Chapter 18

FLUX DETERMINATION IN CELLULAR BIOREACTION NETWORKS: APPLICATIONS TO LYSINE FERMENTATIONS

Joseph J. Vallino and Gregory Stephanopoulos

TABLE OF CONTENTS

Library of Congress Cataloging-in-PublicationData

Frontiers in bioprocessing/editors, Subhas K. Sikdar, Milan Bier, Paul Todd.
p. cm.
Plenary lectures and poster presentations presented at a conference held June 28—July 2, 1987, Boulder, Colo.
Bibliography: p.
Includes index.
ISBN 0-8493-5839-6
1. Biotechnology—Technique—Congresses. 2. Separation (Technology)—Congresses. I. Sikdar, Subhas K. II. Bier, Milan. III. Todd, Paul, 1936-
TP248.24.F76 1989
660'.6—dc19

Direct all inquiries to CRC Press, Inc., 2000 Corporate Blvd., N.W., Boca Raton, Florida, 33431.

International Standard Book Number 0-8493-5839-6

Library of Congress Card Number 89-30528
Printed in the United States

89-30528
CIP

I. INTRODUCTION

Mutation/selection or genetic engineering techniques employed to increase product yield in fermentations often only alleviate metabolic regulation occurring at or near the end of a biosynthetic chain that leads to that particular product. However, microorganisms have evolved biosynthetic pathways to produce energy and metabolites necessary for growth and replication, not for the overproduction of specific biocompounds. Consequently, not only does the regulation of the product of interest need to be removed, but also the main fueling reactions of the cell must be rerouted in such a way as to channel the main carbon flux into the biosynthetic pathway required for product synthesis.

To achieve, or at least approach, theoretical product yields in an optimal manner, one must know how carbon partitioning occurs between the fueling reactions, biomass, and product. This necessitates the development of means to determine the partitioning of carbon flux between the primary biosynthetic pathways. Once acquired, this information can be used to rechannel the carbon in the fueling reactions into the product pathway by amplification or attenuation of the enzymes associated with the limiting reactions. This type of global modification process has been termed metabolic engineering, yet there does not exist any simple way to determine the flux of carbon through the primary metabolic pathways.

Although the fluxes can be estimated, in theory, by constructing a set of differential equations to model the concentrations of the important metabolites in the network, these models rely on the regulatory and kinetic properties of individual enzymes in the network. These enzymes are poorly understood or not known at all for most microbes. Even though such models are useful, kinetic and regulatory information of individual enzymes *in vitro* does not necessarily indicate how the overall biosynthetic network functions as a whole *in vivo*. As will be shown, models this detailed are not necessary to estimate the fluxes in the network.

Experimentally, it is possible, to a limited extent, to calculate the carbon flux on the basis of radiocarbon labeling experiments; however, these experiments often disrupt the cellular environment, are difficult and expensive to conduct, and cannot be used in a fermentation environment due to the large volumes involved.

We have developed a methodology that determines the carbon flux through the primary biosynthetic pathways employed for product synthesis from measurements of extracellular metabolites only. This algorithm, presented in Section II, couples measurements of extracellular metabolites with the known metabolic pathways of the microbe of interest, and a pseudo-steady-state (PSS) approximation for intracellular metabolites is used to generate a bioreaction network equation. Once constructed, singularity and sensitivity analysis routines determine if the system is well posed. Often these algebraic systems have more equations than unknowns; consequently, redundancy analysis is used to check the consistency of measurements and the PSS approximations. Least-squares or quadratic programming techniques are employed to solve the equation and produce an estimate of the carbon flux.

The actual fermentation being studied is the production of lysine from glucose by *Corynebacterium glutamicum*. This process is a good example of mutation/selection techniques employed to construct production mutants free of feedback regulation by lysine and threonine on aspartate kinase. However, these fermentations have molar yields (moles lysine produced per mole glucose consumed) of approximately 30 to 40%,[1] while the theoretical maximum is about 75%[2] when NADPH, is assumed not to be equivalent to $NADH_2$ and about 86%[3] when NADPH, and $NADH_2$ are assumed to be equivalent. The difference in theoretical yields is caused by CO_2 evolution in the pentose phosphate pathway (PPP) when it is used to generate NADPH,. Since essentially no other products are produced and biomass synthesis is minimal, glucose is simply being oxidized to carbon dioxide and water. Consequently, it might be possible to redirect this carbon flux into lysine. Literature data from

