BIOINFORMATICS ORIGINAL PAPER Vol. 23 no. 3 2007, pages 351–357
doi:10.1093/bioinformatics/btl619

Systems biology

Bayesian-based selection of metabolic objective functions

Andrea L. Knorr, Rishi Jain and Ranjan Srivastava*

Department of Chemical, Materials and Biomolecular Engineering, University of Connecticut, 191 Auditorium Road U3222, Storrs, CT 06269-3222, USA

Received on October 3, 2006; revised on November 28, 2006; accepted on November 29, 2006
Advance Access publication December 6, 2006
Associate Editor: Martin Bishop

ABSTRACT

Motivation: A critical component of *in silico* analysis of underdetermined metabolic systems is the identification of the appropriate objective function. A common assumption is that the objective of the cell is to maximize growth. This objective function has been shown to be consistent in a few limited experimental cases, but may not be universally appropriate. Here a method is presented to quantitatively determine the most probable objective function.

Results: The genome-scale metabolism of *Escherichia coli* growing on succinate was used as a case-study for analysis. Five different objective functions, including maximization of growth rate, were chosen based on biological plausibility. A combination of flux balance analysis and linear programming was used to simulate cellular metabolism, which was then compared to independent experimental data using a Bayesian objective function discrimination technique. After comparing rates of oxygen uptake and acetate production, minimization of the production rate of redox potential was determined to be the most probable objective function. Given the appropriate reaction network and experimental data, the discrimination technique can be applied to any bacterium to test a variety of different possible objective functions.

Contact: srivasta@engr.uconn.edu

Supplementary information: Additional files, code and a program for carrying out model discrimination are available at http://www.engr.uconn.edu/~srivasta/modisc.html.

INTRODUCTION

Amid the explosion of information now available from the publication of genome sequences, several projects have begun to create models of large-scale networks or whole-cell behavior (Kitano, 2002; Schaff and Loew, 1999; Zimmer, 2005). Additionally, many attempts have been made to model cellular metabolism, either in whole or in part (Aiba and Matsuoka, 1979; Bonarius *et al*., 1996; Edwards *et al*., 2001; Edwards and Palsson, 1999, 2000; Fell and Small, 1986; Fischer and Sauer, 2005; Forster *et al*., 2003; Goel *et al*., 1993, 1999; Holms, 1996; Ibarra *et al*., 2002; Majewski and Domach, 1990; Mavrovouniotis *et al*., 1992; Raman *et al*., 2005; Reed *et al*., 2003; Sauer *et al*., 1996, 1998; Savinell and Palsson, 1992; Schilling *et al*., 2000, 2002; Schilling and Palsson, 2000; Vallino and Stephanopoulos, 2000). Such analyses are beneficial because a variety of *in silico* experiments can be run in less time than would be necessary to carry out the same experiments using traditional laboratory methods.

*To whom correspondence should be addressed.

Flux balance analysis (Bonarius *et al*., 1997; Edwards and Palsson, 1999; Varma and Palsson, 1994) is a common method of modeling bacterial metabolism, and it was used in the current analysis. A vast number of feasible metabolic flux distributions exist, so optimization may be used to predict the phenotypic behavior of the cell subject to various constraints (Edwards and Palsson, 2000). The genome-scale metabolism of *Escherichia coli* growing on succinate was the model system analyzed, and five objective functions were compared.

This paper introduces the technique of objective function discrimination to allow researchers to determine the most probable objective function for any cell given the appropriate experimental data and reaction set. Most analyses thus far have assumed that maximization of growth is the appropriate objective function (Edwards *et al*., 2001; Edwards and Palsson, 2000; Forster *et al*., 2003; Ibarra *et al*., 2002; Reed *et al*., 2003; Schilling *et al*., 2002; Varma and Palsson, 1993, 1994). However, it has been acknowledged that optimization based on growth may not occur on all substrates (Ibarra *et al*., 2002). Others have suggested optimization (Burgard and Maranas, 2003) or heuristic approaches (Savinell and Palsson, 1992) for comparing objective function candidates. In this paper, a Bayesian approach to identify the most probable objective function was utilized. The objective functions compared were growth rate maximization, minimization of ATP production rate, maximization of ATP production rate, minimization of substrate utilization rate and minimization of the production rate of redox potential for *E.coli* growing on succinate. Note that any objective function can be considered, such as those related to minimization of metabolic adjustment (MOMA) or regulatory on/off minimization (ROOM), as long as there is experimental data available for comparison. Benefits of this rigorous quantitative approach include ease of implementation as well as extremely low computational cost. To our knowledge, this is the first comparison of objective functions for genome-scale metabolism in *E.coli*. It is also the first use of a Bayesian approach to identify potential metabolic functions.

