiology and cell function, described primarily by the vector of intracellular metabolic fluxes (referred to as the *metabolic phenotype*). An important recent development is the introduction of DNA microarrays for the routine, genome-wide measurement of gene expression (Schena et al., 1995). These microarrays are glass microscope slides containing thousands of specific probes for individual genes whose expression can be detected in parallel through hybridization of the microarray probes with the corresponding cDNA strands of the sample properly labeled by fluorescent dyes. DNA microarrays can thus yield data on the expression levels of a large number of genes from cellular and tissue samples, and these data collectively define the *expression phenotype.*

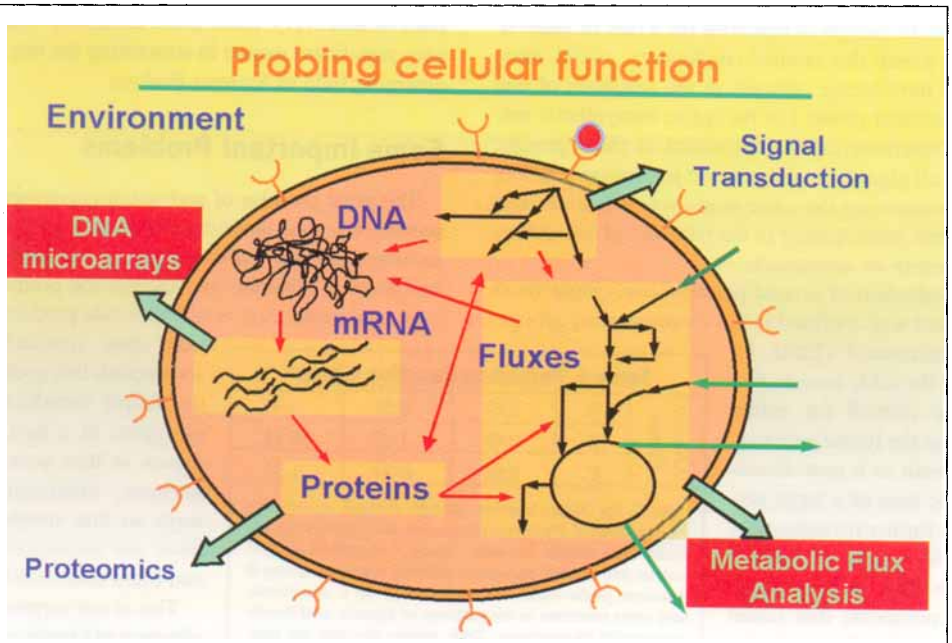

Figure 3. Cellular functions by distinct groups of molecules (e.g., genes, mRNA, proteins, storage compounds, and small metabolites), hierarchically organized and regulated by molecular interactions involving molecules from all organizational levels.

A central problem in the quest for understanding the relationship between gene expression and cellular function is the *linkage between the expression and metabolic phenotypes.* It is not clear how this linkage will be most efficiently accomplished. Multivariate analysis and methods of pattern discovery will certainly play important roles. The importance of this endeavor stems from the fact that it will allow, for the first time, the derivation of *relationships* between aspects of cellular function (exemplified by the metabolic phenotype, among other parameters) and a specific set of genes whose activity is reflected in the expression phenotype. Elucidation of such linkages will provide important information about gene function especially in relationship to desirable cellular phenotypes and yield candidate genes for conferring to cells desirable industrial or medical traits.

Figure 3 shows the main intracellular processes. It shows how a ligand-receptor binding event initiates a cascade of signaling reactions leading to the activation of a transcription factor that initiates transcription of one or more genes. Thus, genes are activated to generate mRNA templates (transcripts) that in turn provide the code for the formation of proteins responsible for vital cellular functions, such as metabolic conversions or transmission of signals, among many others. The individual small metabolite molecules thus affected by such reactions modulate the activity of enzymes and, along with other regulatory factors, regulate the activity of genes. These processes can be conceptualized as occurring at four distinct, but interacting, layers of hierarchy: genes, transcripts (mRNA), proteins, and small metabolites. While this organization is useful in describing the dominant hierarchical structure of cellular reactions, it is important to note that there is extensive interaction among the classes of molecules at all levels, primarily through modulation of the processes of gene expression, translation, and protein activity by gene products and small metabolites. Until recently, no means were available to probe the diverse interactions shown in the figure.