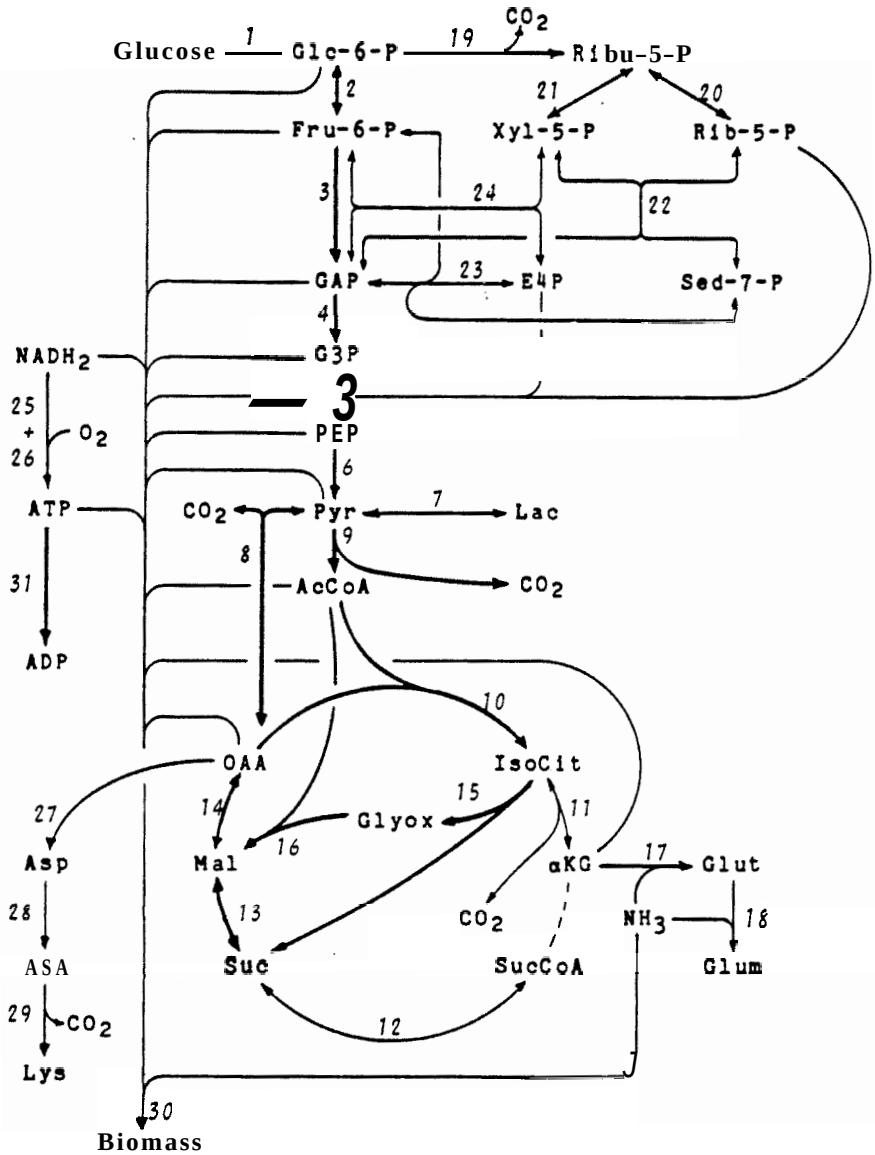

FIGURE 1. Bioreaction network for lysine production by glutamic acid bacteria.

a lysine fermentation of *Brevibacterium* are used to illustrate the methodology, since *Brevibacterium* species that produce lysine or glutamate are basically identical to C. *glutamicum* for our purposes.

II. METHODOLOGY

Consider the simplified biosynthetic reaction network for lysine production by glutamic acid bacteria illustrated in Figure 1. The estimation of the carbon flux through each reaction in this network from extracellular measurements is the desired goal. Although this simplified network is used to represent cellular metabolism, it is quite obvious that not all biosynthetic reactions have been incorporated. There are thousands of such reactions; to include them all would be impractical. The first step in the development of the methodology is to extract

those reactions that represent the major carbon fluxes. This information can usually be found in the literature.

To construct a simplified network, the main fueling and metabolite-generating bioreactions, such as the Embden-Meyerhof-Parnas pathway, the PPP, the Krebs or tricarboxylic acid (TCA) cycle, and the glyoxylate shunt, are assembled. In order to maintain observability of the overall network from extracellular measurements, the pathways that couple extracellular metabolites (including the product) to the fueling reactions must be included. Since a PSS approximation will be used for intracellular metabolites, regenerating reactions such as those for ATP and NAD via the respiratory chain are included to insure that no intracellular metabolite has a net production or consumption. Elaborate or ill-defined pathways, such as biomass synthesis or maintenance requirements, must be expressed as lumped reactions (the details of lumping will be explained in Section II.A). Finally, to minimize the dimensionality of the system, only metabolites that are involved at branch points in the biosynthetic pathways are considered in the network. For example, there are several metabolites between aspartylsemialdehyde (ASA) and lysine (Lys); however, they need not be considered, since they comprise a nonbranching sequence of reactions that must all proceed at the same rate if PSS assumption is to be satisfied.

The reactions used for the lysine fermentation as well as the important metabolites are presented in the Appendix and illustrated in Figure 1.

A. BIOMASS AND MAINTENANCE EQUATIONS

Of the multitude of intracellular reactions, the majority lead to the synthesis of large biomolecules needed for cell growth and maintenance. Although essential, the flux through any one such reaction represents only a small fraction of the total carbon entering the cell. The sum total of these reactions, however, cannot be neglected, for they represent the pathways through which carbon and other essential nutrients are incorporated into biomolecules. In our methodology, all such reactions are lumped into one overall biomass equation and coupled into the main fueling reactions.

Ingraham et al.[4] have shown that biomass can be synthesized from 12 precursor metabolites, and the researchers have calculated the amounts required to form 1 g dry weight of biomass, as well as the amounts of ATP, $NADH_2$, and NADPH, required for biosynthesis. Their analysis is similar to that of Stouthamer.[5] Even though the reported biomass yields are for *Escherichia coli,* they are adequate approximations for C. *glutamicum* for the following two reasons. First, in most lysine fermentations biomass synthesis only requires approximately 10 to 20% of the initial glucose, and the growth rate is quite small or zero during lysine production. Consequently, the sensitivity of the flux estimates on the biomass yields is quite small, especially during lysine production. Second, we have been able to duplicate the biomass yields reported by Ingraham et al.[4] for threonine and methionine to within less than 5%. Therefore, the use of E. *coli* biomass yields for C. *glutamicum* is warranted; however, this will probably not be true for all microorganisms and/or fermentations. From the tables given by Ingraham, the lumped equation for biomass synthesis for C. *glutamicum* is shown in Table 1.