METHODS

Modeling approach

Several mathematical methods have been developed to study metabolism, including biochemical systems theory (Savageau, 1969a,b), cybernetic modeling (Kompala *et al*., 1986), temporal decomposition (Palsson *et al*., 1987), network rigidity (Stephanopoulos and Vallino, 1991), metabolic control analysis (Fell, 1997; Kacser and Burns, 1973), metabolic pathway analysis (Liao *et al*., 1996; Schilling *et al*., 1999; Schuster *et al*., 1999), metabolic flux analysis (Follstad *et al*., 1999; Lee *et al*., 1999, 2000) and flux balance

analysis (Alper *et al*., 2005; Bonarius *et al*., 1997; Edwards and Palsson, 1999; Varma and Palsson, 1994). However, the rate constants of most metabolic systems are unknown, although there are some notable exceptions (Joshi and Palsson, 1989a,b; Joshi and Palsson, 1990a,b; Palsson *et al*., 1989). Although ^{13}C-labeled substrate may be used to measure fluxes throughout a metabolic system (Klapa *et al*., 1999, 2003; Koffas and Stephanopoulos, 2005; Park *et al*., 1999; Sauer *et al*., 1997; Stephanopoulos, 1999) and therefore reduce the degrees of freedom, often only extracellular metabolite concentrations are measured due to lack of expertise or resources. As a result, flux balance analysis was used for this research since it only requires reaction stoichiometry. Given that flux balance analysis predicts all feasible metabolic flux distributions, optimization theory may be used to find a solution that satisfies a biologically relevant objective function and its concomitant set of flux constraints.

Flux balance analysis considers all reactions in which a metabolite participates. This can be expressed as a mass balance equation of the form

$$\frac{dx_i}{dt} = \sum_j S_{ij}\phi_j, \tag{1}$$

where S_{ij} is the stoichiometric coefficient of metabolite x_i for the flux ϕ_j. Equation (1) neglects metabolite dilution due to cell growth because this effect on concentration is negligible compared to the fluxes for each reaction in which the metabolite participates (Stephanopoulos *et al*., 1998). The other major assumption of flux balance analysis is the system is at a pseudo steady-state. High-turnover allows metabolite concentrations to respond rapidly to perturbations (Edwards, 1999; Stephanopoulos *et al*., 1998), justifying this assumption. As a result, the system of ordinary differential equations is reduced to a set of linear algebraic equations:

$$\underline{\underline{S}} \cdot \underline{\phi} = \underline{0}, \tag{2}$$

where $\underline{\underline{S}}$ is the $m \times n$ matrix of stoichiometric coefficients for all m metabolites in all n fluxes, $\underline{\phi}$ is the vector of fluxes, and $\underline{0}$ is a vector of zeros. To ensure that all internal fluxes are greater than zero, all internal reversible reactions must be decomposed into forward and reverse reactions. Exchange fluxes, which are reactions involving transport of metabolites into or out of the cell, are allowed to be negative, denoting direction. For the genome-scale metabolism of *E.coli* MG1655 there are 1320 fluxes—1177 irreversible reactions and 143 exchange fluxes—that involve 625 metabolites, as specified in iJR904 (Reed *et al*., 2003).

Most metabolic systems are underdetermined due to a greater number of fluxes than metabolites; in Equation (2) $\underline{\underline{S}}$ is non-square with $n>m$. Therefore, to solve for $\underline{\phi}$, an optimization strategy, such as linear programming must be used (Savinell and Palsson, 1992). An objective function is chosen and either minimized or maximized, as appropriate, subject to a variety of constraints. If Z is the objective function, and one wants to make Z as large as possible, the linear programming problem has the following form:

Maximize Z subject to

$$\sum_{i,j} S_{ij}\phi_j = 0 \text{ and } \alpha_j \le \phi_j \le \beta_j, \tag{3}$$

where α_j and β_j are the minimum and maximum values, respectively, of the jth flux.

Five objective functions were examined in this work: maximization of growth rate, minimization of the production rate of redox potential, minimization of ATP production rate, maximization of ATP production rate and minimization of nutrient uptake rate. Growth rate was defined as the combination of various metabolites, such as amino acids and cofactors, in stoichiometric proportion to produce ATP, inorganic phosphate, pyrophosphate and water (Reed *et al*., 2003). The specific equation used is listed in the supplementary files provided with the genome-scale model iJR904 (Reed *et al*., 2003). Maximizing growth rate was chosen for examination not only because it has been utilized in the past, but also because it may be argued that a cell is best able to out-compete other cells if it can grow quickly and overtake competitors.

