Toward a Science of Metabolic Engineering

JAMES E. BAILEY

Application of recombinant DNA methods to restructure metabolic networks can improve production of metabolite and protein products by altering pathway distributions and rates. Recruitment of heterologous proteins enables extension of existing pathways to obtain new chemical products, alter posttranslational protein processing, and degrade recalcitrant wastes. Although some of the experimental and mathematical tools required for rational metabolic engineering are available, complex cellular responses to genetic perturbations can complicate predictive design.

THE METABOLIC ACTIVITIES OF LIVING CELLS ARE ACCOMplished by a regulated, highly coupled network of ~1000 enzyme-catalyzed reactions and selective membrane transport systems. However, metabolic networks that evolved in natural settings are not genetically optimized for the objectives important in practical applications. Hence, performance of bioprocesses can be enhanced by genetic modification of the cells.

Metabolic engineering is the improvement of cellular activities by manipulation of enzymatic, transport, and regulatory functions of the cell with the use of recombinant DNA technology. The opportunity to introduce heterologous genes and regulatory elements distinguishes metabolic engineering from traditional genetic approaches to improve the strain. This capability enables construction of metabolic configurations with novel and often beneficial characteristics. Cell function can also be modified through precisely targeted alterations in normal cellular activities. Examples in the manipulation of protein processing pathways, as well as of pathways involving smaller metabolites, will be highlighted here.

At present, metabolic engineering is more a collection of examples than a codified science. Results to date promise future technological benefits, as well as contributions to basic science, agriculture, and medicine. However, many studies have shown the feasibility of metabolic engineering methods without achieving the yields, rates, or titers (final concentrations) required for a practical process. Most experiments explore changes in a single gene, operon, or gene cluster. After a new strain has been created by such a manipulation, limitations arise that can in principle be addressed by subsequent genetic manipulation. An iterative cycle of a genetic change, an analysis of the consequences, and a design of a further change, analogous to that articulated for protein engineering (*1*), can be used to find an optimized strain. The few cases to date in which such a metabolic engineering cycle has been implemented have achieved success. An emerging base of strategies, tools, and experiences will aid in identifying, implementing, and refining which particular set of genetic manipulations is most effective in accomplishing a desired change in cellular function.

The author is the Chevron Professor of Chemical Engineering at the California Institute of Technology, Pasadena, CA 91125.

Recruiting Heterologous Activities for Strain Improvement

Cloning and expression of heterologous genes can serve several useful purposes, including extending an existing pathway to obtain a new product, creating arrays of enzymatic activities that synthesize a novel structure, shifting metabolite flow toward a desired product, and accelerating a rate-determining step. Introduction of a functional heterologous enzyme or transport system into an organism can result in the appearance of new compounds that may subsequently undergo further reactions. Difficulties in anticipating these further reactions are a central limitation of metabolic engineering.

Expression of a heterologous protein does not guarantee appearance of the desired activity. The protein must avoid proteolysis, fold properly, accomplish any necessary assembly and prosthetic group acquisition, be suitably localized, have access to all required substrates, and not encounter an inhibitory environment. Despite these potential barriers to the successful recruitment of heterologous cellular activities, the number and scope of positive experiments encourage further application of this approach.

Synthesis of new products is enabled by completion of partial pathways. The genetic and metabolic diversity that exists in nature provides a collection of organisms with a spectrum of substrate assimilation and product synthesis capabilities. However, many natural strains are imperfect from an applied perspective. Their performance can sometimes be enhanced by extension of their native pathways. Native metabolites can be converted to preferred end products by the genetic installation of a few well-chosen heterologous activities (Table 1).

For example, the final precursor in a current commercial process for ascorbic acid (vitamin C) synthesis is 2-keto-L-gulonic acid (2-KLG). One route to 2-KLG involves two successive fermentations. The first converts glucose to 2,5-diketo-D-gluconic acid (2,5-DKG) in *Erwinia herbicola*; the second fermentation, carried out in a species of *Corynebacterium*, transforms 2,5-DKG to 2-KLG. Researchers devised a way to convert glucose to 2-KLG in a single fermentation step by cloning the *Corynebacterium* enzyme 2,5-DKG reductase, which catalyzes the 2,5-DKG to 2-KLG conversion, into *E. herbicola* (*2*). A similar goal was achieved for 7-aminocephalosporanic acid (7ACA), the precursor for several semisynthetic cephem antibiotics (*3*).

Posttranslational modifications can influence the function of proteins. The types of modifications that occur can be affected by expression of cloned protein processing enzymes. For example, expression in Chinese hamster ovary (CHO) cells of β-galactoside α2,6-sialyltransferase (*4*) allows the formation of sialyl α2,6-galactosyl linkages on its surface glycoproteins. These terminal glycosylation linkages are normally absent from proteins produced in this industrial

cell line, including cloned erythropoietin. Thus, this strategy should enable erythropoietin made in recombinant CHO cells to more closely resemble human erythropoietin, which is rich in these linkages (*5*). In another study, mouse cells displayed the human H blood group antigen after transfection with human DNA (*6*).

Transferring multistep pathways: Hybrid metabolic networks. The transfer of genes that encode entire biosynthetic pathways to a heterologous host can provide more industrially robust strains, enhance productivity, or permit the use of less costly raw materials. Moreover, such experiments are useful for exploring the regulation and function of a multistep metabolic pathway in a particular species.

Transferring entire antibiotic biosynthetic pathways to heterologous hosts has been facilitated by the clustering of the genes involved (*7*). Genes for the biosynthesis of actinorhodin were transferred from *Streptomyces coelicolor* to *Streptomyces lividans*, enabling the latter strain to produce actinorhodin (*8*). Subsequently, clustered erythromycin biosynthetic genes from *Streptomyces erythreus* were transferred to *S. lividans*, which then synthesized an antibiotic indistinguishable from erythromycin A (*9*). *Escherichia coli* carrying this cloned gene cluster did not synthesize the antibiotic, possibly because of low transcriptional activity of *Streptomyces* promoters in *E. coli*. The fungi *Neurospora crassa* and *Aspergillus niger*, which normally do not produce β-lactam antibiotics, synthesized penicillin V after transformation with a cosmid containing *Penicillium chrysogenum* DNA that encoded enzymes in the penicillin biosynthetic pathway (*10*).

Polyhydroxybutyrate (PHB), a storage product sequestered in large amounts by some bacteria under growth-limiting, carbon source–excess conditions, is a biodegradable polyester that already has small-scale applications. *Alcaligenes eutrophus* can produce not only PHB but, when supplied with different precursors, can synthesize various polyhydroxyalkanoate copolymers as well (*11*). Metabolic engineering of the synthesis of these and related polymers should provide greater control over the nature and quantity of the polymer produced and should also offer alternative production organisms. The PHB synthesis operon from *A. eutrophus*, which encodes PHB polymerase, thiolase, and reductase activities, has been used to transform *E. coli* (*12*). As in *A. eutrophus*, this recombinant *E. coli* accumulates PHB when the nitrogen source is depleted; PHB concentrations in these cultures reach 50% of the dry cell weight.