The ATP requirement in the lumped biomass equation is the theoretical amount and is equivalent to an ATP yield of 28 g of biomass synthesized per mole of ATP consumed. The lumped equation for biomass does not account for ATP consumption due to maintenance, futile cycles, transport costs, or the energy required to maintain concentration gradients across the cell wall. Consequently, an excess amount of ATP is usually produced by the fueling reactions. To maintain a PSS approximation for ATP, this excess is removed by incorporating the following equation into the network for the conversion of ATP into ADP:

$$\text{ATP} \rightarrow \text{ADP} \qquad (1)$$

TABLE 1
Lumped Biomass Equation from Ingraham's Data[4]

Substrates (moles)	Products (moles)
GLC6P (0.0205)	
FRU6P (0.0071)	
RIB5P (0.0898)	
E4P (0.0361)	
GAP (0.0129)	
G3P (0.150)	
AKG (0.058)	
OAA (0.107)	
PEP (0.0519)	Biomass (1.0)
PYR (0.125)	ADP (3.89)
ACCOA (0.327)	NADP (1.37)
NH_3 (0.7%)	NADH (0.312)
ATP (3.89)	
NADPH (1.37)	
NAD (0.312)	
LEU (0.043)[a]	
MET (0.015)[a]	
THR (0.024)[a]	
LYS (0.033)[b]	
GLUT (0.025)[b]	
GLUM (0.025)[b]	

Note: Mol wt of biomass = 100.

[a] Auxotroph.
[b] Pathway included.

Although the ATP balance is not necessary for flux determination, Equation 1 is informative since it indicates how much excess ATP, above the theoretical, is produced by the cells. For the lysine fermentation, a P/O ratio of 2 is used.

The reactions illustrated in Figure 1 and presented in the Appendix comprise the metabolism of the cell for the production of lysine by glutamic acid bacteria. A similar approach can be applied to any other organism for any product so long as the biochemistry is known. For protein products, a lumped equation similar to the biomass equation would probably have to be used.

Since it is desired to estimate the extent (i.e., the flux) of each reaction in the network from extracellular measurements only, a metabolite balance based on the biochemistry is used to construct an algebraic equation for the metabolism of the cell. The development of this equation, called the bioreaction network equation (BRNE), is described in the next section and follows a similar approach established for butyric acid bacteria by Papoutsakis[6,7] for the construction of the fermentation equation.

B. THE BIOREACTION NETWORK EQUATION

The extent or flux of each reaction listed in the Appendix is unknown and can be represented as x_i for the ith reaction in the network. The BRNE represents a metabolite balance and is constructed by determining the time rate of change of each metabolite in the network as a function of all the unknown flows, x_i, producing or consuming that metabolite. For example, in the lysine fermentation the rate of production of pyruvate based on the reactions given in the Appendix is

$$r_{pyr} = x_6 - x_7 + x_8 - x_9 - x_{29} - 0.13x_{30} \qquad (2)$$

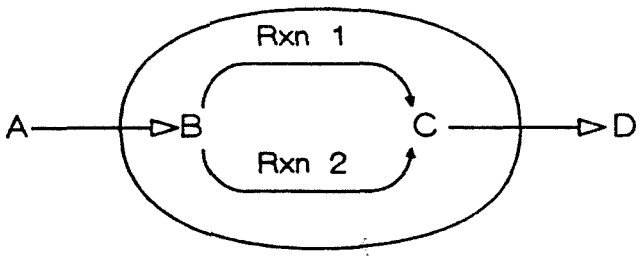

FIGURE 2. Example of a singular reaction network.

The resulting set of equations for the rate of change of each metabolite in the network is placed in matrix form as follows:

$$\mathbf{Ax} = \mathbf{r} \tag{3}$$

where A is the bioreaction network matrix reflecting the assumed biochemistry, **r** is the production rate vector for all the metabolites in the network, and **x** is the unknown flux vector to be determined.

The A matrix is determined by the biochemistry of the microbe and the fermentation, as described above, and has a dimension of m × n, where m is the number of metabolites in the network, n is the number of reactions, and $m \geq n$. For the lysine fermentation, $m = 33$ and $n = 31$.

1. Singularities

A solution to the BRNE will exist provided the A matrix is nonsingular. Singularities in the A matrix may arise due to reaction dependence or network observability problems. Consider, for example, the simplified network depicted in Figure 2. In this case, Reactions 1 and 2 are indistinguishable from the extracellular measurements of A and D, and the resulting A matrix would be singular. Reactions in the network that produce such singularities must be either lumped together or removed; in the above example, both yield the same result.

When the complete TCA cycle and the glyoxylate shunt are incorporated into the network for lysine fermentation, a singularity arises indicating that the determination of carbon flux distribution between these two branches is not possible from available extracellular measurements. It has been shown, however, that a-ketoglutarate dehydrogenase (EC 1.2.4.2) of the TCA cycle is not active in glutamic acid bacteria[8] (shown as the dashed line in Figure 1). When this reaction is removed from the network, so is the singularity. Similarly, carboxylation reactions, such as the malate enzyme (EC 1.1.1.40), pyruvate carboxylase (EC 6.4.1.1), oxaloacetate decarboxylase (EC 4.1.1.3), and phosphoenolpyruvate carboxylase and carboxykinase (EC 4.1.1.31 and 4.1.1.32), must be lumped into one reaction (Reaction 8 in the Appendix). It should also be pointed out that matrix singularities are one of the main reasons that biosynthetic reactions are lumped into a single equation.

Once the bioreaction network matrix, A, is assembled and reactions that produce singularities lumped together or removed, the production rate vector, r, must be determined.