Minimization of redox potential was also selected as a plausible objective function because it may be interpreted as maximization of energy efficiency by the cell. The redox potential objective function included all reactions that produce $FADH_2$, NADH or NADPH (Savinell and Palsson, 1992). This objective function is expressed as follows:

Minimize{5DGLCNR_R + ABUTD + ACALDI + 2∗ADHER_R + AGPR_F + ALCD19_R + ALDD19X + ALDD2X + ASAD_F + BETALDHX + BETALDHY + DAAD + DHBD_F + DHCIND + DHPPD + DHRF_R + E4PD_F + 6∗FAO1 + 7∗FAO2 + 8∗FAO3 + FAO4 + G3PD2_F + G6PDH2R_F + GAPD_F + GCALDD + GLTPD_F + GLUDY_F + GLYCDX + GLYCL + GND + GTHOR_R + 2∗HISTD + HSDY_F + ICDHYR_F + IDOND_R + IMPD + IPMD + LCAD_F + LCADI + LCAR_R + LDH_D_F + M1PD_F + MANAO_F + MDH_F + ME1 + ME2 + MTHFD_F + P5CD + PDH + PDX5PS + PERD_F + PGCD + PPND + PROD2 + SBTPD_F + SGSAD + SHCHD2 + SHK3DR_R + SSALX + SSALY + SUCD1I + SUCD4_R + 3∗SULR_F + TAGURR_F + TEST_AKGDH + TEST_NADTRDH + THD2 + THRD + 2∗UACMAMO + 2∗UDPGD},
where the reaction name is that specified in iJR904 (Reed *et al*., 2003), and '_R' and '_F' refer to the reverse and forward reaction, respectively, for reversible reactions. A cell may minimize redox potential to decrease the number of oxidizing reactions that occur thus conserving its energy or using its energy in the most efficient way possible (Savinell and Palsson, 1992).

Minimization of ATP production was examined because this implies efficient use of energy. The goal of this objective function is for the cell to grow while using the minimum amount of energy necessary, thereby conserving ATP. As a result the cell uses ATP as efficiently as possible. Though this objective function is concerned with efficient operation, it is different than minimizing redox potential because ATP is the molecule in question instead of one of the adenine dinucleotides. This objective function includes all reactions that produce ATP (Savinell and Palsson, 1992), and appears as follows:

Minimize{ACCOACR_R + ACKR_R + ADK1_R + ALAALAR_R + AP5AH + ASPK_R + ATPS4R_F + CBIAT_R + CBLAT_R + CYTK1_R + CYTK2_R + DADK_R + DBTSR_R + DGK1_R + DHBSR_R + DTMPK_R + GALK_R + GK1_R + NDPK1_R + NDPK2_R + NDPK3_R + NDPK4_R + NDPK5_R + NDPK6_R + NDPK7_R + NDPK8_R + PGK_R + PPAKR_F + PRAGSR_R + PRASCS_R + PRPPS_R + PYK + SERASR_R + SUCOAS_R + TMKR_R + TMPKR_R + UMPK_R + URIDK2R_R}.

Once again, the reaction names are those listed in the supplementary files of iJR904 (Reed *et al*., 2003), and '_R' and '_F' refer to the reverse and forward reaction, respectively, of reversible reactions. It has been proposed that instead, the rate of ATP production should be as large as possible, signifying the production of as much ATP as possible. However, this objective function may make less sense intuitively, as alluded to in the description of minimizing the ATP production rate. Though often used as an objective function when modeling mitochondria (Ramakrishna *et al*., 2001; Vo *et al*., 2004), it might not be in a bacterial cell's best interest to produce an abundance of ATP, especially if the cell's resources could be diverted to other processes and the ATP made would be wasted excess. Nonetheless, maximization of ATP production rate was also analyzed.

Minimization of nutrient uptake was investigated as a fifth objective function. The rationale behind this hypothesis is that efficient use of the carbon source leads to a longer-lasting food supply in the case of nutrient-limiting environments. However, it can be argued that the cell may not employ this method when in nutrient-rich conditions. This objective function was implemented by optimizing the exchange flux for uptake of external nutrients (Savinell and Palsson, 1992), succinate in this case.

Objective function discrimination

The method used for objective function discrimination was adapted from a Bayesian-based technique developed by Stewart and coworkers detailed in

several previous works (Knorr and Srivastava, 2005; Stewart *et al*., 1992, 1996, 1998; Stewart and Sørenson, 1981). Using this technique, the posterior probability for each objective function was calculated and normalized to the sum of all the posterior probabilities. These normalized values were referred to as the posterior probability shares. The objective function with the highest posterior probability share was considered the most probable.

To calculate the probability share for an objective function, M_j, the posterior probability must first be determined:

$$p(M_j \mid \underline{\underline{Y}}) \propto p(M_j) 2^{-p_j/2} \mid \underline{\underline{v}}_j \mid^{-DOF/2}. \quad (4)$$

In Equation (4) M_j was objective function j; $p(M_j)$ was the prior probability of M_j; $\underline{\underline{Y}}$ was the matrix of weighted experimental data; p_j was the number of parameters estimated in M_j; and DOF was the number of degrees of freedom. For the case of flux balance analysis, the stoichiometric coefficients of all reactants were known and it was unnecessary to estimate any parameters; therefore p_j was zero in all cases (Bailey, 2001).

