Systems Properties of the *Haemophilus influenzae* Rd Metabolic Genotype*

(Received for publication, January 14, 1999, and in revised form, March 17, 1999)

Jeremy S. Edwards and Bernhard O. Palsson‡

From the Department of Bioengineering, University of California, San Diego, La Jolla, California 92093-0412

*This work was funded by the University of California Biotechnology Program (96–28), the National Institutes of Health (NIH 1RGM57089-01A1), and the National Science Foundation (MCB9873384). The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

‡ To whom correspondence should be addressed: Dept. of Bioengineering, University of California, San Diego, 9500 Gilman Dr., La Jolla, CA. 92093-0412. Tel.: 619-534-5668; Fax: 619-822-3120; E-mail: palsson@ucsd.edu.

*Haemophilus influenzae* Rd was the first free-living organism for which the complete genomic sequence was established. The annotated sequence and known biochemical information was used to define the *H. influenzae* Rd metabolic genotype. This genotype contains 488 metabolic reactions operating on 343 metabolites. The stoichiometric matrix was used to determine the systems characteristics of the metabolic genotype and to assess the metabolic capabilities of *H. influenzae*. The need to balance cofactor and biosynthetic precursor production during growth on mixed substrates led to the definition of six different optimal metabolic phenotypes arising from the same metabolic genotype, each with different constraining features. The effects of variations in the metabolic genotype were also studied, and it was shown that the *H. influenzae* Rd metabolic genotype contains redundant functions under defined conditions. We thus show that the synthesis of in silico metabolic genotypes from annotated genome sequences is possible and that systems analysis methods are available that can be used to analyze and interpret phenotypic behavior of such genotypes.

Genomics is a rapidly growing field. The complete genome sequence for 19 microorganisms is now available (1–4), and the first multicellular organism, *Caenorhabditis elegans* (5), has just been completely sequenced. It is expected that the full DNA sequences will become available for many human pathogens, as well as for several well known multicellular organisms, within just a few years (6, 7). Additionally, there are many ongoing efforts to identify the genes and assign putative function to their products (8–10), which will result in an essentially complete “parts catalogue” of the molecular components found in a multitude of living cells.

With the growing availability of defined genotypes, the question arises of whether the genotype-phenotype relation can be studied based on the genomic data. The experience with an increasing number of experimental systems shows that the relation between the genotype of an organism and its overall function is not simple (11). Genomics provides detailed information regarding the composition of an organism’s genome, but it does not provide knowledge on the dynamic and systemic characteristics that define the physiological function of a living system. Physiological processes are the result of multiple gene products working in a coordinated fashion, leading to the integrated functions of the system. Thus, the complex relation between the genotype and the phenotype cannot be predicted by cataloging and assigning functions to the genes found in a genome (11).

Although the genome sequence per se does not provide direct information about physiology, the definition of complete genotypes opens the possibility of systematically studying the genotype-phenotype relation using novel experimental and computational techniques. These novel approaches include methods to identify regulatory motifs and coregulated genes (12–20), to identify genes that are essential to support bacterial growth (21, 23), and to develop simulators to describe integrated cellular functions (24–26).

The results presented in this work utilized the *Haemophilus influenzae* annotated genome sequence, biochemical information, and a systems science-based analysis technique to further our understanding of the metabolic physiology of this bacterium. A high percentage (over 80%) of the ORFs identified in the bacterium *H. influenzae* have functional assignments (27–29), and the biochemical functions of the metabolic gene products are well known. Additionally, there is a long history of developing systems science descriptions of metabolic function (30–34). Therefore, it is logical to begin with metabolism for an analysis of integrated cellular functions. We have formulated an in silico description of the *H. influenzae* metabolic genotype from the available annotated genome sequence (27). Using the in silico metabolic genotype, we examined the systems characteristics of the metabolic network, studied the optimal phenotypic behavior, and examined the effects of in silico gene deletions on the ability of the metabolic network to support the growth of the cell.

MATERIALS AND METHODS

Formulation of the *H. influenzae* Metabolic Genotype—The metabolic genotype for *H. influenzae* was generated using its annotated genome sequence (27). The genes included in the metabolic genotype for *H. influenzae* Rd are shown in Table I. Of the enzymes included in the in silico metabolic genotype, 27 have not been identified by the genome annotations. Fourteen of these were included because of evidence in the literature, and six were included based on physiological evidence (Table I). The remaining seven enzymes, for which the genes have not been characterized, were included because there is evidence for these reactions being present (Table I). Based on the annotated genetic sequence and biochemical data, the *H. influenzae* metabolic genotype catalyzes 488 metabolic reactions and transport processes operating on a network of 343 metabolites.