Assembly of pathways for simultaneous degradation of chloro- and methylaromatics by combining and refining of cloned pathway segments and regulatory systems from several different organisms exemplifies the iterative design of an effective hybrid organism (*13*). The ultimate strain thus far constructed contains five pathway segments obtained from three organisms. The biochemical and metabolic complexities of the degradation of mixed substrates and the resulting rationale behind each portion of this construction offer useful general perspectives on metabolic engineering strategies (*14*).

Creating new products and new reactants. Expression of biosynthetic genes for a secondary metabolite in a heterologous host that synthesizes its own different secondary metabolite can result in the construction of an array of enzymatic activities that yield novel products. Among the novel antibiotics that have been produced in recombinant strains of *Streptomyces* by such manipulations are mederrhodins A and B and dihydrogranatirhodin (*15*), 2-norerythromycins A, B, C, and D (*16*), and isovaleryl spiramycin (*17*).

Compounds new to the cell that result from a heterologous activity often undergo further reactions. In some cases, such as in the biosynthesis of indigo by *E. coli* that express *Pseudomonas putida* naphthalene dioxygenase (*18*), these subsequent reactions are essential components of the desired pathway. Another illustration of metabolic engineering to introduce a novel intermediate into a host involves recombinant *E. coli* that express the cloned tyrosinase gene from *Streptomyces antibioticus* (*19*). Synthesis by the recombinant *E. coli* strain of the pigment melanin, an ultraviolet light–absorbing compound with material and cosmetic applications, depends on a single critical catalytic step: the oxidation of tyrosine and L-dopa to dopaquinone by tyrosinase; the remaining reactions that yield melanin are apparently nonenzymatic. Melanin production is increased when another protein from *S. antibioticus* is coexpressed with tyrosinase. Although definitive evidence is not yet available, this second protein may provide a copper-donor function that activates apotyrosinase. Thus, increasing the expression of a cofactor-requiring protein as part of a metabolic engineering scheme may require the engineering of an increased supply of the cofactor as well.

Biodegradation of undesirable compounds can often be accomplished by host enzymes after a heterologous activity provides the initial attack on the target compound or compounds. For example, the expression of *Pseudomonas mendocina* toluene monooxygenase in *E. coli* enabled the efficient degradation of trichlorethylene, a suspected carcinogen and widespread pollutant (*20*). In *E. coli*, degradation can be induced by isopropyl-1-thio-β-D-galactoside or by a temperature shift, rather than by toluene, as occurs in *P. mendocina*. In addition, the engineered *E. coli* has degradation kinetics (no competitive inhibition, as with toluene) and cosubstrate

Table 1. Heterologous activities recruited to alter small metabolite and protein end products. The original metabolite serves as the substrate for the synthesis of the new product through a pathway involving the new intermediate. It is difficult to prove that the inserted activity alone is responsible for an altered phenotype. In each case discussed here, the observed change in cell function is consistent with the expected consequence of the newly installed gene or genes. *A. chrysogenum*, *Acremonium chrysogenum*; GDP, guanosine diphosphate, *F. solani*, *Fusarium solani*; *P. diminuta*, *Pseudomonas diminuta*.

Host organism	Original metabolite	Heterologous enzymes added (source organism)	New product (new intermediate)
E. herbicola (*2, 21, 22*)	2,5-DKG	2,5-DKG reductase (*Corynebacterium*)	2-KLG
A. chrysogenum (*3*)	Cephalosporin C	D-Amino acid oxidase (*F. solani*), cephalosporin acylase (*P. diminuta*)	7ACA [7-β-(5-carboxy-5-oxopentanamido)-cephalosporanic acid]
CHO cells (*4*)	Terminal β-galactosyl residues in *N*-acetyllactosamine sequences	β-Galactoside α2,6-sialyltransferase (rat)	Sialyl α2,6-galactosyl linkages
Mouse L cells (*6*)	Unsubstituted type II, *N*-acetyllactosamine glycoconjugate end groups	GDP-L-fucose: β-D-galactoside 2-α-L-fucosyltransferase (A431 human cell line)	H fucosyl α1→2 galactosy linkages

requirements (glucose instead of toluene) that are superior to those of the native host of this toluene monooxygenase activity. This example illustrates that transfer of a crucial enzyme activity to a different regulatory environment can render that activity useful for biotechnology.

New metabolites arising from the action of cloned heterologous enzymes may also undergo undesirable side reactions. The precursor of 7ACA in engineered *Acremonium chrysogenum*, produced as a consequence of the cloned D-amino acid oxidase, can also react with hydrogen peroxide to give a useless by-product, dramatically reducing 7ACA yield (*3*). Cloned degradation enzymes have led to metabolic dead ends in the sense that the host cannot convert their products further; in some cases these recalcitrant intermediates inactivate key catabolic enzymes (*14*). Other unexpected complications can arise when desired end products are similar to some native metabolite and are converted to another product by host enzymes. After observations of unexpectedly low yields of 2-KLG in a recombinant strain (*2*), it was found that 2-KLG was converted to L-idonic acid by endogenous 2-ketoaldonate reductase (2KR). Cloning, deletion mutagenesis, and homologous recombination of the mutated gene for 2KR into the chromosome were part of several steps undertaken to develop an engineered organism able to accumulate large amounts of 2-KLG (>120 g/liter) (*21, 22*). The present engineered metabolic pathway involving these constructs (Fig. 1) shows complex interactions of enzymes and substrates that were identified, characterized, and engineered in an iterative process.

Perfecting strains by altering nutrient uptake and metabolite flow. Increased growth rates, decreased nutrient demands for cell growth, and higher attainable cell densities have advantages in many different applications. The use of metabolic engineering to realize these objectives has been based on increasing the efficiency of nutrient assimilation, enhancing the efficiency of adenosine triphosphate (ATP) production, and reducing the production of inhibitory metabolic end products. In one of the earliest applications of recombinant DNA to the improvement of the metabolism of the commercial strain, the goal was improvement of the efficiency of carbon conversion into cell mass by *Methylophilus methylotrophus*, a strain developed as an animal feed material. The native route of nitrogen assimilation used by this bacterium is the glutamate synthase pathway, which consumes one ATP per nitrogen incorporated into glutamate. Nitrogen assimilation by means of glutamate dehydrogenase, a process absent from this organism, does not require ATP. In an effort to improve cell yield, glutamate dehydrogenase from *E. coli* was expressed in a glutamate synthase mutant of *M. methylotrophus* (*23*). The efficiency of carbon conversion was increased 4 to 7%.