$\underline{\underline{v}}_j$ was the matrix of the products of the deviation of the data from the predicted value for M_j, evaluated at the maximum likelihood of the parameter vector $\underline{\theta}$. The ikth element of v_j was calculated by:

$$v_{ik}(\underline{\theta}_j) = \sum_{u=1}^{n} [Y_{iu} - F_{ji}(\underline{\xi}_u, \underline{\theta}_j)][Y_{ku} - F_{jk}(\underline{\xi}_u, \underline{\theta}_j)]. \quad (5)$$

F_{ji} was the weighted value predicted by M_j, which was a function of the vectors of independent variables, $\underline{\xi}_u$, and parameters, θ_j. The subscripts i and k indicated a specific response value, such as the acetate production rate, and u represented the event, or experiment, during which the data was collected. The experimental data, Y_{iu}, and the predicted data, F_{ji}, were weighted by the reciprocal standard deviation of the corresponding response value, i or k, as described by Stewart *et al*. (1998). Standard deviations were provided in the supplementary data of Edwards *et al*. (2001).

Each posterior probability was normalized to form the posterior probability share, π:

$$\pi(M_j \mid \underline{\underline{Y}}) = \frac{p(M_j \mid \underline{\underline{Y}})}{\sum_k p(M_k \mid \underline{\underline{Y}})}. \quad (6)$$

The objective function with the largest <*t*>π</*t*> was selected as the most probable. It should be noted that objective function discrimination can predict only which objective function is the most appropriate of those considered for the system of interest.

Experimental data

Experimental data were obtained from previously published work (Edwards *et al*., 2001) for growth of *E.coli* MG1655 on minimal M9 media supplemented with succinate rather than glucose as the carbon source. Twenty-four batch experiments were run at various temperatures (27.5–37°C) and succinate levels (0.05–4 g/l) to give different succinate uptake rates. Growth rate, oxygen uptake rate, succinate uptake rate and acetate production rate were measured for each trial.

Flux predictions

To generate flux predictions, flux balance analysis was carried out for the genome-scale metabolism of *E.coli* growing on succinate. Experimentally determined values were used as constraints to calculate the predicted distribution of metabolic fluxes for each objective function. To some extent, values that were constrained were dependent upon the objective function being evaluated. For example, cell growth was fixed when substrate uptake was minimized. However, when evaluating maximization of cell growth as the objective function, substrate uptake was fixed. Both substrate uptake and cell growth were fixed in all other cases.

RESULTS

Five objective functions were examined for the behavior of *E.coli* on succinate using the reactions involved in genome-scale metabolism as detailed in the supplementary files of iJR904 (Reed *et al*., 2003). Linear programming and flux balance analysis were utilized to predict fluxes for all reactions using the GNU Linear Programming Solver software. The predictions for rates of acetate production and oxygen uptake were compared to experimental data of Edwards *et al*. (2001) for growth on succinate. Each objective function generated a flux vector of predicted behavior, and comparisons to experimental data are shown in Figures 1 and 2. Experiments 1–7 and 10 from Edwards *et al*. (2001) were not included in the current study because solutions to the flux balance analysis problem could not be found when minimization of ATP production rate, maximization of ATP production rate, or minimization of the production rate of redox potential was the objective function. A potential reason why solutions could not be found for some experiments include measurement error leading to points that lie outside the phenotype phase plane.

Figure 1 shows measured and predicted acetate production rates. Maximization of ATP predicted the results of experiments 19–22 and 24 fairly accurately, but gave low-values for experiments 11, 14, 18 and 23. In addition, acetate production was predicted in experiments 12, 13 and 17, though none was observed. Values of zero were predicted for experiments 8 and 9, failing to reproduce the acetate production observed experimentally. Minimization of ATP incorrectly predicted acetate export for experiments 12, 13 and 15–17, although none is observed experimentally. While this objective function was the only one to predict acetate production in experiments 8 and 9, the predictions were more than 2- and 4-fold greater, respectively, than the true values. The prediction for experiment 21 was very close to the measured value, and experiments 23 and 24 were predicted exactly. Maximization of growth rate poorly predicted acetate production; only two experiments yielded nonzero predictions, and only that of experiment 24 was close to the true value. Minimization of redox incorrectly predicted acetate production for experiments 12, 13 and 17, as there was no measured acetate production in all cases. Further, though acetate production was predicted in experiment 11, this value was <6% of the observed value. Predictions for experiments 19–21, 23 and 24 were close to the measured vales, but predictions for experiments 14, 18 and 22 were not. Interestingly, minimization of succinate uptake rate never predicted production of acetate.

Figure 2 shows values predicted by all objective functions for oxygen uptake rate. Maximization of ATP production and minimization of the production rate of redox potential generally exhibited the same behavior: oxygen uptake rate was over estimated in experiments 13, 14, 17–24, under estimated in experiments 8, and 16, and very close to the true value in experiments 9, 11, 12 and 15. Minimization of ATP production rate under estimated oxygen uptake rate for every experiment with the exception of experiment 20. Maximization of growth under predicted the oxygen uptake rate in every experiment except for 21, and results for experiments 19–23 were very close to the observed values. Minimization of succinate uptake rate greatly under estimated the value measured in every experiment. Using the objective function discrimination technique described here, all the information in Figures 1 and 2 was combined to yield the most probable objective function.