Methods for Analyzing the Capabilities of Defined Metabolic Genotypes—Flux-balance analysis (FBA) is a method for assessing the capabilities and systemic properties of a metabolic genotype. The fundamentals of FBA have recently been reviewed (32, 35, 36).

---

1 J. S. Edwards and B. O. Palsson, submitted for publication.
2 The abbreviations used are:FBA, flux-balance analysis; PPP, pentose phosphate pathway.
matrix equation describes the steady-state mass balances of the metabolic network and is central to FBA.

\[ S \cdot v = b \]  
(Eq. 1)

where \( S \) is the stoichiometric matrix \((m \times n)\), \( v \) is the vector of \( n \) metabolic fluxes, and \( b \) is the vector representing \( m \) transport fluxes (i.e. known consumption rates, by-product production rates, and uptake rates). The stoichiometric matrix is derived directly from the defined metabolic genotype (Table I) \((m = 343 \text{ and } n = 488 \text{ for } H. \text{ influenzae})\).

The stoichiometric matrix, \( S \), is underdetermined \((n > m)\), and thus Equation 1 does not have a unique solution. Mathematically, this non-uniqueness is reflected in the null-space for \( S \). All metabolic flux solutions reside in the solution set, which is the null-space translated by a single vector \( (37) \). The solution set contains all metabolic flux distributions that satisfy the mass balance constraints (defined by Equation 1).

In addition to the mass balance constraints, there are physicochemical constraints on the metabolic fluxes that are defined by linear inequalities \((a_i \leq v_i \leq b_i)\). The physicochemical constraints are used to define maximum and minimum flux values. In this analysis \( a_i \) was set to zero for irreversible fluxes, and in all other cases \( a_i \) and \( b_i \) were unconstrained. The intersection of the solution set (mass balance constraints) and the region defined by the linear inequalities (physicochemical constraints) defines the feasible set. Each particular solution must be contained within the feasible set, and the particular solution represents the metabolic phenotype \((32, 38)\).

The genotype properties of interest can be studied by examining the feasible set of the metabolic system. Such an assessment is formulated as a linear programming problem \((32, 35)\).

\[ \text{Maximize } Z = \sum c_i v_i = (c \cdot v) \]  
(Eq. 2)

where \( Z \) is the objective function, representing a phenotypic property, and \( c \) is a vector of weights. LINDO was used to solve the linear programming problems \((LINDO \text{ Systems, Inc., Chicago})\). The objective, \( Z \), is maximized subject to the mass balance and physicochemical constraints. The objective functions utilized in this analysis are the maximization of biomass production, \( \delta \text{ biomass (Eq. 4)} \)

\[ \sum_{t=1}^{m} d_m X_m V_{\text{biomass}} \]  
(Eq. 4)

where \( d_m \) are derived from the biomass composition of each metabolite \( (X_m) \), and the maximization of the production of the charged form of the metabolic cofactors. The biomass composition for Escherichia coli is used in the computations \((39–41)\). It has been shown that the FBA results are not sensitive to biomass composition \((42)\), and therefore this should not have a significant effect on the results. However, given the flexibility of the approach described herein, the \( d \) vector can be adjusted to account for any differences.

Phenotype Phase Diagram—The phenotype phase diagram was generated using the sensitivity analysis of the linear programming package LINDO. The sensitivity analysis defines the amount by which a given component of the \( b \) vector can change without changing the basis

\( (32, 45) \)
various carbon substrates. The capabilities of a metabolic genotype is the ability of the metabolic genotypes to produce an important metric in the comparison of different metabolic cellular metabolism in the above described connectivity characteristics show, the metabolic metabolites with the largest difference in participation between E. coli and H. influenzae are listed in the Table inset. TCA, tricarboxylic acid cycle.

The shadow prices, or the dual variables, in a linear programming problem arise from the solution of the dual problem (43). These variables are used to interpret the metabolic state of the cell. The shadow prices are interpreted as the intrinsic value of a given metabolite to the cell (44). The shadow prices also undergo discontinuous changes at the demarcation lines.