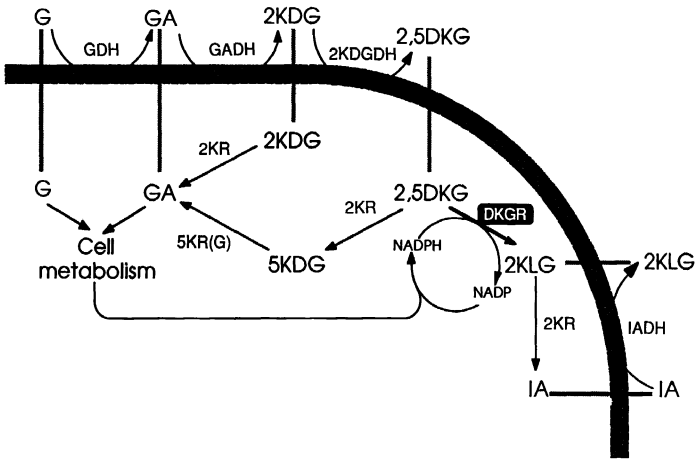

Fig. 1. Summary of the enzymes, intermediates, and by-products encountered in the synthesis of 2-keto-L-gulonic acid (2-KLG) from glucose in genetically engineered *E. herbicola*. The control of cloned (white printing on black) 2,5-DKG reductase activity, which requires the reduced form (NADPH) of nicotinamide adenine dinucleotide phosphate (NADP) supplied by the cell metabolism, directs metabolite flow to 2-KLG. Wide band with dots, cell membrane; straight bars, transport systems. Abbreviations are as follows: GDH, D-glucose dehydrogenase; GADH, D-gluconate dehydrogenase; 2KDGDH, 2-keto-D-gluconate dehydrogenase; DKGR, 2,5-diketo-D-gluconate reductase; IADH, L-idonate dehydrogase; 5KR(G), 5-keto-D-gluconate reductase (D-gluconate–producing); G, D-glucose; GA, D-gluconate; 2KDG, 2-keto-D-gluconate; 2,5KDG, 2,5-diketo-D-gluconate; IA, L-idonate; and 5KDG, 5-keto-D-gluconate. Reprinted by permission from (*21*).

End products of carbon catabolism (acetate, ethanol, and lactate) that inhibit cell growth are produced by bacteria, yeasts, and mammalian cells under conditions of oxygen limitation or carbon source excess. The final optical density of *E. coli* grown under shake-flask aerobic conditions was increased threefold after introduction of a plasmid that expressed pyruvate decarboxylase and alcohol dehydrogenase from *Zymomonas mobilis* (*24*). The former activity, absent in unmodified *E. coli*, redirects catabolite fluxes from pyruvate and results in a shift from acetate production, which strongly inhibits cell growth, to production of ethanol, which is less inhibitory.

Microbial catabolic products such as ethanol, acetone, and butanol are important industrial chemicals. Large increases in ethanol yields from pentose and hexose sugar substrates from *E. coli* (*25, 26*) and *Erwinia chrysanthemi* (*27*) have been achieved by transformation with plasmids that encode pyruvate decarboxylase from *Z. mobilis*, in some cases coexpressed with *Z. mobilis* alcohol dehydrogenase. The *E. coli* so engineered have the potential practical advantages of rapid and efficient conversion of several sugars found in biomass (*26*).

α-Acetohydroxy acids, synthesized during fermentation by brewers' yeast, leak into the medium where spontaneous hydroxylation produces diacetyl, which has an undesirable flavor. On the basis of suggestions that the time required for beer lagering is determined by the time required for the enzymatic reduction of diacetyl by the yeast, genes for the enzyme α-acetolactate decarboxylase (α-ALDC) were cloned from *Klebsiella terrigena* or *Enterobacter aerogenes*, fused to yeast promoters, and inserted into *Saccharomyces cerevisiae* on multicopy plasmids. This enzyme converts α-acetolactate to acetoin, rather than diacetyl; acetoin influences flavor only at relatively high concentrations. Pilot brewing studies with these engineered strains that express α-ALDC yielded beer of quality equal to that produced by controls, but in a process time of 2 weeks, as compared to 5 weeks for the conventional process. The lagering step could be omitted when the recombinant brewers' yeasts were used because of low diacetyl production by these organisms (*28*).

Enabling a cell to utilize alternative materials as nourishment is another capability of metabolic engineering. In order to produce microbial surfactants from industrial waste raw materials, *E. coli* β-galactosidase and lactose permease were stably integrated into the chromosome of two *Pseudomonas aeroginosa* strains. These recombinant strains synthesize biosurfactants when grown in lactose and whey-based minimal media (*29*).

Yeast ornithine decarboxylase was cloned and expressed in cultured roots of *Nicotiana rustica* in order to direct a greater metabolite flux from ornithine to putrescine, a precursor of nicotine (*30*). Some clones showed approximately two times as much nicotine accumulation as the controls. Rearrangement of the native fluxes in the hyoscyamine-rich *Atropa belladonna* was motivated by greater commercial demand for scopolamine, the 6,7-epoxide of hyoscyamine. Expression of *Hyoscyamus niger* hyoscyamine 6β-hydroxylase in an *A. belladonna* hairy-root clone produced three to ten times as much scopolamine as did wild-type clones (*31*).

Systematic genetic manipulation of protein processing pathways has proven effective in increasing the quantity of active protein recovered. Overexpression of the *E. coli* chaperone proteins GroES and GroEL provided a five- to tenfold increase in assembled cyanobacterial Rubisco (D-ribulose-1,5-bisphosphate carboxylase/oxygenase) enzyme coexpressed in *E. coli*. In vitro studies of interactions among Rubisco and the GroE proteins implicate Mg^{2+}-ATP as a requirement for assembly (*32*). The failure to achieve assembly of Rubisco from higher plants in altered *E. coli* signals future challenges in the transfer of heterologous protein processing pathways (*33*). Other challenges include genetic manipulations of processing pathways of bacteria that alter the solubility of recombinant proteins (*34*). In addition, opportunities exist for extending such strategies to eukaryotic hosts (*35*).

Transfer of promising natural motifs: Vitreoscilla *hemoglobin*. Because of the constant drive toward maximum cell densities to maximize volumetric productivity, growth and product synthesis in many industrial processes are limited by oxygen supply. The Gram-negative aerobic bacterium *Vitreoscilla*, which lives in poorly aerated environments, synthesizes increased quantities of a hemoglobin molecule in oxygen-limited cultures (*36*). Although the function of this protein in its natural host has not been established, this pattern of regulation of expression, combined with the oxygen-binding and release characteristics of the protein, suggest a possible beneficial physiological activity in poorly oxygenated environments.

Motivated by this hypothesis and the premise that this beneficial function might be genetically transferred to industrial microorganisms, the gene for *Vitreoscilla* hemoglobin (VHb) was cloned and expressed in *E. coli* (*37*). *Escherichia coli* that carried a single copy of this gene integrated in the chromosome synthesized total cell protein more rapidly than an isogenic wild-type strain in oxygen-limited cultivations (Fig. 2), a response attributed to an increased efficiency of net ATP synthesis in the hemoglobin-expressing strain (*38*). Facilitation of oxygen transfer to the respiratory center (*39*) and modification of some aspect of cellular redox chemistry (*38*) have been suggested as contributing mechanisms for these phenomena. Coexpression of VHb increases the expression of cloned β-galactosidase, chloramphenicol acetyltransferase (CAT) (*38*), and α-amylase (*40*) by 1.5- to 3.3-fold relative to controls in oxygen-limited *E. coli* cultures, probably as a result of enhanced net ATP synthesis.