The advent of microarray technology and the use of stable isotopes in conjunction with NMR and/or GC-MS allow quantitative measurements of two important groups of variables, the pool of mRNA transcripts, and the rates (fluxes) of metabolic reactions, respectively. Proteomics promises to provide measures of protein profiles and fractions of activated molecules in signal transduction. Although we are some distance from realizing this vision, one can foresee the time when sufficient measurements will be available to allow the detailed modeling of a simple unicellular organism such as *Escherichia coli.* From a conceptual standpoint, this is a realistic, albeit immense, undertaking that could be materialized with a *coherent and directed* effort of several laboratories with partially overlapping expertise in the measurement and mathematical analysis of complex reacting systems. The chemical engineering paradigm has been applied successfully to similar systems in the past.

Future Outlook

Metabolic engineering, along with the broader biological sciences, has entered a period of rapid change. Changes are precipitated by a number of driving forces.

1. The rapid sequencing of whole genomes *(genomics)* is generating fundamental data of immense value in cataloguing the basic blueprint of life. Genomics thus is supplying information about genes and gene regulation, both important ingredients of metabolic engineering.

2. Applied molecular biological methods have reached a point of maturity where they can be applied on a routine basis in implementing biological constructions that introduce controls of metabolic processes at the genetic level.

3. There is increased emphasis on renewable resource utilization, an area that has traditionally benefited from advanced biotechnological processes.

4. An enormous array of opportunities in the medical, specialty chemicals, materials, fuels and environmental areas are being revisited as the capabilities of enzymes and cells continue to expand. In this regard it is important to note that there is no molecule of commercial interest that some microbe under some conditions is not able to produce and many novel molecules can be synthesized to meet present and future needs.

The issue with all such applications has been one of economics, and this is precisely the area that metabolic engineering is addressing through its focus on the construction of better biocatalysts. Finally, one should not lose sight of the fact that all this activity is fueled by an enormous public investment in the life sciences which is conservatively estimated to approximately one-quarter of a trillion U.S. dollars over the past 10 years in the U.S. alone. Such a concentration of R&D funding is bound to create enormous opportunities for new products and processes and valuable intellectual property centered around biology as enabling science.

So, why might chemical engineers be interested in metabolic engineering? First, I believe that metabolic engineering combines the intellectual framework and implementation tools required to capture the enormous potential of biology for industrial and medical applications. Its concepts and tools should be familiar to chemical engineers as metabolic engineering borrows heavily from chemical reaction engineering. The importance of metabolic engineering in materials, fuels, and specialty chemicals (pharmaceuticals and chiral compounds) is undeniable as evidenced by a growing number of applications in these areas. The greatest impact of metabolic engineering in the medical field will be in the development of methods for the rigorous assessment of the physiological state and determination of reasonable enzymatic targets for the treatment of disease. This will be implemented either by direct therapeutic intervention or screening programs for the discovery of new drugs.

In addition to the above rather obvious applications of metabolic engineering, the second reason that justifies the continuing interest of chemical engineers in this area is that it is an excellent entry for them into a very rich field of scientific inquiry. Think for a minute of one of the simplest systems of transcriptional regulation, that of the *lac* operon, that regulates the expression of the two genes whose enzyme products are responsible for the transport of lactose into the cell and its breakdown into the two simple catabolizable sugars of glucose and galactose. When there is no lactose present, the transcription of these two genes is blocked by a constitutively expressed protein that prevents attachment of the RNA polymerase by binding on an operator binding site just upstream of the above two structural genes. When lactose is present, it intercepts this binding protein thus freeing the operator and allowing the transcription process to proceed unimpeded.