RESULTS

The systems characteristics of the H. influenzae metabolic genotype can be studied based on the properties of its stoichiometric matrix. Here we present a study of its 1) connectivity properties, 2) ability to produce charged forms of metabolic cofactors, 3) optimal use of its metabolism to meet the growth requirements, and 4) sensitivity to the loss of gene product function in central intermediary metabolism.

Connectivity of Metabolic Intermediates—The number of metabolic reactions in which each of the 343 metabolites in the H. influenzae in silico metabolic genotype is involved varies across several orders of magnitude (Fig. 1). The metabolites can be rank-ordered by the number of reactions in which they participate. H. influenzae metabolism revolves around relatively few highly connected metabolites. The metabolites involved in the largest number of metabolic reactions are ATP, ADP, inorganic phosphate, and pyrophosphate. Even though H. influenzae does not possess the isocitrate dehydrogenase enzyme to synthesize α-ketoglutarate, α-ketoglutarate is a highly connected metabolite and participates in 17 reactions. Glutamate and glutamine also participate in a relatively large number of metabolic reactions, 31 and 10, respectively.

The degree of interconnectivity illustrates how metabolism must be coordinated around a few key metabolites that represent phosphate (energy), carbon, nitrogen, and redox metabolism (Fig. 1). Therefore, it is likely that metabolic regulation will revolve around the careful control of these metabolites.

Production of Metabolic Energy and Redox Potential—As the above described connectivity characteristics show, the metabolic cofactors play an important role in the function and coordination of cellular metabolism in H. influenzae. Thus, we can expect that an important metric in the comparison of different metabolic genotypes is the ability of the metabolic genotypes to produce charged forms of the cofactors, ATP, NADH, and NADPH, on various carbon substrates. The capabilities of a metabolic genotype to produce these cofactors can be determined by optimizing (using proper weights as shown in Equation 3) their production on a given substrate (45). The vector was defined to allow only the single carbon source to enter the metabolic system. Cofactor production capabilities were determined in this way, and the results are summarized in Table II.

The optimal production of ATP by H. influenzae from fructose (the only carbohydrate for which a PTS transporter was identified in the DNA sequence (27)) was determined to be 9.3 mol/mol. Approximately half of this ATP production is the result of substrate level phosphorylation, whereas the other half is produced by oxidative phosphorylation. The maximal production of both NADH and NADPH is 8.0 mol/mol.

This cofactor production ability of the H. influenzae metabolic genotype compares with 20.5 mol/mol maximal ATP production in E. coli, and a maximal production of NADH 11.6 and 12.0, respectively, with fructose and glucose as the energy sources. The maximal production of NADPH is 10.8 and 11.4 mol/mol, respectively, with fructose and glucose as the energy sources. These comparisons show that the reduced metabolic genotype of H. influenzae has a decreased ability to generate charged forms of the metabolic cofactors from the same substrate.

Flux Distributions for Optimal Growth—The metabolic flux distributions for optimal growth were determined in silico for H. influenzae Rd in defined media. The medium components required for the growth demands of the in silico H. influenzae Rd strain are shown in the legend to Fig. 2. The in silico determined medium is similar to the experimentally determined defined medium for the growth of other strains of H. influenzae (46–49). Several of the experimentally determined defined media contain additional compounds (47–49); however, the defined medium discussed by Klein and Lugtenburg (46) is considered a defined “minimal” medium and differs from our in silico defined media by glutathione (replaced by cysteine) and inosine.

H. influenzae in silico requires multiple substrates for its growth, with fructose and glutamate being the two key substrates. The vector describing the uptake of the metabolites is described in the legend of Fig. 2. The optimal use of the metabolic pathways for the growth of H. influenzae on the defined media was determined using established methods (32, 35, 50, 51). The metabolic flux distributions were calculated for all combinations of fructose and glutamate uptake rates. The optimal utilization of the metabolic genotype to meet the cellular growth requirements was determined to be dependent on the uptake rates of these two substrates.

Fig. 2 is a phenotype phase diagram showing the different optimal metabolic phenotypes and their characteristics that can be derived from the H. influenzae metabolic genotype depending on the substrate (fructose and glutamate) uptake rates. The six regions are described in the following paragraphs.