Aeration of bioreactors used in the synthesis of antibiotics is frequently complicated by the thick broths that result from growth of filamentous fungi and *Streptomyces*. Success of the strategy for enhancing aerobic metabolism in other bacteria prompted cloning and intracellular expression of VHb in two different *Streptomyces* species (*41*). *Streptomyces lividans* with a multicopy hemoglobin expression plasmid achieved final cell densities up to 54% greater than the untransformed host in shake-flask cultivations. The presence of cloned intracellular VHb in *S. coelicolor* markedly increased secondary metabolite accumulation, without affecting cell growth relative to a control strain containing a mutated VHb gene (Fig. 2).

These examples suggest a general genetic strategy for addressing stresses and corresponding productivity limitations encountered in bioprocessing: after identifying a response in nature to a similar stress (most likely involving a different organism), genes that specify that response can be transferred to the organism of choice.

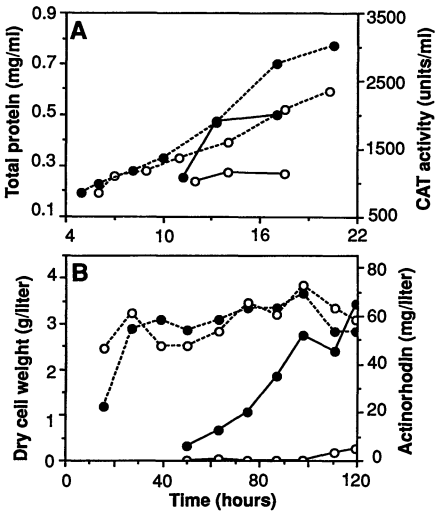

Fig. 2. (**A**) Time trajectories of total *E. coli* cell protein (dashed lines) and cloned CAT activity (solid lines) in the wild type (open symbols) and an engineered host that expresses VHb from a single gene copy integrated in the chromosome (closed symbols) in oxygen-limited fed-batch fermentations [reprinted from (*38*)]. (**B**) Cell densities (dashed lines) and concentration of actinorhodin in the medium (solid lines) in batch fermentations of *S. coelicolor*. Closed symbols are from a transformant expressing active VHb; open symbols are from a control transformant not expressing VHb [reprinted from (*41*)].

Redirecting Metabolite Flow

Typically the route of reactions to a desired product passes several forks where intermediates can enter alternative pathways. At such bifurcations of metabolite flow, a common resource—for example, substrate, enzyme, transport system, or ribosome—contributes to two or more parallel processes. Maximizing product formation requires that the desired route at each fork be made a priority and that traffic in alternative pathways be minimized to the extent possible without decreasing cell viability.

Directing traffic toward the desired branch. Amplification of the activity initiating a desired process at a fork in a metabolic flow is a common strategy of metabolic engineering. Whereas isolation of mutant enzymes that are desensitized to feedback repression was achieved with classical methods, such mutants may now be obtained more rapidly with the use of cloned genes. This approach also avoids the complication of uncharacterized additional mutations that are often obtained with classical, whole-cell mutagenesis.

The past decade has seen a new generation of strain improvements in amino acid–producing coryneform bacteria with metabolic engineering (also called molecular breeding) (*42, 43*). Central to the success achieved was the development of new vectors and transformation procedures.

Genetic engineering of improved threonine production by *Brevibacterium lactofermentum* illustrates some of the strategies useful in redirecting metabolite flow to the desired product. Figure 3 presents an abbreviated diagram of the reactions involved in the synthesis of the aspartate family of amino acids and a few key reactions that feed into the synthesis pathway for this family. Homoserine dehydrogenase (HD) was amplified by cloning and transformation into a threonine- and lysine-producing mutant (designated M-15). This mutant organism was selected for its lack of feedback inhibition of aspartokinase by threonine and lysine and of HD by threonine (*44*). The respective final concentrations of threonine, homoserine, and lysine from benchtop fermentations were 25.0, 2.8, and 1.1 g/liter for the recombinant strain compared to 17.5, 0.5, and 12.1 for M-15. Subsequent further engineering to coexpress cloned homoserine kinase (HK) with HD further increased the final threonine concentration to 33 g/liter and reduced homoserine and lysine levels, relative to the strain with cloned HD alone (*45*). In another study, the coryneform gene for HD was mutagenized to eliminate feedback inhibition by threonine. Introduction of this mutated HD gene into a lysine producer shifted the final lysine concentration from 65 g/liter to 4 g/liter and the final threonine concentration from 0 g/liter to 52 g/liter (*43*). Threonine production by M-15 was increased 12% by the expression of cloned phosphoenolpyruvate (PEP) carboxylase (PEPCase) (*46*). This manipulation was motivated by

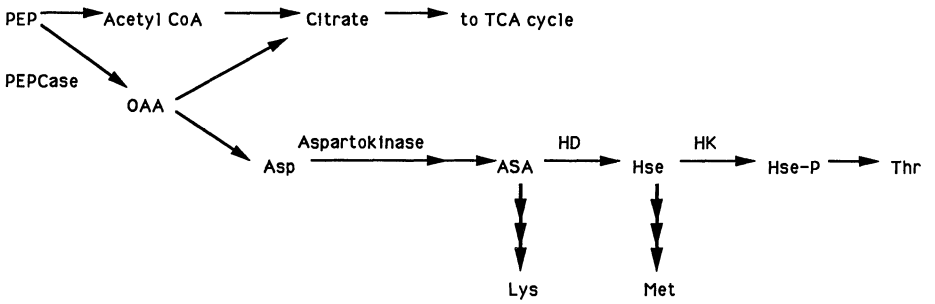

Fig. 3. Pathways of the biosynthesis of the aspartate family of amino acids. Metabolite abbreviations are as follows: acetyl CoA, acetyl coenzyme A; TCA, tricarboxylic acid cycle; Asp, aspartate; ASA, aspartate semialdehyde; Hse, homoserine; Lys, lysine; Met, methionine; Hse-P, *O*-phosphohomoserine; and Thr, threonine.

the researchers' desire to increase oxaloacetate (OAA) production and thereby to increase carbon flow into amino acid production. Further improvements in rates of amino acid synthesis and yields will depend on a better understanding of mechanisms of regulation of gene expression and metabolite flow in these bacteria (*47, 48*).

Because of metabolic engineering, *E. coli* has become an industrially important producer of amino acids. Transformation by multicopy plasmids that contain tryptophan (*49*) and threonine (*50*) biosynthetic genes have increased production of these amino acids. A project to engineer phenylalanine production in *E. coli* showed that overexpression of some genes in the phenylalanine biosynthetic pathway could cause a decrease in phenylalanine production and that inducible excision vector technology can be used to manipulate the biosynthesis of tyrosine, an inhibitor of the desired pathway (*51*). An intermediate in a metabolic pathway can be overproduced by combining a mutation that blocks that intermediate's use by the cell and by genetic augmentation of precursor flow into that pathway; although this concept has been used extensively in classical genetic production of organisms that are amino acid overproducers, it can be implemented in other contexts by metabolic engineering (*52*).

The gene *eryF* in *Saccharopolyspora erythrae* encodes the first enzyme in the pathway from 6-deoxyerythronolide B to the antibiotic erythromycin. After the targeted disruption of this gene using an integrative plasmid, 6-deoxyerythronolide B was converted to an erythromycin derivative that is more stable at the low pH of the stomach (*53*).