There are very few nonbiological systems possessing anything that remotely resembles the efficiency and beauty of the *lac* operon. Yet, all biological systems owe their exceptional properties to specific chemical reactions catalyzed by enzymes that, in a growing number of cases, can be uniquely prescribed from genomic information. In other words, genomics provide the means to define the specific steps of the chemical reaction system that can be subsequently analyzed using the tools of metabolic engineering. This is a profound difference from typical chemical reacting systems where defining the actual reaction steps is a major challenge. Obviously, chemical engineers are in a unique position to help elucidate such systems and realize the vision of a comprehensive cell model presented in the previous section. The vision of a single cell model that accurately simulates all cellular functions and predicts its responses to all environmental stimuli is a challenge that is commensurate in importance and magnitude to genome sequencing. Just the same, it represents a noble and feasible goal in which chemical engineers can play a central role.

Chemical reactions are, for the most part, responsible for the wonders of biology. Metabolic engineering combines the tools and concepts of reaction engineering and molecular biology for the analysis and purposeful modification of bioreaction networks. It also provides a framework for integrating and quantifying genomic information and cell-wide data generated from modern technologies. As such, it is the natural vehicle for capturing the enormous potential of biology and transforming it into the enabling science of many new industrial and medical applications. Chemical engineers are in a unique position to extend their educational and research paradigm into the most exciting field of scientific inquiry. This will require that they embrace biology as a foundational science equal to chemistry and modify the curriculum such as to reflect this fundamental change into chemical and biological engineering.

Acknowledgment

Support from the National Science Foundation (Grant Number: BES - 9985421), the Department of Energy (Engineering Research Program, Office of Basic Energy Science Grant Number: DE-FG02-94ER14487), and the DuPont-MIT Alliance is gratefully acknowledged.

Literature Cited

Bailey, J. E., "Toward a Science of Metabolic Engineering," *Science,* **252**(5013), 1668 (1991).

Christensen, B., and J. Nielsen, "Metabolic Network Analysis of *Penicillium chrysogenum* Using C-13-Labeled Glucose," *Biotechnol. Bioeng.,* **68**(6), 652 (2000).

Farmer, W. R. and J. C. Liao, "Improving Lycopene Production in *Escherichia coli* by Engineering Metabolic Control," *Nat. Biotechnol.,* **18**(5), 533 (2000).

Fell, D. A., *Understanding the Control of Metabolism,* Portland Press, London (1997).

Ingram, L. O., H. C. Aldrich, A. C. C. Borges, T. B. Causey, A. Martinez, F. Morales, A. Saleh, S. A. Underwood, L. P. Yomano, S. W. York, J. Zaldivar, and S. D. Zhou, "Enteric Bacterial Catalysis for Fuel Ethanol Production," *Biotechnol. Prog.,* **15**(5), 855 (1999).

Kacser, H., and J. A. Burns, "The Control of Flux," *Symp. Soc. Exp. Biol.,* **27,** 65 (1973).

Keasling, J. D., S. J. Van Dien, P. Trelstad, N. Renninger, and K. McMahon "Application of Polyphosphate Metabolism to Environmental and Biotechnological Problems," *Biochem.-Moscow.*, **65**(3), 324 (2000).

Kelleher, J. K., "Flux Estimation Using Isotopic Tracers: Common Ground for Metabolic Physiology and Metabolic Engineering," *Metabolic Engineering*, **3**(2), 100 (2001).

Klapa, Maria Ioanni, "High Resolution Metabolic Flux Determination using Stable Isotopes and Mass Spectrometry," PhD Diss., Mass. Inst. Tech., Cambridge, MA (2001).

Koffas, Mattheos A. G., "Metabolic Engineering of *C. glutamicum* for Amino Acid Production," PhD Diss., MIT, Cambridge, MA (2000).

Kuchner O., and F. H. Arnold, "Directed Evolution of Enzyme Catalysts," *Trends Biotechnol.*, **15**(12), 523 (1997).

Mavrovouniotis, M., G. Stephanopoulos, and Gr. Stephanopoulos, "Synthesis of Biochemical Production Router," *Computers in Chemical Engineering*, **16,** 605 (1992).