In region 1, the capability of the H. influenzae metabolic genotype to meet growth requirements is limited by its ability to generate the biosynthetic precursors derived from fructose. A low CO2 production and a low acetate production characterize the optimal metabolic phenotype in this region. The optimal
utilization of the metabolic pathways results in a low production of the metabolic by-products because of the large demand for the metabolic precursors. The optimal flux distribution also utilizes the nonoxidative branch of the PPP, thus reducing the production of CO₂.

In region 2, cellular growth is limited by the ability of the metabolic network to produce high energy phosphate bonds and redox potential. The optimal metabolic phenotype in this region is characterized by cycling of the PPP for the generation of energy. The transhydrogenase reaction is utilized to convert the redox potential into energy in the form of the proton motive force. There is a high CO₂ production, and acetate production is still low (although it is increased relative to region 1).

Region 3 is also limited in terms of the generation of metabolic energy and redox potential. The oxygen demands in this region surpass the ability of oxygen to reach the cell because of diffusion constraints. There is an increased demand for high energy phosphate bonds and a decreased demand for redox potential relative to region 2. The optimal metabolic phenotype in this region is characterized by decreased fluxes through the oxidative branch of the PPP, decreased CO₂ production, and increased acetate production.

Region 4, similar to regions 2 and 3, is limited by the ability to generate high energy phosphate bonds and redox potential. However, in this region there is a shift in the demand for NADPH relative to NADH. The NADPH demand for biosynthesis is increased relative to NADH when compared with the other energy-limited regions (region 2 and 3), which is evident by the utilization of the transhydrogenase to convert the NADH into NADPH (a reversal from region 3). Fluxes with arrows are zero, fluxes shown in light gray are decreased, and fluxes shown with thick lines are increased relative to the next lower region. Fluxes in black are unchanged with respect to the next lower region.

Region 5 is characterized by the excess redox potential. The large glycolytic flux leads to a condition in which the ability to eliminate the redox potential is limiting growth. The oxidative branch of the PPP is not utilized under optimal conditions for this region, and thus the biosynthetic precursors are generated by the nonoxidative branch. Similar to region 4, the NADPH for the biosynthetic reactions is optimally generated, using the transhydrogenase reaction to convert the excess redox potential in the form of NADH into NADPH. The CO₂ production is low, and a high acetate production is optimal. Additionally, the

**FIG. 2.** Phenotype phase diagram of the *H. influenzae* Rd metabolic phenotype. The qualitative optimal metabolic phenotype is represented. The growth of the bacteria is simulated in the following defined media: fructose, arginine, cysteine, glutamate, putrescine, spermidine, thiamin, NAD, haem, pantothenate, ammonia, and phosphate. The b vector elements for arginine, cysteine, putrescine, spermidine, thiamin, NAD, haem, and pantothenate were assigned an inequality constraint restricting the maximal uptake rate below 2 mmol/g dry weight (DW/h); the oxygen b vector element was assigned an inequality constraint restricting the maximal uptake rate below 20 mmol/g dry weight/h; and the b vector elements for carbon dioxide, phosphate, and ammonia were unconstrained. The b vector elements were set to allow the metabolic by-products (acetate, formate, succinate, lactate, and pyruvate) to leave the system. The metabolic phenotype is represented as a function of two metabolic uptake rates. The uptake (b vector value) of fructose and glutamate was varied to generate the phase plane. *H. influenzae* is shown to exhibit six different phenotypes. A shadow price analysis (43) was used to construct the phase portrait. The boundaries between the metabolic phenotypes is likely to be a “gray area” in which a switch between the qualitative regions occurs. The qualitative metabolic flux map for each region is shown in the insets. The metabolic fluxes were normalized with respect to the growth rate and color-coded to indicate the qualitative changes that occur when moving from a lower to a higher number region (i.e. flux changes when moving from region 2 to region 3). Fluxes with arrows are zero, fluxes shown in light gray are decreased, and fluxes shown with thick lines are increased relative to the next lower region. Fluxes in black are unchanged with respect to the next lower region.
The results also demonstrate that there is flexibility in the cellular genome to achieve the optimal metabolic performance. FBA can define the capabilities of the metabolic genotypes, and additionally, it also will suggest the optimal utilization of the cellular genome to achieve the optimal metabolic performance. The loss of enzymatic function resulted in a range of different behaviors, which were grouped into three different categories: lethal, critical, or redundant (Fig. 3). It was determined that during growth under conditions defined by region 3 (point A shown in Fig. 2), 33% (12 of 36) of the gene products are essential, meaning that the deletion of any of these gene products is lethal to H. influenzae growing in the defined medium. 25% (10 of 36) of the gene products were found to be critical; loss of function of these gene products was nonlethal, but it resulted in a decreased ability to grow. 42% (14 of 36) of the gene products are considered redundant for growth in the defined media.