Because enzyme activities involved in secondary metabolite production are regulated at both the gene and protein levels, identifying genetic changes that accelerate synthesis of these metabolites is challenging. One successful strategy is based on measuring the biosynthetic pathway intermediate concentrations in the growth medium. Relatively high extracellular concentrations of the intermediate penicillin N suggested that the activity that converts this intermediate to cephalosporin C (encoded in *cef*EF) may limit the rate of the overall pathway (*54*). Thus, expression of *cef*EF was elevated through increased gene dosage in a production strain of *Cephalosporium acremonium*. This recombinant fungus exhibited a 15-fold reduction in penicillin N production and an increase of ~15% in cephalosporin C production.

Routing through protein processing pathways has also been altered by manipulation of host genes. The overproduction lethality commonly observed with exported β-galactosidase fusion proteins in *E. coli* is suppressed by the overproduction of *E. coli prl*F (*55*), and the expression of *E. coli* DnaK enables export of *lacZ*-hybrid proteins that are otherwise confined to the cytoplasm (*56*). An NH_2-terminal methionine often differentiates cloned polypeptides synthesized in *E. coli* from their native human counterparts. Coexpression of cloned *E. coli* methionine aminopeptidase with human interleukin-2 in *E. coli* has substantially reduced the fraction of product with methionine at its NH_2-terminus (*57*). Observation of large quantities of a variant of human tissue plasminogen activator (tPA) associated with GRP78 in the rough endoplasmic reticulum of CHO cells suggested that GRP78 binding was a rate-limiting step in tPA secretion. (GRP78 is the 78-kD glucose-regulated protein, one of the stress-response proteins.) Coexpression of antisense GRP78 message resulted in smaller quantities of GRP78 and faster tPA secretion (*58*).

A quantitative study was conducted of *S. cerevisiae* isolates that contain different numbers of the phosphoglycerate kinase (PGK) gene (*PGK1*). In some cases a 10 to 15% increase in PGK activity gave rise to a higher (30%) overall cell mass yield when the yeast were grown on glucose. However, in another construct that contained more copies of *PGK1*, yield was depressed by 40% (*59*). These results show the importance of fine-tuning the amount of gene or enzyme amplification to achieve the desired benefit.

Reducing competition for a limiting resource. Computer simulations with a detailed single-cell model suggested that competition between vector- and host-encoded messages for a common pool of ribosomes could limit cloned gene expression (*60*), a prediction consistent with experimental observations of reduced ribosome content in recombinant *E. coli* that contain more plasmids per cell (*61, 62*). By expressing a cloned mutant 16*S* ribosomal RNA, a population of ribosomes is created that is specialized for expression of only those cloned gene transcripts that bear a corresponding mutation in the Shine-Dalgarno sequence (*63*). Large amounts of messenger RNA transcribed from the cloned gene will not interact with the primary population of native ribosomes; therefore the expression of cloned genes will not interfere with the simultaneous expression of host cell genes important for protein synthesis. With this approach, expression of cloned β-galactosidase was increased by 35%. A 30% reduction in specific growth rate occurred after expression of the specialized 16*S* RNA; however, no growth rate reduction was observed on induction of cloned β-galactosidase synthesis (*64*).

Revising metabolic regulation. Positive control genes have been found so far in the biosynthetic gene clusters for actinorhodin, bialaphos, streptomycin, and undecylprodigiosin, which are all secondary metabolites produced by *Streptomyces* species. Cloning additional copies of an activator gene in the wild-type host can substantially increase antibiotic production, as indicated for undecylprodigiosin (*8*).

Modification of the regulation of expression of maltose permease and maltase is the basis for a metabolically engineered baker's yeast intended to reduce the time for leavening of sweet doughs (*65*). Glucose normally represses expression of these proteins and thereby blocks simultaneous maltose utilization. The engineered strain uses constitutive yeast promoters for these enzymes to enable simultaneous uptake and catabolism of both sugars. This is one of few examples in which a transport system has been manipulated.

A genetic engineering strategy to stimulate CO_2 production by bakers' yeast seeks to consume ATP (*66*). This could relieve ATP inhibition of phosphofructokinase (PFK) and pyruvate kinase, two regulatory enzymes in sugar catabolism. A futile cycle with PFK

was created by expressing cloned yeast fructose 1,6-bisphophatase (FBPase) from a yeast glycerophosphate dehydrogenase promoter that is induced by glucose; FBPase is not normally expressed in the presence of high glucose concentrations. This yeast strain produced 20 to 25% more CO_2 than the wild type.

In order to construct a pathway for processing the pollutant 4-ethylbenzoate, it was necessary to alter the regulation of the alkylbenzoate degradation pathway encoded on the *Pseudomonas* TOL plasmid (*67*). Originally 4-ethylbenzoate did not induce transcription of the crucial *meta* operon. Mutations were introduced into the positive regulator of *Pm* (the promoter of the *meta* operon) that enabled *Pm* activation by 4-ethylbenzoate.

Completing the Metabolic Engineering Cycle: Potentials and Perils of Rational Design

The iterative cycle of genetic modification, analysis of the metabolic consequences of this change, and choice of the next genetic modification has been successfully implemented in a few instances with promising results. Contemporary concepts and technologies for each function in this cycle are summarized next.

Cloning in industrial strains. The lack of suitable vectors and methods for the introduction of exogenous DNA limits the application of metabolic engineering in many important industrial organisms (*68*). Electroporation and conjugation have proven useful in introducing DNA into diverse organisms. The stable propagation of cloned genes remains problematic even in such a well-studied system as *Bacillus subtilis* and is apparently a result of the error-prone rolling-circle replication mechanism used by many plasmids in Gram-positive bacteria. Extensive rearrangements and deletions of both chromosomal and plasmid DNA occur frequently in some species, complicating their systematic manipulation. Restriction (cleavage) of heterologous DNA is a limitation in efficient engineering of many cells of practical interest.

The hurdles to be surmounted in developing the necessary genetic tools are illustrated in research that is establishing a foundation for engineering the complex catabolic metabolism of *Clostridium acetobutylicum*. This bacterium is the basis for the biological production of the industrial chemicals acetone and butanol. Efficient transformation of this organism required optimization of an electroporation protocol, and it was discovered that, because of a clostridial restriction enzyme system, *E. coli* is not a suitable organism for the cloning of clostridial DNAs, whereas *B. subtilis* is (*69*). Technology for chromosomal integration should soon follow, as several *C. acetobutylicum* genes have now been cloned.

Dissecting physiological responses. For the most effective design of a subsequent genetic manipulation, it is useful to know the concentrations of intracellular proteins and metabolites. The concentrations of many cellular proteins can be determined in principle from two-dimensional gel electrophoresis, but data bases are necessary to identify individual proteins (*70*). In vitro assays of changes in activities of key enzymes have been widely applied.

A broad spectrum of analytical methods can be applied for determining metabolite concentrations. The measurement of concentrations and in some cases of fluxes in particular pathways of interest can often be aided by the application of isotopically labeled precursors. For example, with the use of labeled acetate and glutamate, along with quasi-steady state conservation equations for intracellular metabolites, the velocities for carbon flow through *E. coli* growing on acetate can be determined (*71*).