Murdock, D., B. D. Ensley, C. Serdar, and M. Thalen, "Construction of Metabolic Operons Catalyzing the De-novo Biosynthesis of Indigo in *Escherichia-coli*," *Biotechnology (N.Y.)*, **11**(3), 381 (1993).

Ostergaard, S., L. Olsson, M. Johnston, and J. Nielsen, "Increasing Galactose Consumption by *Saccharomyces cerevisiae* through Metabolic Engineering of the GAL Gene Regulatory Network," *Nat. Biotechnol.*, **18**(12), 1283 (2000).

Pfeifer, B. A., S. J. Admiraal, H. Gramajo, D. E. Cane, and C. Khosla, "Biosynthesis of Complex Polyketides in a Metabolically Engineered Strain of *E-coli*," *Science*, **291**(5509), 1790 (2001).

Schena M., D. Shalon, R. W. Davis, and P. O. Brown, "Quantitative Monitoring of Gene-Expression Patterns with a Complementary-DNA Microarray," *Science*, **270**(5235), 467 (1995).

Schlatter, S., J. E. Bailey, and M. Fussenegger, "Novel Surface Tagging Technology for Selection of Complex Proliferation-Controlled Mammalian Cell Phenotypes," *Biotechnol. Bioeng.*, **75**(5), 597 (2001).

Simpson, T. W., H. Shimizu, and G. Stephanopoulos, "Experimental Determination of Group Flux Control Coefficients in Metabolic Networks," *Biotechnol. & Bioengineering*, **58,** 149 (1998).

Simpson, T. W., B. D. Follstad, and G. Stephanopoulos, "Analysis of the Pathway Structure of Metabolic Networks," *J. Biotechnol.*, **71**(1–3), 207 (1999).

Stafford, D. E., K. S. Yanagimachi, P. A. Lessard, S. K. Rijhwani, A. J. Sinskey, and G. Stephanopoulos, "Optimizing Bioconversion Pathways through Systems Analysis and Metabolic Engineering," *Proc. Natl. Acad. Sci. U.S.A.*, **99**(4), 1801 (2002).

Stemmer, W. P. C., "Molecular Breeding of Genes, Pathways, and Genomes by DNA Shuffling," *Abstr. Pap. Am. Chem. S.*, **219,** 104 (2000).

Stephanopoulos, G., A. Aristidou, and J. Nielsen, *Metabolic Engineering: Principles and Methodologies*, Academic Press, Inc., San Diego, CA (1998).

Stephanopoulos, G., and A. J. Sinskey, "Metabolic Engineering - Methodologies and Future-Prospects," *Trends Biotechnol.*, **11**(9), 392 (1993).

Weikert, S., D. Papac, J. Briggs, D. Cowfer, S. Tom, M. Gawlitzek, J. Lofgren, S. Mehta, V. Chisholm, N. Modi, S. Eppler, K. Carroll, S. Chamow, D. Peers, P. Berman, and L. Krummen, "Engineering Chinese Hamster Ovary Cells to Maximize Sialic Acid Content of Recombinant Glycoproteins," *Nat Biotechnol*, **17**(11), 1116 (1999).

Yarmush, D. M., A. D. MacDonald, B. D. Foy, F. Berthiaume, R. G. Tompkins, and M. L.Yarmush, "Cutaneous Burn Injury Alters Relative Tricarboxylic Acid Cycle Fluxes in Rat Liver," *J. of Burn Care & Rehabilitation*, **20**(4), 292 (1999).

Ye, X., S. Al-Babili, A. Kloti, J. Zhang, P. Lucca, P. Beyer, and I. Potrykus, "Engineering the Provitamin A (beta-carotene) Biosynthetic Pathway into (carotenoid-free) Rice Endosperm," *Science*, **287**(5451), 303 (2000).

Zaslavskaia L. A., J. C. Lippmeier, C. Shih, D. Ehrhardt, A. R. Grossman, and K. E. Apt, "Trophic Obligate Conversion of an Photoautotrophic Organism through Metabolic Engineering," *Science*, **292**(5524), 2073 (2001).