The consequences of alterations in the metabolic genotype can be assessed. The in silico strain of H. influenzae Rd was subjected to deletions in the gene products of the central metabolic pathways of glycolysis, pentose phosphate pathway, tricarboxylic acid cycle, and respiration processes. The optimal growth performance was evaluated while each of the gene products involved in the aforementioned pathways was removed from the system. Genes that code for isozymes or genes that code for components of the same enzyme complex were simultaneously removed (i.e. aceEF, sucCD). The genes that are considered in the analysis are set in nonitalic (i.e. glgA), or they are not utilized in the conditions that were examined and thus will not provide any additional information (i.e. eda). A set of 36 different enzymes in the H. influenzae genotype was considered in the analysis. The ability of the altered metabolic genotype to compensate for the loss of enzymatic function was evaluated in silico during growth in the defined media.

The loss of enzymatic function resulted in a range of different behaviors, which were grouped into three different categories: lethal, critical, or redundant (Fig. 3). It was determined that during growth under conditions defined by region 3 (point A shown in Fig. 2), 33% (12 of 36) of the gene products are essential, meaning that the deletion of any of these gene products is lethal to H. influenzae growing in the defined medium. 25% (10 of 36) of the gene products were found to be critical; loss of function of these gene products was nonlethal, but it resulted in a decreased ability to grow. 42% (14 of 36) of the gene products are considered redundant for growth in the de-
fined medium because essentially equivalent flux distributions can be implemented without the presence of any of the respective enzymatic functions.

These results show that H. influenzae, compared with E. coli, is less capable of overcoming the loss of function of gene products during growth in a defined medium. For E. coli, 14, 18, and 69% of the gene products are considered essential, critical, and redundant, respectively. Of the seven gene products in E. coli that were determined to be essential for growth in glucose minimal medium, three are not present in H. influenzae. These gene products are involved in the first three reactions of the tricarboxylic acid cycle. The lack of these functions in H. influenzae has created a requirement for glutamate in the growth media. The other four essential gene products in E. coli were also determined to be essential for H. influenzae during growth in the defined medium. These essential gene products are transketolase, ribose-5-phosphate isomerase, glyceraldehyde-3-phosphate dehydrogenase, and phosphoglycerate kinase.

Optimal growth performance was also determined for all possible combinations of the simultaneous loss of two (630 combinations) (Fig. 3) and three gene products (7140 combinations). There are very few nontrivial lethal double (7 of 361 lethal gene pairs) and triple (7 of 5270 lethal gene triplets) gene deletions. A nontrivial lethal double deletion is defined as a combination of two genes that when removed from the metabolic genotype results in a lethal phenotype. However, the removal of either gene individually from the metabolic genotype does not result in a lethal phenotype. Similarly, for nontrivial lethal triple deletions, the additional condition in which the double deletion of any two of the gene products does not result in a lethal phenotype is also considered. This result is noteworthy and suggests that there are relatively few critical gene products in metabolic pathways considered for this pathogenic microorganism while growing in the defined media.

DISCUSSION

The work presented herein demonstrates a novel methodology for the exploitation of the biological databases and annotated genome sequence information to gain an understanding of the complex relation between the metabolic genotype and the optimal phenotypes derived therefrom. We have defined an in silico metabolic genotype for H. influenzae Rd based on the annotated DNA sequence and biochemical information. This in silico representation of the H. influenzae metabolic machinery was used to study its systemic characteristics. The analysis of the in silico H. influenzae metabolic genotype has introduced a methodology for the utilization of DNA sequence information to gain an understanding of the integrated physiology of a cellular system. More specifically, the results presented above illustrate the systemic effect of the reduced metabolic network on the production of the charged form of the metabolic cofactors and utilization of the finite number of optimal metabolic phenotypes to identify the essential feature of flexibility in the metabolic network.