2. Production Rate Vector

Since biomass is treated as a product just like any other, each element of the **r** vector represents the time rate of change in total concentration of a metabolite in the network minus the intracellular component of that metabolite which has already been accounted for in the measurement of biomass, or

$$r_j = \frac{dC_j^T}{dt} - \mu C_j^I \qquad \text{and} \qquad r_{biomas} = \frac{dC_{biomass}}{dt} \tag{4}$$

$$r_j = \frac{dC_j^E}{dt} + \frac{dC_j^I}{dt} - \mu C_j^I \qquad \text{and} \qquad r_{biomas} = \frac{dC_{biomass}}{dt} \tag{5}$$

where C_j^T is the total concentration of metabolite j (total moles of metabolite j per total volume of broth) and is equal to $C_j^E + C_j^I$, where the superscripts E and I refer to extracellular and intracellular components, respectively. The subscript j refers to all metabolites listed in the Appendix except for biomass, and μ is the specific growth rate. Since it is desired to base the methodology on measurements of extracellular metabolites only, a PSS (i.e., balanced growth) approximation is used for all intracellular metabolites, given by

$$\frac{dC_j^I}{dt} = \mu C_j^I \tag{6}$$

Consequently, after substitution of Equation 6 into Equation 5, the production rate vector, **r,** simplifies to

$$\begin{aligned} r_i &= 0 && \text{for all intracellular metabolites} \\ r_k &= \frac{dC_k^E}{dt} && \text{for all extracellular metabolites and biomass} \end{aligned} \tag{7}$$

that is, r_i is set to zero for all intracellular metabolites and r_k is determined by measuring the production rate of the extracellular component of metabolite k. Of course, dilution and sampling effects on concentration data must be accounted for when appropriate. The BRNE is now determined completely and can be solved.

Before presenting the solution, it is worth noting that the bioreaction network algorithm (BRNA) does not require information concerning intracellular control mechanisms or kinetic rate constants for reactions. The overall network is simply a metabolic balance governed by the particular biochemistry of the microbe. Consequently, the only uncertainties are the lumped equations and the PSS hypothesis, both of which are reasonable approximations. Furthermore, Equation 6, the PSS approximation, need not be held to rigorously. As long as r_j for all intracellular metabolites, given by Equation 5, is small compared to the fluxes producing or consuming metabolite **j,** the estimated fluxes will be a valid representation of the true state of the cellular metabolism.

3. Solution and Sensitivity Analysis

The estimated flux vector, $\bar{\mathbf{x}}$, is given by the least squares solution[9] to the BRNE **3:**

$$\bar{\mathbf{x}} = (\mathbf{A}^T\mathbf{A})^{-1}\mathbf{A}^T\mathbf{r} \tag{8}$$

A unique solution exists to this equation, since **A** is of full **rank** and r is defined. Even though a solution exists, it might be very sensitive to slight perturbations in the measurements, **r,** or perturbations in the biochemistry, **A.** As a first check, the condition number of the system is calculated. The condition number is defined as

$$C(\mathbf{A}) = \|\mathbf{A}\| \cdot \|\mathbf{A}^{\#}\| \tag{9}$$

where $\| \|$ is any matrix norm and $\mathbf{A}^{\#}$ is the pseudoinverse, given as

$$A^{Y} = (\mathbf{A}^{T}\mathbf{A})^{-1}\mathbf{A}^{T} \tag{10}$$

If $C(\mathbf{A})$ is small (say <100, Euclidian norm), then the system is considered well posed; however, if $C(\mathbf{A})$ is large (say >1000, Euclidian norm), then there are possible sensitivity problems that should be investigated further (see Reference 9 for more details on the condition number). It should be noted that no measurements are needed for the determination of the condition number, indicating that modifications to the biochemistry or new networks can be tested for sensitivity problems without conducting experiments. The condition number for the lysine fermentation is **62** based on the Euclidian norm.

The sensitivity of the solution with respect to biochemistry and measurements is determined from the appropriate derivatives shown here:

$$\frac{\partial \bar{\mathbf{x}}}{\partial a_{ij}} = (\mathbf{A}^{T}\mathbf{A})^{-1}\left[\frac{\partial \mathbf{A}^{T}}{\partial a_{ji}} - \left(\frac{\partial \mathbf{A}^{T}}{\partial a_{ji}} + \mathbf{A}^{T}\frac{\partial \mathbf{A}}{\partial a_{ij}}\right)(\mathbf{A}^{T}\mathbf{A})^{-1}\mathbf{A}^{T}\right]\mathbf{r} \tag{11}$$

$$\frac{\partial \bar{\mathbf{x}}}{\partial \mathbf{r}} = (\mathbf{A}^{T}\mathbf{A})^{-1}\mathbf{A}^{T} \tag{12}$$

where $\frac{\partial \mathbf{A}}{\partial a_{ij}}$ represents an $m \times n$ matrix whose ijth element is a 1 and all others 0. These equations are similar to those presented by Boot[10] and can be used to isolate sensitivity problems. To remove such sensitivity problems, one must again either lump reactions together or remove them. Equation **12** allows one to determine which metabolites require accurate measurement, since the derivative indicates how sensitive the solution is to a particular measurement.

4. Redundancy Analysis

This analysis follows the original work of Romagnoli and Stephanopoulos[11] and its application to fermentation data discussed by Wang and Stephanopoulos.[12] When the row dimension of A, m, is greater than the column dimension, n, and the matrix is of full rank, then there are more equations than unknowns, indicating an over-determined system and, hence, the least squares solution. The $m - n$ dependent equations can be used to check the consistency of the measurements, the validity of the PSS hypothesis, and possibly other assumptions.