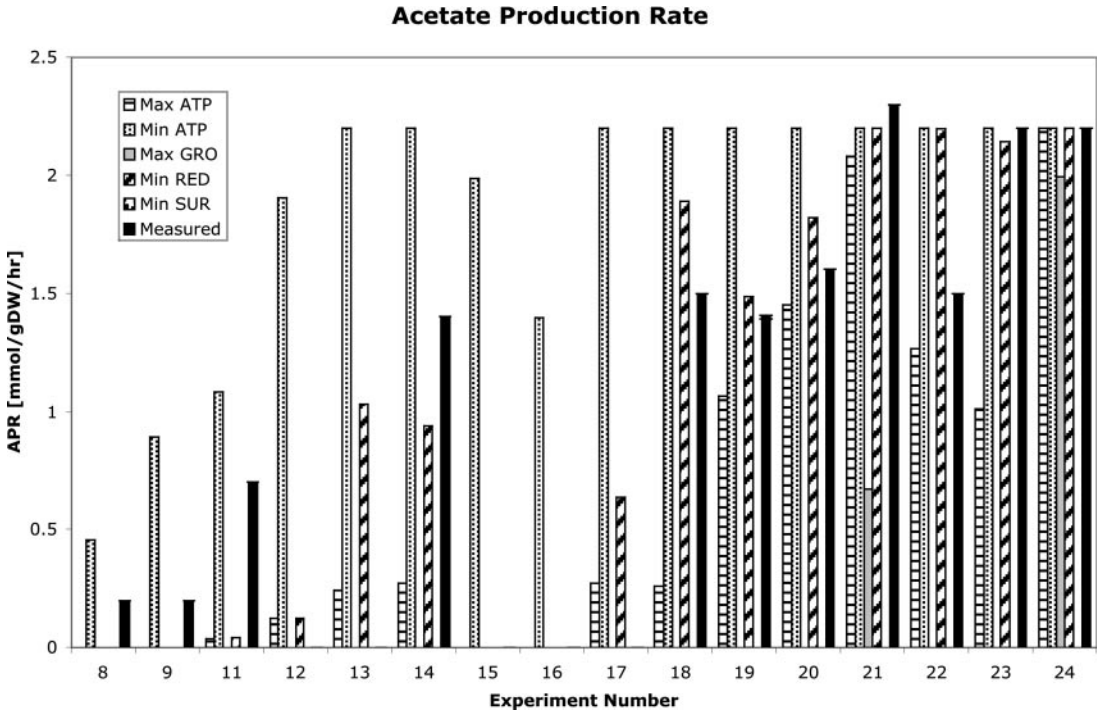

Fig. 1. Acetate production rate as predicted by each of the five objective functions: 'Max ATP' and 'Min ATP' are maximizing and minimizing ATP, respectively; 'Max GRO' is maximization of growth rate; 'Min RED' is the minimization of the production of redox potential; 'Min SUR' is minimization of succinate uptake rate; 'Measured' are the measurements from Edwards *et al*. (2001). Predictions varied widely and were often wrong.