The ability of the in silico H. influenzae metabolic network to produce the high energy phosphate bonds and redox potential was assessed and can potentially be used as a metric for comparative functional genomics. It was determined that the H. influenzae metabolic genotype has a reduced capacity to produce the charged form of the metabolic cofactors, which leads to several important physiological consequences. For instance, the FBA, presented above for the H. influenzae metabolic genotype, demonstrates that the optimal metabolic network flux distribution to support biomass production utilizes the PPP in a variety of different primary metabolic roles, suggesting that the physiological role of the metabolic pathways will be dependent on the overall genotype and the environment in which it operates. Thus, the physiologic role of a metabolic pathway is not simply a function of its absence or presence in cell, but rather it is a function of the entire genome as well as the environmental conditions. This emphasizes the utility of redefining the metabolic pathways in the different completely sequenced organisms with a functional, rather than historical, definition (53, 54).

The global effect on the metabolic network of in silico gene deletions was also assessed with the H. influenzae metabolic genotype. The in silico approach described herein provides a method for determining the genes that are essential for bacterial growth. This question is important, and several experimental strategies are available (21, 23, 55). However, for examining an entire genome, these experimental programs are ambitious and will be time-consuming. Therefore, an in silico program can be used to aid in the design of an experimental strategy.

The in silico deletion analysis was performed on the H. influenzae gene products involved in central intermediary metabolism. The results suggested that, under a single well defined condition, there is redundancy in the H. influenzae metabolic genotype, but it is unlikely that truly redundant functions would be evolutionarily conserved. However, we have also shown (Fig. 2, phenotype phase diagram) that the optimal metabolic pathway utilization is a function of the substrate availability. Thus, if the deletion analysis is spanned across the phenotype phase diagram, the number of redundant genes is reduced to 9 of 36 (results not shown). Additionally, if the regions in another phenotype phase diagram (fructose versus oxygen, not shown) are analyzed using the in silico deletion analysis, the number of redundant gene products is further reduced to 5 of 36 (frd, dld, pch, pfk, sfc). Thus, it is likely that this apparent redundancy provides an essential feature, here called flexibility, in the metabolic pathways. The metabolic flexibility is likely a beneficial feature that the bacteria use to adjust to different conditions. H. influenzae, a parasitic organism, which sees a relatively constant environment, has retained some degree of flexibility. Thus, the benefit to the bacteria to be able to adjust to changing conditions must be greater than the metabolic burden of maintaining these genes in the genome.

The future of many areas of biological study will depend greatly upon the ability to capitalize on the wealth of genetic and biochemical information currently being generated from the fields of genomics and, similarly, proteomics. With such detailed information available about an organism’s arsenal of metabolic reactions, the ability to perform detailed studies of the systemic metabolic capabilities has been demonstrated. This development is significant from a fundamental and conceptual standpoint, as it yields a holistic definition of biochemical processes. Additionally, this perspective for studying cellular processes will play a role in 1) gaining insight into the regulatory logic implemented by the cell to control its metabolic pathways and 2) analyzing the production capabilities of the global metabolic network along with understanding the robustness and sensitivity of the network to alteration in its metabolic genotype. Undoubtedly, studies of this nature hold potential value for research in various fields, including metabolic engineering for bioprocesses and therapeutics, bioremediation, and antimicrobial research.

We have presented a method of analysis to aid in the understanding of this complex relation. However, the construction of in silico cells and the analysis considered herein should be considered to be only the first step toward the integrative analysis of bioinformatic data bases to predict and understand cellular function based on the underlying genetic content. Continued prediction and experimental verification will be an in-
H. influenzae Metabolic Genotype

(Neidhardt, F. C., ed) Vol. 1, pp. 721–730, ASM Press, Washington, D. C.

Akerley, B. J., Rubin, E. J., Camilli, A., Lampke, D. J., Robertson, H. M., and Mekalanos, J. J. (1998) Proc. Natl. Acad. Sci. U. S. A. 95, 8927–8932