Refemng to the above papers for a complete description of the gross error identification methodology, the essentials of the algorithm can be summarized as follows. If the redundant equations are satisfied by the available measurements, the latter are consistent with the assumed bioreaction network structure and associated assumptions. If inconsistencies are detected, the redundant equations may be used to identify the source of gross error among the measurements and PSS assumptions. This is done by systematically eliiating, one at a time, the measurements or steady-state assumptions and using one of the redundant equations to determine the eliminated measurement. The remaining redundant equations are used to test the consistency of the resulting system of equations and measurements through the calculation of a statistically balanced consistency index.[12] A sharp decrease in the value of the consistency index resulting from the elimination of a measurement or assumption is a strong indication of the latter being a source of error, suggesting that it be dropped from further consideration.

To implement the consistency analysis described by Wang and Stephanopoulos,[12] the independent equations of the BRNE must be removed and the resulting set of redundant

equations placed in the form $\mathbf{Br} = \mathbf{O}$ via the following procedure. Through Gauss elimination, the **A** matrix can be partitioned into an $n \times n$ upper diagonal matrix, U, and an $m - n \times n$ zero matrix by multiplication of the appropriate $m \times m$ permutation matrix, E. Equation **3** can then be expressed as

$$\mathbf{EAx} = \begin{bmatrix} \mathbf{U} \\ \hline \mathbf{O} \end{bmatrix} \mathbf{x} = \mathbf{Er} = \begin{bmatrix} \mathbf{E}_1\mathbf{r} \\ \hline \mathbf{E}_2\mathbf{r} \end{bmatrix}$$

$$\mathbf{O} = \mathbf{E}_2\mathbf{r} \tag{13}$$

where E_2 is the lower part of the permutation matrix whose dimension is $m - n \times m$. Equation **13** is now in the appropriate form.

The net result of the consistency analysis is the ability to detect and isolate errors in measurements, PSS approximations, and, to a limited extent, biochemistry.

Following the redundancy analysis and consistency tests, a final check of the PSS approximation can be performed utilizing the reaction fluxes obtained from the solution of Equation **3.** These fluxes can be used to back calculate the apparent time rates of change of all intracellular metabolites. The latter can then be compared to the fluxes of all reactions producing or consuming the corresponding metabolite. PSS will be a valid approximation for a metabolite if its rate of change is found to be small compared to the fluxes of the reactions affecting that metabolite. If a PSS approximation is found to be invalid, further information is needed on that particular metabolite in the form of additional measurements or kinetic expressions. Such metabolites can also be removed from the network, thereby removing the PSS constraint, provided that this does not produce a singularity in **A.**

5. Additional Information

In general, the more knowledge that can be incorporated into the flux determination algorithm, the more reliable the flux estimates will be. Thus far, we have not taken into account the directionality of reactions. It is well known that some reactions are completely reversible, while others are considered irreversible. Also, in rare cases a maximum rate might be known for a particular reaction. In either case, directionality of a reaction, or upper bounds on reactions rates, imposes inequality constraints on the BRNE of the type

$$\mathbf{Ax} = \mathbf{r} \qquad \text{subject to} \qquad \mathbf{Cx} \geqq \mathbf{b} \tag{14}$$

where C is the constraint matrix that dictates which reactions are constrained to be greater than or equal to **b,** which is usually taken as the zero vector. This inequality constrained least squares (ICLS) problem is not amenable to solution by analytical methods and must be solved numerically. Liew[13] describes a technique for solving ICLS problems. Simply stated, a linearized equation is produced from the derivative of the quadratic least squares equation, $J = (\mathbf{Ax} - \mathbf{r})^T(\mathbf{Ax} - \mathbf{r})$, with respect to **x.** The linearized equation is solved by a modification of the simplex method, which is but one of the many ways to solve such quadratic programming problems.'' The solution to the ICLS problem is

$$\bar{\mathbf{x}} = (\mathbf{A}^T\mathbf{A})^{-1}\mathbf{A}^T\mathbf{r} + (\mathbf{A}^T\mathbf{A})^{-1}\mathbf{C}^T\mathbf{z} \tag{15}$$

where **z** is a numerically determined *k*-dimensional dual vector.

It is emphasized that the constraints should not be used to overcome an ill-posed system, but rather to fine tune the unconstrained solution to Equation 8. For example, if the solution to Equation 8 produces a slightly negative flow entering the PPP (i.e., ribulose-5-phosphate

TABLE 2
Production Rates of Extracellular Metabolites[15]

Time (h)	Production rate (mmol/h/l)				
	Glucose	Biomass	Lysine	Oxygen	CO_2
0—12	−2.32[a]	2.92	1.10	−12.5	13.4
12—24	−7.83[a]	3.66	2.28	−49.1	52.5
24—36	−14.3	2.58	1.53	−55.6	59.5
36—48	−16.1	1.25	2.38	−57.8	61.9

[a] Glucose measurement not used due to inconsistency.

producing glucose-6-phosphate), then the above constraints can be used to set this flow to zero. Ideally, the constraints should not be violated for a perfectly posed system (i.e., $z = 0$), but this is hard to achieve, and the constraints may add information that is not directly obtainable from the measurements.

For the lysine fermentation data, no constraints were violated; consequently, Equation 8 was used instead of Equation 15.

C. SUMMARY OF THE ALGORITHM

The following is a summary of the algorithm presented in the previous sections. For a particular microbe and fermentation, one constructs the primary biosynthetic pathways necessary for product synthesis and energy production, including those pathways necessary to generate the intermediate metabolites for the lumped biomass equation and those necessary to regenerate intracellular metabolites. These pathways can usually be found in the literature. The ones for the lysine fermentation are illustrated in Figure 1 and listed in the Appendix. These reactions are then placed in an equational form called the BRNE. The BRNE is then tested for singularity and sensitivity problems. Once a well-posed system is obtained (through lumping, etc.), production rates of all extracellular metabolites are measured experimentally as a function of time to determine r and are checked for consistency. This vector is then substituted into Equation 8 or 15, depending on whether system constraints are invoked, to obtain an estimate, $\bar{x}$, of the carbon flux through the primary pathways.