Nuclear magnetic resonance (NMR) spectroscopy has been applied to estimate metabolite concentrations in whole cells, cell extracts, and growth media (*72*). For example, ^{31}P NMR measurements of *S. cerevisiae* cells converting glucose to end products under anaerobic conditions, in concert with a methodology for extracting individual component information from the sugar phosphate portion of the spectrum, provided estimates of metabolite concentrations that were essential for analysis of the pathway (*73*). The time and instrumentation required to evaluate metabolite concentrations presently limits rational metabolic engineering.

Design principles and cell models: Coping with complexity and coupling. No universal principles have emerged from metabolic engineering research to guide the choice of the next useful genetic alteration. Attempts to address these problems with artificial intelligence have shown that there is no substitute for knowledge of the pathways involved, their regulation, and their kinetics. Some useful approaches include measurements of intermediate concentrations to indicate possible rate-determining reactions, genetic transfer of natural stress response motifs, and applications of organisms that can be used over wide ranges of temperature and pH (*74*).

Alternatively, if a mathematical description of the system is available, sensitivity analysis can be applied to calculate the expected response of the pathway to changes in the individual steps or pathway segments. An advantage of such an approach is its simultaneous determination of the sensitivities of the desired flux to many different participating reactions, permitting the identification of situations in which several genetic modifications in concert are required to achieve a desired response.

A body of theoretical developments known as metabolic control theory is well suited to the requirements of rational metabolic engineering (*75*). A central result provided by this theory is a sensitivity calculation that provides the flux control coefficients, defined as the fractional changes of flux expected for a unit fractional change in the amount of each enzyme participating in a given pathway. In addition, it is possible to evaluate the sensitivity of flux through the pathway to individual parameters in kinetic expressions for each of the enzymes, thereby providing guidance for useful protein engineering to accelerate the pathway. Analysis of several simple examples that involve unbranched sequences of reactions showed that sole control of flux by any single step (in other words, the existence of a single, rate-limiting step) is in general not expected. Instead, the flux through the pathway is usually influenced by the activities of several individual steps. This result, augmented by specific model calculations outlined below, provides motivation for sequential improvement of metabolic pathways.

In one of the few cases in which detailed kinetic expressions for each step in the reaction network, as well as the concentrations of all substrates and effectors, are known or estimated, flux control coefficients were determined for nongrowing yeast converting glucose to ethanol and other end products (*76*). Several general points are suggested by this investigation. First, the sensitivity of pathway flux to individual step changes depends on the environment in which the cell is grown. It is therefore important to carry out modeling and measurement under the expected industrial conditions. Second, flux control can be extraordinarily sensitive to some parameters such as intracellular pH. Third, in any system with interacting pathways (and it is difficult to envision any case where this does not occur), the most general version of metabolic control theory must be used (*77*). Such coupling is extensive in the usual case of growing cells, where the pathway of interest interacts with all of the other metabolic processes in the cell. A strategy for accomplishing flux control coefficient calculations in this situation has been presented (*78*). Finally, calculations with the kinetic model formulated in a yeast biocatalysis study indicate that amplification of the activity of one enzyme results in a shift of flux control to other steps in the pathway. New theory that presumes linear approximations for all rate expressions provides estimates of flux control coefficients,

without requiring knowledge of kinetics and using time-resolved metabolite concentration measurements instead (*79*); the practical merits of this approach have not yet been evaluated.

Both sensitivity to small changes and simulations of responses to large changes in intracellular activities can be calculated from a detailed and reliable mathematical model of the cell. Large quantities of biological information have been integrated into computer models for single cells (*60, 80*) and several molecular control systems (*81*). These have successfully simulated consequences of several genetic and environmental changes. Although useful initial directions for genetic improvement have been suggested by such models, they have not yet been used as the central tool in an iterative metabolic engineering study. In spite of their obvious limitations, these mathematical structures are the only way that the net consequences of simultaneous, coupled, and often counteracting processes can be simulated and evaluated consistently and quantitatively.

Minimizing response cascades. Unanticipated cell responses to a genetic modification may complicate rational practice of the metabolic engineering cycle. Introduction of a cloning vector alone may result in a large cascade of metabolic changes, many of which are difficult to anticipate. For example, the introduction of multicopy plasmids into *E. coli*, even without overexpression of a cloned product, has been shown to cause substantial changes in growth rates, cell cycle regulation, amounts of many individual proteins, glucose uptake, and carbon catabolite production rates (*82*). Transformation of yeast with multicopy plasmids can introduce lesions that persist after the plasmids have been eliminated from the transformants (*83*). Different mammalian cell clones transfected by the same vector often exhibit different growth rates and cell sizes. Therefore, introduction of a desired genetic change should be carefully configured to minimize perturbation of the host, using the lowest gene dosage and lowest expression level that give the desired result. The apparatus used for selecting the modified strain should also be carefully considered. For example, ampicillin resistance used for the maintenance of many laboratory recombinant *E. coli* strains is provided by cloned β-lactamase. This precursor must be processed at the cytoplasmic membrane in competition with host cell preprotein, which often results in a major physiological disruption.

Even if the genetic manipulation is accomplished in a relatively well-controlled fashion, the regulatory apparatus of the cell at both the gene and protein levels may confound the intended change or even alter cellular activities. For example, amplification of citrate synthase in *E. coli* did not increase the flux through the citric acid cycle because of a compensating modulation of the activity of isocitrate dehydrogenase (*84*). Expression of even low concentrations of unnatural proteins can activate stress responses, influencing many cell functions (*85*). Anticipating and accounting for such regulatory responses to genetic intrusions are fundamental challenges for the future.