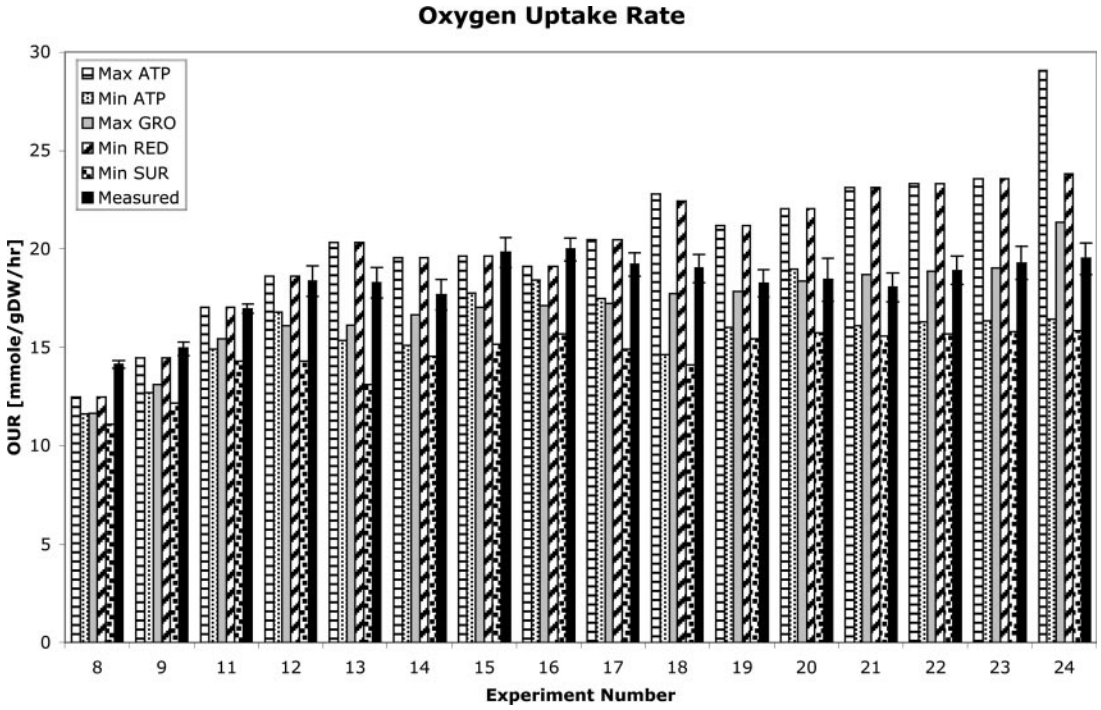

Fig. 2. Oxygen uptake rate as predicted by each of the five objective functions: 'Max ATP' and 'Min ATP' are maximizing and minimizing ATP, respectively; 'Max GRO' is maximization of growth rate; 'Min RED' is the minimization of the production of redox potential; 'Min SUR' is minimization of succinate uptake rate; 'Measured' are the measurements from Edwards *et al*. (2001).

Table 1. Posterior probability share, $\pi(M_j \mid Y)$, of each objective function, listed in descending order of probability

Objective function	$\pi(M_j \mid Y)$
Minimize production of redox	9.97×10^{-1}
Maximize growth	2.47×10^{-3}
Maximize ATP production	1.86×10^{-4}
Minimize succinate uptake	6.55×10^{-9}
Minimize ATP production	1.32×10^{-9}

Interestingly, minimization of the production rate of redox potential was the most probable, with growth rate maximization the second most probable, as shown in Table 1. Although this result is surprising, given the prominence in the literature of using maximization of growth rate as the objective function, the results in Table 1 may be due to constraining only a small number of the 1320 total fluxes. A noteworthy result is that maximizing ATP production had a higher posterior probability share than minimizing ATP production, which was the opposite of what was expected. The more efficient objective function of minimizing ATP production was less probable than maximizing ATP production even though the most probable objective function overall was based on energy efficiency. Further, efficient use of resources, in the form of succinate, yielded poor results.

DISCUSSION

Using a Bayesian-based objective function discrimination method, it was quantitatively demonstrated that minimizing the production rate of redox potential was a more probable objective function than maximizing growth rate as well as three other objective functions for genome-scale metabolism of *E.coli* growing on succinate. This result was unexpected since maximization of growth is the most-widely used objective function for flux balance analysis.

Figures 1 and 2 show the experimentally measured fluxes and the fluxes predicted by each objective function for 16 experiments. Maximization of growth rate often predicted no acetate production though it was observed experimentally. Minimization of the production rate of redox potential predicted values of acetate production that were close to experimental values for Experiments 14 and 18–24. Minimizing the production rate of redox potential had closer predictions for oxygen uptake rate than maximizing growth rate for Experiments 8–13 and 15–17, but maximizing growth rate was the better predictor over the entire range of data. By taking into consideration all the data from Figures 1 and 2, objective function discrimination predicted that minimization of redox production was the most probable objective function for genome-scale metabolism of *E.coli* growing on succinate. These results may change if additional fluxes are measured experimentally and used as constraints. Further, while alternate optimal solutions to the linear programming problem exist (Lee *et al*., 2000; Reed and Palsson, 2004), the alternate optima that have been found for *E.coli* growing on succinate do not have different exchange fluxes, and therefore the results presented here are not affected by the alternate optima as described previously in the literature (Reed and Palsson, 2004).

It should be noted that other methods have been proposed to determine the most appropriate objective function for modeling bacterial growth. Burgard and Maranas have developed a method of calculating what were termed the coefficients of importance for various fluxes (Burgard and Maranas, 2003). The coefficients of importance are a measure of how much a particular flux contributes to an assumed objective function. Calculating the coefficients of importance requires all metabolic fluxes to be known, which was not the case for the experimental system (Edwards *et al*., 2001). Objective function discrimination is an attractive alternative to coefficients of importance because results can be generated after measurement of only a few fluxes. Combining coefficients of importance with objective function discrimination, in the form of a non-uniform prior probability for example, may yield further insight in future studies. In the current study the prior probability, $p(M_j)$, was assumed equal for all objective functions. However, a coefficient of importance analysis may suggest a different $p(M_j)$ for each objective function, which would weight their posterior probability in Equation (4). Such weighting would subsequently affect the posterior probability shares calculated in Equation (5).