III. APPLICATIONS

The algorithm was tested with literature data reported by Erickson et al.[15] for a batch lysine fermentation of *Brevibacterium* (species not reported), since it has been well established[16] that glutarnic-acid-producing strains of *Brevibacterium* and *Corynebacterium* have similar biochemistry and should be classified under C. *glutamicum*. Data on the production rates of biomass, lysine, and carbon dioxide and the consumption rates of glucose and oxygen were used from that paper. These data, reported every 12 h for the first 48 h, are summarized in Table 2. When these data were checked for consistency, it was found (see Reference 12) that the calculated performance index was outside the 90% co dence level for the 0- to 12-h and 12- to 24-h time periods. When the glucose measurement was removed, the index improved to within acceptable limits; therefore, the glucose measurement was not used for the first two time periods.

Upon substitution of the corrected rate data into Equation 8, the flux map through the lysine bioreaction network of Figure 1 was generated, with the results shown in Table 3. The reaction numbers refer to the reactions given in the Appendix and correspond to the numbers listed on Figure 1.

To compare the estimate of the flux network with experimental data, we have calculated the percentage of glucose entering the PPP from the flux estimate, assuming partial recycling

TABLE 3
Estimated Fluxes for the Bioreaction Network in mmol/h/l

Reaction no.	Time periods (h) 0—12[a]	12—24[a]	24—36	36—48
1	5.39	13.0	14.2	15.8
2	1.11	5.62	8.00	8.60
3	3.70	10.2	10.9	11.9
4	8.60	22.5	23.3	25.6
5	8.14	22.0	22.9	25.4
6	7.96	21.8	22.8	25.3
7	0.00	0.00	0.090	0.135
8	0.621	11.7	15.4	14.7
9	7.04	30.7	35.7	36.7
10	3.31	14.9	17.4	18.2
11	0.407	0.318	0.297	0.366
12	−1.17	−2.42	−2.03	−3.00
13	2.79	14.6	17.2	17.7
14	5.53	29.3	34.7	35.8
15	2.83	14.6	17.2	17.9
16	2.79	14.6	17.4	18.0
17	2.56	4.96	4.17	6.17
18	0.0958	0.0721	0.106	0.145
19	4.22	7.27	5.88	6.92
20	1.62	2.69	2.27	2.58
21	2.60	4.58	3.47	4.12
22	1.35	2.36	1.87	2.22
23	1.35	2.36	1.66	1.90
24	1.25	2.23	1.44	1.67
25	11.1	41.8	47.2	49.2
26	1.40	7.31	8.59	8.84
27	1.22	2.38	1.93	2.95
28	1.17	2.42	2.03	3.00
29	1.15	2.44	2.02	2.94
30	2.91	3.67	2.72	1.45
31	40.7	184	213	226

[a] Glucose measurement not used.

of fructose-6-phosphate (F6P). It has been shown by radiolabeling experiments[17,18] that the best method for calculating the fraction of glucose entering the PPP is to assume partial recycling of F6P. This assumes that the fraction of F6P that is recycled back into the PPP is the same as the fraction of glucose that enters the PPP. From the network shown in Figure 1, the percentage of glucose entering the PPP, based on partial recycling, is given by

$$PPP\ (\%) = \frac{100(x_1 - x_2)}{x_1 + x_{23} + x_{24}} \tag{16}$$

where x_i represents the flux in reaction i, depicted in Figure 1 and listed in the Appendix. Oishi and Aida[19] have measured, using radiorespirometry,[20] the percentage of glucose entering the PPP for B. *ammoniagenes* in the stationary phase (i.e., no growth) and under high biotin concentration (20 μg/l). The *PPP* (%) calculated from the flux estimates and Equation 16 along with the measurements of *PPP (%)* of Oishi and Aida are listed in Table 4.

The conditions under which Oishi and Aida conducted their experiments correspond to

TABLE 4
Estimated **and** Measured Glucose Entering the Pentose Phosphate Pathway

	Percent Glucose Entering PPP	
Time (h)	**PPP (%)[a]**	**PPP (%)[b]**
0—12	54[c]	—
12—24	42[c]	—
24—36	35	—
36—48	37	38[d]

[a] Calculated by partial recycle of F6P, Equation 16.
[b] Datum from Oishi and Aida.[19]
[c] Glucose measurement not used.
[d] Resting cells, with 20 μg/l biotin.

the conditions that prevail during the end of the fermentation, i.e., stationary phase and high biotin concentration. From Table 4 we can see that the estimated result of 35 to 37% glucose entering the PPP corresponds extremely well with the experimentally measured value of 38% for cells in stationary phase. This indicates that the estimated flux obtained from the BRNA does represent the true state of the metabolism in the cell for the PPP.

Furthermore, there are some other interesting features of this flux network. Reaction 31 indicates the amount of excess ATP produced over that quantity theoretically required. This flow is an order of magnitude greater than any other flux in the network and represents approximately 64, 82, 85, and 86% of ATP produced for the 0- to 12-h, 12- to 24-h, 24- to 36-h, and 36- to 48-h time periods, respectively. This finding indicates that most of the glucose consumed by the cells is simply oxidized to carbon dioxide and water, and it accounts for the large flux of carbon cycling through the modified TCA cycle (i.e., the TCA cycle plus the glyoxylate shunt and the decarboxylation reaction) and the low lysine yield (20% molar). Consequently, if this flux can be redirected to oxaloacetate (OAA), the lysine yield may be increased.