Lee, Y., and Yin, J. (1996) Nat. Biotechnol. 14, 491–493

Endy, D., Kong, D., and Yin, J. (1997) Biotechnol. Bioeng. 55, 375–389

Trivedi, B. (1998) Nat. Biotechnol. 16, 1316–1317

Fleischmann, R. D., Adams, M. D., White, O., Clayton, R. A., Kirkness, E. F., Kerlavage, A. R., Bult, C. J., Tomb, J. F., Dougherty, B. A., Merrick, J. M., McKenney, K., Sutton, G., FitzHugh, W., Fields, C., Gocayne, J. D., Scott, J., Sherrill, R., Liu, L. L., Gloekle, A., Kelley, J. M., Weidman, J. P., Phillips, C. A., Spriggs, T., Hedblom, E., Cotton, D. M., Utterback, T. R., Hanna, M. C., Nguyen, D. T., Saudek, D. M., Brandon, R. C., Fine, L. D., Fritchman, J. L., Fuhrmann, J. L., Geoghegan, N. S. M., Gnehm, C. L., McDonald, L. A., Small, K. V., Fraser, C. M., Smith, H. O., and Venter, J. C. (1995) Science 269, 496–498, 507–512

Tatusov, R. L., Mushegian, A. R., Bork, P., Brown, N. P., Hayes, W. S., Borodovsky, M., Rudd, K. E., and Koonin, E. V. (1996) Curr. Biol. 6, 79–91

Robison, K., Gilbert, W., and Church, G. M. (1996) Science 271, 1302–1304

Reich, J. G., and Selker, E. E. (1981) Energy Metabolism of the Cell, 2nd Ed., Academic Press, New York

Fell, D. (1996) Understanding the Control of Metabolism, Portland Press, London

Varma, A., and Palsson, B. O. (1994) Bio/Technology 12, 994–998

Heinrich, R., and Schuster, S. (1996) The Regulation of Cellular Systems, Chapman and Hall, New York

Shuler, M. L., and Domach, M. M. (1985) in Foundations of Biochemical Engineering (Blanch, H. W., Papoutsakis, E. T., and Stephanopoulos, G., eds), p. 101, American Chemical Society, Washington, D. C.

Bonarius, H. J., Schmid, G., and Tramer, J. (1997) Trends Bio/Technol. 15, 308–314

Edwards, J. S., Ramakrishna, R., Schilling, C. H., and Palsson, B. O. (1998) in Metabolic Engineering (Lee, S. Y., and Papoutsakis, E. T., eds), Springer-Verlag, Berlin, in press

Strang, G. (1988) Linear Algebra and Its Applications, 3rd Ed., Saunders College Publishing, Fort Worth, TX

Edwards, J. S., and Palsson, B. O. (1998) Bio/Technol. Bioeng. 58, 162–169

Neidhardt, F. C., Ingraham, J. L., and Schaechter, M. (1990) Physiology of the Bacterial Cell, Sinauer Associates, Inc., Sunderland, MA

Neidhardt, F. C., and Umbarger, H. E. (1996) in Escherichia coli and Salmonella: Cellular and Molecular Biology (Neidhardt, F. C., ed) Vol. 1, 2nd Ed., pp. 13–16, ASM Press, Washington, D. C.

Chvatal, V. (1983) Linear Programming, W. H. Freeman and Company, New York

Varma, A., and Palsson, B. O. (1993) J. Theor. Biol. 165, 477–502

Varma, A., Boeche, B. W., and Palsson, B. O. (1993) Bio/Technol. Bioeng. 42, 59–73

Klein, R. D., and Lugtenbuhl, G. H. (1979) J. Gen. Microbiol. 113, 409–411

Herriott, R. M., Meyer, E. Y., Vogt, M., and Modan, M. (1970) J. Bacteriol. 101, 513–516

Talmadge, M. B., and Herriott, R. M. (1980) Biochem. Biophys. Res. Commun. 2, 203–206

Butler, I. O. (1962) J. Gen. Microbiol. 27, 51–60

Savinell, J. M., and Palsson, B. O. (1992) J. Theor. Biol. 154, 421–454

Savinell, J. M., and Palsson, B. O. (1992) J. Theor. Biol. 154, 455–473

Xars, A., Elkmann, B. J., Sahm, H., de Graaf, A. A., and Eggeling, L. (1999) Metab. Eng. 1, 38–48

Schilling, C. H., and Palsson, B. O. (1998) Proc. Natl. Acad. Sci. U. S. A. 95, 4193–4198

Schuster, S., and Hilgetag, C. (1994) J. Biol. Syst. 2, 165–182

Link, A. J., Phillips, D., and Church, G. M. (1997) J. Bacteriol. 179, 6228–6237

Macduffy, L. P., and Redfield, R. J. (1996) Res. Microbiol. 147, 541–551