Savinell and Palsson compared the accuracy of three objective functions for hybridoma cells using a heuristic analysis (Savinell and Palsson, 1992). Results were generated for minimizing ATP production rate, minimizing nutrient uptake rate and minimizing the production rate of redox potential. For each objective function, Savinell and Palsson compared which nutrients were used, mass yields, redox production, oxygen uptake and ATP production. Interestingly, after taking all this information into consideration, the authors decided that minimizing redox was a realistic and appropriate cellular objective. Unlike objective function discrimination where a quantitative criterion, the posterior probability share, was used to select an objective function, the heuristic approach was somewhat subjective and qualitative, as multiple pieces of information were considered to reach a conclusion.

One may wonder what the most appropriate objective function is for their system of interest. Objective function discrimination is convenient because it can be applied to any bacterium and can compare any plausible objective function, regardless whether the objective function is linear, nonlinear, mixed-integer, etc. The only requirements are a set of metabolic reactions, which will be used in linear programming to generate predicted fluxes, and experimentally measured fluxes for comparison. Of course, as more data are gathered, a more comprehensive analysis can be made. In addition, instead of using a heuristic approach to choose the best objective function (Savinell and Palsson, 1992), a single value, the posterior probability share, is used. Although objective function discrimination is a very useful technique, one must remember that it only considers the reactions and objective functions specified by the user. It cannot consider additional reactions or objective functions without doing a separate analysis. The caveat to this approach is if a group of poor objective functions is analyzed, the discrimination technique will still calculate the best of the poor objective functions. However, this does not change the fact that the selected objective function is ultimately a poor one.

In conclusion, minimization of the production rate of redox potential has been shown to be the most probable objective function for the genome-scale metabolism of *E.coli* growing on succinate. This is the first attempt to compare objective functions using genome-scale metabolism, and it is also the first to use a probabilistic approach. Because a small number of fluxes were constrained, further study is vital for validating or disproving the

results presented here. Additional experimental data would allow more fluxes to be constrained, thereby providing a more accurate depiction of cellular behavior.

ACKNOWLEDGEMENTS

Conflict of Interest: none declared.

REFERENCES

Aiba,S. and Matsuoka,M. (1979) Identification of metabolic model: citrate production from glucose by *Candida lipolytica*. *Biotechnol. Bioeng.*, **21**, 1373–1386.

Alper,H. *et al.* (2005) Identifying gene targets for the metabolic engineering of lycopene biosynthesis in *Escherichia coli*. *Metab. Eng.*, **7**, 155–164.

Bailey,J.E. (2001) Complex biology with no parameters. *Nat. Biotechnol.*, **19**, 503–504.

Bonarius,H.P.J. *et al.* (1996) Metabolic flux analysis of hybridoma cells in different culture media using mass balances. *Biotechnol. Bioeng.*, **50**, 299–318.

Bonarius,H.P.J. *et al.* (1997) Flux analysis of underdetermined metabolic networks: the quest for the missing constraints. *Trends Biotechnol.*, **15**, 308–314.

Burgard,A.P. and Maranas,C.D. (2003) Optimization-based framework for inferring and testing hypothesized metabolic objective functions. *Biotechnol. Bioeng.*, **82**, 670–677.

Edwards,J.S. (1999) *Functional Genomics and the Computational Analysis of Bacterial Metabolism*. Department of Bioengineering, University of California, San Diego.

Edwards,J.S. *et al.* (2001) *In silico* predictions of *Escherichia coli* metabolic capabilities are consistent with experimental data. *Nat. Biotechnol.*, **19**, 125–130.

Edwards,J.S. and Palsson,B.O. (1999) Systems properties of the *Haemophilus influenzae* Rd metabolic genotype. *J. Biol. Chem.*, **274**, 17410–17416.

Edwards,J.S. and Palsson,B.O. (2000) The *Escherichia coli* MG1655 *in silico* metabolic genotype: its definition, characteristics, and capabilities. *Proc. Natl Acad. Sci. USA*, **97**, 5528–5533.

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

Fell,D.A. and Small,J.R. (1986) Fat synthesis in adipose tissue. An examination of stoichiometric constraints. *Biochem. J.*, **238**, 781–786.

Fischer,E. and Sauer,U. (2005) Large-scale *in vivo* flux analysis shows rigidity and suboptimal performance of *Bacillus subtilis* metabolism. *Nat. Genet.*, **37**, 636–640.

Follstad,B.D. *et al.* (1999) Metabolic flux analysis of hybridoma continuous culture steady state multiplicity. *Biotechnol. Bioeng.*, **63**, 675–683.

Forster,J. *et al.* (2003) Genome-scale reconstruction of the *Saccharomyces cerevisiae* metabolic network. *Genome Res.*, **13**, 244–253.

Goel,A. *et al.* (1993) Analysis of metabolic fluxes in batch and continuous cultures of *Bacillus subtilis*. *Biotechnol. Bioeng.*, **42**, 686–696.