For glutamic acid production it has been proposed that Reaction 11 generates the required NADPH, for Reaction 17; however, this cannot be true for lysine production since, based on the flux network, Reaction 11 is very small compared to Reaction 17. Reaction 11 in this case is not large, since α-ketoglutaric acid is produced from glutamate in the lysine reactions (see Reactions 27 and 29); consequently, NADPH, needed for Reaction 17 must be generated via the PPP, which could be another limiting reaction in the production of lysine.

It should also be noted that the decarboxylation reaction (Reaction 8) is essential when the glyoxylate shunt is functioning, which is reasonable since Reaction 8 is used to complete the modified TCA cycle, as postulated by Shiio et al.[21]

IV. SUMMARY

The primary metabolic pathways of a cell have been represented by a linear equation, termed the BRNE. Several routines have been developed to check the singularity, sensitivity, and consistency of the model, and techniques have been described as means to correct these problems when they arise, such as lumping or removal of reactions and removal of statistically inconsistent measurements.

The biochemistry of lysine fermentation for glutarnic acid bacteria has been constructed based on available literature data. With the carboxylation reactions lumped into one reaction (Reaction 8), α-ketoglutarate dehydrogenase in the TCA cycle removed, excess ATP removed via Reaction 31, and biomass synthesis represented as one reaction (Table 1), the bioreaction network matrix A is nonsingular and has a dimension of 33 × 31 and a condition number of 62. The corresponding biochemical reactions are listed in the Appendix.

The fluxes generated by the unconstrained solution, Equation 8, from Erickson's data appear reasonable, and the percentage of glucose entering the PPP calculated from the flux estimates correlates well with experimental data presented by Oishi and Aida.[19] It also appears that the low lysine yield of 20% molar might be caused by overproduction of ATP by the oxidative pathways of glutamic acid bacteria.

ACKNOWLEDGMENTS

We thank the Massachusetts Institute of Technology (MIT) Biotechnology Process Engineering Center and the National Science Foundation (PYI grant no. CBT-8514729) for their support of this research.

APPENDIX: BIOCHEMISTRY AND METABOLITES*

The following reactions are used to describe the metabolism of glutarnic acid bacteria for lysine fermentation. Also given are the important metabolites considered. From this information the bioreaction network matrix, A, is constructed.

A. BIOCHEMICAL REACTIONS

Embden-Meyerhof-Parnas Pathway

1. GLC + ATP → GLC6P + ADP
2. GLC6P ↔ FRU6P
3. FRU6P + ATP → 2 GAP + ADP
4. GAP + ADP + NAD → NADH + G3P + ATP
5. G3P ↔ PEP + H_2O
6. PEP + ADP → ATP + PYR
7. PYR + NADH ↔ LAC + NAD

Carboxylation Reactions

8. OAA → PYR + CO_2

TCA Cycle

9. PYR + COA + NAD → ACCOA + CO_2 + NADH
10. ACCOA + OAA + H_2O ↔ ISOCIT + COA
11. ISOCIT + NADP ↔ AKG + NADPH + CO,
12. SUCCOA + ADP ↔ SUC + COA + ATP
13. SUC + H_2O + FAD ↔ MAL + FADH
14. MAL + NAD ↔ OAA + NADH

Glyoxylate Shunt

15. ISOCIT ↔ SUC + GLYOX
16. ACCOA + GLYOX + H_2O → MAL + COA

* **See Appendix Section B for abbreviation designations.**

Glutamate and Glutamine Production
17. NH_3 + AKG + NADPH ↔ GLUT + H_2O + NADP
18. GLUT + NH_3 + ATP → GLUM + ADP

Pentose Phosphate Cycle
19. GLC6P + H_2O + 2 NADP → RIBU5P + CO_2 + 2 NADPH
20. RIBU5P ↔ RIB5P
21. RIBU5P ↔ XYL5P
22. XYL5P + RIB5P ↔ SED7P + GAP
23. SED7P + GAP ↔ FRU6P + E4P
24. XYL5P + E4P ↔ FRU6P + GAP

ATP Generation; P/O = 2
25. 2 NADH + O_2 + 4 ADP → 2 H_2O + 4 ATP + 2 NAD
26. 2 FADH + O_2 + 2 ADP → 2 H_2O + 2 ATP + 2 FAD

Aspartate Family
27. OAA + GLUT ↔ ASP + AKG
28. ASP + NADPH + ATP → ASA + ADP + NADP
29. ASA + PYR + NADPH + SUCCOA + GLUT → SUC + AKG + CO, + H_2O + LYS + NADP + COA

Biomass Production; Assume Mol Wt (Biomass) = 100
30. 0.0205 GLC6P + 0.00709 FRU6P + 0.0898 RIB5P + 0.0361 E4P + 0.0129 GAP + 0.15 G3P + 0.0519 PEP + 0.125 PYR + 0.327 ACCOA + 0.058 AKG + 0.107 OAA + 0.0326 LYS + 0.796 NH_3 + 0.025 GLUT + 0.025 GLUM + 3.89 ATP + 1.37 NADPH + 0.312 NAD = BIOMASS + 3.89 ADP + 1.37 NADP + 0.312 NADH

Unaccounted for ATP Requirements
31. ATP → ADP

B. LIST OF IMPORTANT METABOLITES

Each metabolite listed below corresponds to one element of the r vector.