REFERENCES AND NOTES

1. K. M. Ulmer, *Science* **219**, 666 (1983).
2. S. Anderson *et al.*, *ibid.* **230**, 144 (1985).
3. T. Isogai *et al.*, *Biotechnology* **9**, 188 (1991).
4. E. U. Lee, J. Roth, J. C. Paulson, *J. Biol. Chem.* **264**, 13848 (1989).
5. M. Takeuchi *et al.*, *ibid.* **263**, 3657 (1988).
6. L. K. Ernst *et al.*, *ibid.* **264**, 3436 (1989).
7. F. Malpartida, J. Niemi, R. Navarrete, D. A. Hopwood, *Gene (Amsterdam)* **93**, 91 (1990).
8. F. Malpartida and D. A. Hopwood, *Nature* **309**, 462 (1984).
9. R. Stanzak, P. Matsushima, R. H. Baltz, R. N. Rao, *Biotechnology* **4**, 229 (1986).
10. D. J. Smith, M. K. R. Burnham, J. Edwards, A. J. Earl, G. Turner, *ibid.* **8**, 39 (1990).
11. O. P. Peoples and A. J. Sinskey, in *Novel Biodegradable Microbial Polymers,* E. A. Dawes, Ed. (Kluwer Academic, Dordrecht, Netherlands, 1990), pp. 191–202.
12. O. P. Peoples and A. J. Sinskey, *J. Biol. Chem.* **264**, 15298 (1989).
13. F. Rojo, D. H. Pieper, K.-H. Engesser, H.-J. Knackmuss, K. N. Timmis, *Science* **238**, 1395 (1987).
14. K. N. Timmis, F. Rojo, J. L. Ramos, in *Biotechnology and Biodegradation*, D. Kamely, A. Chakrabarty, G. S. Omenn, Eds. (Portfolio, The Woodlands, Texas, 1990), pp. 61–80.
15. D. A. Hopwood *et al.*, *Nature* **314**, 642 (1985); S. Omura, H. Ikeda, F. Malpartida, H. M. Kieser, D. A. Hopwood, *Antimicrob. Agents Chemother.* **29**, 13 (1986).
16. J. B. McAlpine *et al.*, *J. Antibiot.* **40**, 1115 (1987).
17. J. K. Epp, M. L. B. Huber, J. R. Turner, T. Goodson, B. E. Schoner, *Gene (Amsterdam)* **85**, 293 (1989).
18. B. D. Ensley *et al.*, *Science* **222**, 167 (1983).
19. G. della-Cioppa, S. J. Garger, G. G. Sverlow, T. H. Turpen, L. K. Grill, *Biotechnology* **8**, 634 (1990).
20. R. B. Winter, K.-M. Yen, B. D. Ensley, *ibid.* **7**, 282 (1989).
21. R. A. Lazarus *et al.*, in *Biocatalysis*, D. Abramowitz, Ed. (Van Nostrand Reinhold, New York, 1990), pp. 136–155.
22. R. A. Lazarus *et al.*, in *Proceedings of the Sixth International Symposium on Genetics of Industrial Microorganisms*, H. Heslot *et al.*, Eds. (Société Française de Microbiologie, Strasbourg, 1990), vol. 2, pp. 1073–1082.
23. J. D. Windass *et al.*, *Nature* **287**, 396 (1980).
24. L. O. Ingram and T. Conway, *Appl. Environ. Microbiol.* **54**, 397 (1988).
25. B. Bräu and H. Sahm, *Arch. Microbiol.* **144**, 296 (1986).
26. F. Alterthum and L. O. Ingram, *ibid.* **55**, 1943 (1989).
27. J. S. Tolan and R. K. Finn, *Appl. Environ. Microbiol.* **53**, 2039 (1987).
28. M.-L. Suihko, K. Blomqvist, M. Penttilä, R. Gisler, J. Knowles, *J. Biotechnol.* **14**, 285 (1990).
29. A. K. Koch, J. Reiser, O. Käppeli, A. Fiechter, *Biotechnology* 6, 1335 (1988).
30. J. D. Hamill *et al.*, *Plant Mol. Biol.* **15**, 27 (1990).
31. T. Hashimoto, D. J. Yun, Y. Yamada, personal communication.
32. A. A. Gatenby, P. V. Viitanen, G. H. Lorimer, *Trends Biotechnol.* **8**, 354 (1990).
33. J. E. Rothman, *Cell* **59**, 591 (1989).
34. C. H. Schein, *Biotechnology* **7**, 1141 (1989).
35. R. J. Wise *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **87**, 9378 (1990).
36. S. J. Boerman and D. A. Webster, *J. Gen. Appl. Microbiol.* **28**, 35 (1982).
37. C. Khosla and J. E. Bailey, *Nature* **331**, 633 (1988); *Mol. Gen. Genet.* **214**, 158 (1988).
38. C. Khosla, J. Curtis, J. DeModena, U. Rinas, J. E. Bailey, *Biotechnology* **8**, 849 (1990).
39. C. Khosla and J. E. Bailey, *J. Mol. Biol.* **171**, 5995 (1989).
40. M. Khosravi, D. A. Webster, B. C. Stark, *Plasmid* **24**, 190 (1990).
41. S. K. Magnolo *et al.*, *Biotechnology* **9**, 473 (1991).
42. K. Sano, in *Proceedings of the Fourth European Congress on Biotechnology*, O. M. Neijssel, R. R. van der Meer, K. Ch. A. M. Luyben, Eds. (Elsevier Science, Amsterdam, 1987), vol. 4, pp. 735–747.
43. R. Katsumata *et al.*, *ibid.*, pp. 767–776.
44. S. Nakamori *et al.*, *Agric. Biol. Chem.* **51**, 87 (1987).
45. Y. Morinaga *et al.*, *ibid.*, p. 93.
46. K. Sano, K. Ito, K. Miwa, S. Nakamori, *ibid.*, p. 597.
47. M. T. Follettie, O. P. Peoples, J. A. C. Archer, A. J. Sinskey, in *Proceedings of the Sixth International Symposium on Genetics of Industrial Microorganisms*, H. Heslot *et al.*, Eds. (Société Française de Microbiologie, Strasbourg, 1990), vol. 1, pp. 315–325.
48. G. N. Stephanopoulos, *Science* **252**, 1675 (1991).
49. S. Aiba, T. Imanaka, H. Tsunekawa, *Biotechnol. Lett.* **2**, 525 (1980).
50. K. Miwa *et al.*, *Agric. Biol. Chem.* **47**, 2329 (1983).
51. K. Backman *et al.*, *Ann. N.Y. Acad. Sci.* **589**, 16 (1990).
52. K. M. Draths and J. W. Frost, *J. Am. Chem. Soc.* **112**, 1657 (1990).
53. J. M. Weber, J. O. Leung, S. J. Swanson, K. B. Idler, J. B. McAlpine, *Science* **252**, 114 (1991).
54. P. L. Skatrud *et al.*, *Biotechnology* **7**, 477 (1989).
55. D. R. Kiino, G. J. Phillips, T. J. Silhavy, *J. Bacteriol.* **172**, 185 (1990).
56. G. J. Phillips and T. J. Silhavy, *Nature* **344**, 882 (1990).
57. A. Ben-Bassat *et al.*, *J. Bacteriol.* **169**, 751 (1987).
58. A. J. Dorner, M. G. Krane, R. J. Kaufman, *Mol. Cell. Biol.* **8**, 4063 (1988).
59. P. C. van der Aar *et al.*, *Appl. Microbiol. Biotechnol.* **32**, 577 (1990).
60. S. W. Peretti and J. E. Bailey, *Biotechnol. Bioeng.* **29**, 316 (1987).
61. T. K. Wood and S. W. Peretti, *ibid.* **36**, 865 (1990).
62. S. Birnbaum and J. E. Bailey, *ibid.* **37**, 736 (1991).
63. A. Hui and H. A. de Boer, *Proc. Natl. Acad. Sci. U.S.A.* **84**, 4762 (1987).
64. T. K. Wood and S.W. Peretti, *Biotechnol. Bioeng.*, in press.
65. J. Hodgson, *Biotechnology* **8**, 281 (1990).
66. D. T. Rogers and E. Hiller, in *Proceedings of the Sixth International Symposium on Genetics of Industrial Microorganisms*,H. Heslot *et al.*, Eds. (Societé Française de Microbiologie, Strasbourg, 1990), vol. 1, pp. 533–544.
67. J. L. Ramos *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **83**, 8467 (1986).
68. H. Heslot *et al.*, Eds., *Proceedings of the Sixth International Symposium on Genetics of Industrial Microorganisms*, H. Heslot *et al.*, Eds. (Société Française de Microbiologie, Strasbourg, 1990); A. L. Demain and N. A. Solomon, Eds., *Biology of Industrial Microorganisms* (Benjamin-Cummings, Menlo Park, CA, 1985); C. L. Hershberger, S. W. Queener, G. Hegeman, Eds., *Genetics and Molecular Biology of Industrial Microorganisms* (American Society for Microbiology, Washington, DC, 1989).
69. J. W. Cary, D. J. Petersen, G. N. Bennett, E. T. Papoutsakis, *Ann. N.Y. Acad. Sci.* **589**, 67 (1990).
70. T. A. Phillips, V. Vaughn, P. L. Bloch, F. C. Neidhardt, in *Escherichia coli* and *Salmonella typhimurium*, F. C. Neidhardt *et al.*, Eds. (American Society for Microbiology, Washington, DC, 1987), pp. 919–966.