1. Acetyl coenzyme A (ACCOA)
2. a-Ketoglutarate (AKG)
3. Aspartate semialdehyde (ASA)
4. Aspartate (ASP)
5. Adenosine 5'-triphosphate (ATP)
6. BIOMASS
7. CO_2
8. Erythrose-4-phosphate (E4P)
9. Flavin adenine dinucleotide (FADH)
10. Fructose-6-phosphate (FRU6P)
11. 3-Phosphoglycerate (G3P)
12. Glyceraidehyde-3-phosphate (GAP)
13. Glucose (GLC)
14. Glucose-6-phosphate (GLC6P)
15. Glutamine (GLUM)
16. Glutamate (GLUT)
17. Glyoxylate (GLYOX)
18. Isocitrate (ISOCIT)
19. Lactate (LAC)
20. Lysine (LYS)
21. Malate (MAL)
22. NADH
23. NADPH
24. O,
25. Oxalocetate (OAA)
26. Phosphoenolpyruvate (PEP)
27. Pyruvate (PYR)
28. Ribose-5-phosphate (RIB5P)
29. Rubulose-5-phosphate (RIBU5P)
30. Sedoheptalose-7-phosphate (SED7P)
31. Succinate (SUC)
32. Succinyl CoA (SUCCOA)
33. Xylulose-5-phosphate (XYL5P)

REFERENCES

1. Kinoshita, S. and **Nakayama,** K., Amino acids, in ***Economic Microbiology,*** Vol. 2, Rose, A. H., Ed., Academic Press, London, 1978, 209.
2. **Simon,** J., **Engasser,** J., and Germain, P., Substrate flux mapping in **Corynebacteria** for amino acids production, in ***Third European Congress on Biotechnology, Munchen,*** Vol. 2, Verlag **Chemie, Weinheim,** Federal Republic of Germany. 1984, 49.
3. Shvinka, J., **Viesturs,** U., and **Ruklisha,** M., Yield regulation of lysine in ***Brevibacterium flavum, Biorechnol. Bioeng..*** 22, 897, 1980.
4. **Ingraham,** J. L., Maaloe, O., and **Neidhardt,** F. C., ***Growrh of rhe Bacterial Cell,*** Sinauer Associates, Sunderland, MA, 1983, chap. 3.
5. **Stouthamer,** A. H., A theoretical study on the amount of ATP **required** for synthesis of microbial cell material, ***Antonie van Leeuwenhoek J. Microbiol. Serol.,*** 39, 545, 1973.
6. Papoutsakis, E. T., Equations and calculations for fermentations of **butyric** acid bacteria, ***Biorechnol. Bioeng.,*** 26, 174, 1984.
7. Papoutsakis, E. T., Equations and calculations of product yields and **preferred** pathways for **butanediol** and mixed-acid fermentations, ***Biorechnol. Bioeng..*** 27. 50, 1985.
8. Shiio, I., Significance of a-ketoglutazic **dehydrogenase** on the glutamic acid formation in ***Brevibacterium flavum, J. Biochem. (Tokyo),*** 50, 164, **1961.**
9. Noble, B. and Daniel, J. W., ***Applied Linear Algebra.*** Prentice-Hall, **Englewood** Cliffs, NJ, 1977, 330.
10. Boot, J. C. G., On sensitivity analysis in convex quadratic programming problems, ***Oper. Res..*** 11, 771, 1963.
11. Romagnoli, J. A. and Stephanopoulos, G., Rectification of process measurement data in the presence of gross errors, ***Chem. Eng. Sci.,*** 36, 1849, 1981.
12. Wang, N. S. and Stephanopoutos, G. N., Applications of macroscopic balances to the identification of gross **measurement** errors, ***Biotechnol. Bioeng..*** **25,** 2177, 1983.
13. **Liew, C. K.,** Inequality constrained least-squares estimation, **I.** ***Am. Stat. Assoc.,*** 71, 746, 1976.
14. Boot, J. C. G., Binding constraint **procedures** of quadratic programming, ***Econometrica,*** 31, 464, 1963.
15. **Erickson,** L. E., Selga, S. E., and **Viesturs,** U. E., Application of mass and energy balance regularities to product **formation,** ***Biorechnol. Bioeng.,*** 20, 1623, 1978.
16. Abe, S., **Takayama,** K., and Kinoshita, S., Taxonomical studies on glutamic acid-producing bacteria, ***J. Gen. Appl. Microbiol.,*** 13, 279, **1967.**
17. **Dawes,** E. A. and **Holms,** W. H., Metabolism of ***Sarcina lutea.*** **II.** Isotopic evaluation of the mutes of glucose utilization, ***Biochim. Biophys. Acta,*** 29, 82, 1958.
18. Shiio, I., Otsuka, S., and **Tsunoda,** T., Glutamic acid formation from glucose by bacteria. III. On the pathway of **pyruvate** formation in ***Brevibacterium flavum*** No. **2247,** ***J. Biochem. (Tokyo),*** 47, 414, 1960.
19. **Oishi,** K. and **Aida,** K., Studies on amino acid fermentation. XI. Effect of biotin on the **Embden-Meyerhof-Parnas** pathway and the **hexose-monophosphate** shunt in a glutamic acid-producing bacterium, ***Brevibacterium ammoniagenes*** 317-1, ***Agric. Biol. Chem.,*** 29, 83, 1965.
20. Wang, C. H., Stern, I., **Gilmour,** C. M., **Klungsoyr,** S., Reed, D. **J., Bialy,** J. J., **Christensen,** B. E., and **Cheldelin,** V. H., Comparative study of glucose catabolism by the **radiorespirometric** method, ***J. Bacteriol.,*** 76, 207, 1958.
21. Shiio, I., Otsuka, S., and **Takahashi,** M., Glutamic acid formation from glucose by bacteria. VI. **Metabolism** of the **intermediates** of the TCA cycle and of the glyoxylate bypass in ***Brevibacterium flavum*** No. **2247,** ***J. Biochem. (Tokyo),*** 50, ***34,*** **1961.**