71. K. Walsh and D. E. Koshland, Jr., *J. Biol. Chem.* **259**, 9646 (1984).
72. D. G. Gadian, *Nuclear Magnetic Resonance and Its Applications to Living Systems* (Oxford Univ. Press, Oxford, 1982); E. J. Fernandez and D. S. Clark, *Enzyme Microb. Technol.* **9**, 256 (1987).
73. J. V. Shanks and J. E. Bailey, *Biotechnol. Bioeng.* **32**, 1138 (1988).
74. T. D. Brock, Ed., *Thermophiles: General, Molecular and Applied Microbiology* (Wiley-Interscience, New York, 1986).
75. H. Kacser and J. A. Burns, *Symp. Soc. Exp. Biol.* **27**, 65 (1973); R. Heinrich, S. M. Rapoport, T. A. Rapoport, *Prog. Biophys. Mol. Biol.* **32**, 1 (1977).
76. J. L. Galazzo and J. E. Bailey, *Enzyme Microb. Technol.* **12**, 162 (1990); *ibid.* **13**, 363 (1991).
77. C. Reder, *J. Theor. Biol.* **135**, 175 (1988).
78. P. M. Schlosser and J. E. Bailey, *Math. Biosci.* **100**, 87 (1990).
79. J. P. Delgado and J. Liao, *Biotechnol. Prog.* **7**, 15 (1991).
80. M. M. Domach *et al.*, *ibid.* **26**, 203 (1984).
81. S. B. Lee and J. E. Bailey, *Plasmid* **11**, 151 (1984); *ibid.*, p. 166; *Biotechnol. Bioeng.* **26**, 66 (1984); *ibid.*, p. 1372; *ibid.*, p. 1283.
82. J. E. Bailey, N. Da Silva, S. M. Peretti, J.-H. Seo, F. Srienc, *Ann. N.Y. Acad. Sci.* **469**, 194 (1986).
83. N. Danhash, D. C. J. Gardner, S. G. Oliver, *Biotechnology* **9**, 179 (1991).
84. K. Walsh, M. Schena, A. J. Flint, D. E. Koshland, Jr., *Biochem. Soc. Symp.* **54**, 183 (1987).
85. M. Kosinski, U. Rinas, J. E. Bailey, unpublished data.
86. Supported by NSF and by the Catalysis and Biocatalysis Program of the Advanced Industrial Concepts Division of the U.S. Department of Energy. Advice and assistance in the preparation of this paper were provided by D. S. Clark, B. D. Ensley, T. Hashimoto, R. J. Kaufman, C. Khosla, R. A. Lazarus, A. J. Sinskey, K. N. Timmis, and colleagues at the California Institute of Technology.

Network Rigidity and Metabolic Engineering in Metabolite Overproduction

GREGORY STEPHANOPOULOS AND JOSEPH J. VALLINO*

In order to enhance the yield and productivity of metabolite production, researchers have focused almost exclusively on enzyme amplification or other modifications of the product pathway. However, overproduction of many metabolites requires significant redirection of flux distributions in the primary metabolism, which may not readily occur following product deregulation because metabolic pathways have evolved to exhibit control architectures that resist flux alterations at branch points. This problem can be addressed through the use of some general concepts of metabolic rigidity, which include a means for identifying and removing rigid branch points within an experimental framework.

ALL ORGANISMS USE PRIMARY METABOLIC PATHWAYS TO supply precursor metabolites and energy to anabolic pathways that synthesize cellular constituents that are necessary for growth and maintenance. In many industrial strains of microorganisms (as well as tissue and plant cultures), these anabolic pathways have been exploited for the overproduction of compounds (such as amino and nucleic acids, antibiotics, vitamins, enzymes, and proteins) that cannot be synthetically produced or for which it is not economical to do so. In general, a particular metabolite is overproduced by deregulating the pathway directly associated with the synthesis of that metabolite, or, more recently, by transforming a robust host organism (typically *Escherichia coli*) with the genes that encode for the synthesis of the desired product (*1, 2*). This approach, however, does not necessarily result in high product yields (defined as the moles of product formed per mole of substrate consumed) since carbon flux distributions at key branch points (nodes) in the primary metabolism [such as glycolysis, tricarboxylic acid (TCA) cycle, and pentose phosphate pathway] must often be radically redirected from the flux distributions that are normally associated with balanced growth. Such metabolic flux alterations are often directly opposed by mechanisms for controlling enzyme activity that have evolved to maintain flux distributions that are optimal for growth. We refer to this inherent resistance to flux alterations as metabolic or network rigidity and to the genetic modifications of specific nodes in the primary metabolism for the purpose of enhancing yield and productivity as metabolic engineering (*3*). Although genetic manipulations can now be readily performed, there are relatively few accounts of successful metabolic flux alterations because of the complex, nonlinear nature of the metabolic control architectures.

The nature and types of metabolic rigidity are reviewed in this article along with methods to identify and possibly circumvent such undesirable nodal controls. The overproduction of lysine by *Corynebacterium glutamicum* [and related strains (*4*)] is used as a vehicle to illustrate key points because of: (i) the lack of compartmentalization in bacteria; (ii) the need for significant flux alterations to optimize lysine biosynthesis; and (iii) the apparent marginal success of mutation-selection (*5, 6*) or genetic engineering (7) techniques used to that end. The concepts, however, are of general value, and the methods are applicable to other metabolic products as well.

Basis of Metabolic Rigidity

Although intracellular metabolite concentrations can fluctuate during growth, on average, the distributions of the major cellular groups (proteins, RNA, DNA, lipids, and so forth) remain relatively proportional to one another throughout balanced growth (*8*). In fact, metabolites and energy required to synthesize an *E. coli* cell have been calculated on the basis of its known composition (*9*). In order to preserve this regularity in cellular composition, the primary metabolism has evolved coordination of pathway control, such that building-block metabolites, energy [such as adenosine triphosphate (ATP)], and biosynthetic reducing power [such as nicotinamide adenine dinucleotide phosphate (NADPH)] are synthesized in approximate stoichiometric ratios during balanced growth. Al-

The authors are in the Department of Chemical Engineering, Massachusetts Institute of Technology, Cambridge, MA 02139.

*Present address: Scripps Institution of Oceanography, University of California at San Diego, La Jolla, CA 92093–0202.