Goel,A. *et al.* (1999) Metabolic fluxes, pools, and enzyme measurements suggest a tighter coupling of energetics and biosynthetic reactions associated with reduced pyruvate kinase flux. *Biotechnol. Bioeng.*, **64**, 129–134.

Holms,H. (1996) Flux analysis and control of the central metabolic pathways in *Escherichia coli*. *FEMS Microbiol. Rev.*, **19**, 85–116.

Ibarra,R.U. *et al.* (2002) *Escherichia coli* K-12 undergoes adaptive evolution to achieve *in silico* predicted optimal growth. *Nature*, **420**, 186–189.

Joshi,A. and Palsson,B.O. (1989a) Metabolic dynamics in the human red cell. Part I—A comprehensive kinetic model. *J. Theor. Biol.*, **141**, 515–528.

Joshi,A. and Palsson,B.O. (1989b) Metabolic dynamics in the human red cell. Part II—Interactions with the environment. *J. Theor. Biol.*, **141**, 529–545.

Joshi,A. and Palsson,B.O. (1990a) Metabolic dynamics in the human red cell. Part III—Metabolic reaction rates. *J. Theor. Biol.*, **142**, 41–68.

Joshi,A. and Palsson,B.O. (1990b) Metabolic dynamics in the human red cell. Part IV—Data prediction and some model computations. *J. Theor. Biol.*, **142**, 69–85.

Kacser,H. and Burns,J.A. (1973) The control of flux. *Symp. Soc. Exp. Biol.*, **27**, 65–104.

Kitano,H. (2002) Systems biology: a brief overview. *Science*, **295**, 1662–1664.

Klapa,M.I. *et al.* (2003) Systematic quantification of complex metabolic flux networks using stable isotopes and mass spectrometry. *Eur. J. Biochem.*, **270**, 3525–3542.

Klapa,M.I. *et al.* (1999) Metabolite and isotopomer balancing in the analysis of metabolic cycles: I. Theory. *Biotechnol. Bioeng.*, **62**, 375–391.

Knorr,A.L. and Srivastava,R. (2005) Evaluation of HIV-1 kinetic models using quantitative discrimination analysis. *Bioinformatics*, **21**, 1668–1677.

Koffas,M. and Stephanopoulos,G. (2005) Strain improvement by metabolic engineering: lysine production as a case study for systems biology. *Curr. Opin. Biotechnol.*, **16**, 361–366.

Kompala,D.S. *et al.* (1986) Investigation of bacterial growth on mixed substrates: experimental evaluation of cybernetic models. *Biotechnol. Bioeng.*, **28**, 1044–1055.

Lee,K. *et al.* (2000) Metabolic flux analysis of postburn hepatic hypermetabolism. *Metab. Eng.*, **2**, 312–327.

Lee,K. *et al.* (1999) Metabolic flux analysis: a powerful tool for monitoring tissue function. *Tissue. Eng.*, **5**, 347–368.

Lee,S. *et al.* (2000) Recursive MILP model for finding all the alternate optima in LP models for metabolic networks. *Comput. Chem. Eng.*, **24**, 711–716.

Liao,J.C. *et al.* (1996) Pathway analysis, engineering, and physiological considerations for redirecting central metabolism. *Biotechnol. Bioeng.*, **52**, 129–140.

Majewski,R.A. and Domach,M.M. (1990) Simple constrained-optimization view of acetate overflow in *E.coli*. *Biotechnol Bioeng.*, **35**, 732–738.

Mavrovouniotis,M. *et al.* (1992) Synthesis of biochemical production routes. *Comp. Chem. Eng.*, **16**, 605–619.

Palsson,B.O. *et al.* (1987) Reducing complexity in metabolic networks: making metabolic meshes manageable. *Fed. Proc.*, **46**, 2485–2489.

Palsson,B.O. *et al.* (1989) Computer model of human erythrocyte metabolism. *Prog. Clin. Biol. Res.*, **319**, 133–150, 151–154.

Park,S.M. *et al.* (1999) Metabolite and isotopomer balancing in the analysis of metabolic cycles: II. Applications. *Biotechnol. Bioeng.*, **62**, 392–401.

Ramakrishna,R. *et al.* (2001) Flux-balance analysis of mitochondrial energy metabolism: consequences of systemic stoichiometric constraints. *Am J. Physiol. Regul. Integr. Comp. Physiol.*, **280**, R695–R704.

Raman,K. *et al.* (2005) Flux balance analysis of mycolic acid pathway: targets for anti-tubercular drugs. *PLoS Comput. Biol.*, **1**, e46.

Reed,J.L. and Palsson,B.O. (2004) Genome-scale in silico models of *E.coli* have multiple equivalent phenotypic states: assessment of correlated reaction subsets that comprise network states. *Genome Res.*, **14**, 1797–1805.

Reed,J.L. *et al.* (2003) An expanded genome-scale model of *Escherichia coli* K-12 (iJR904 GSM/GPR). *Genome Biol.*, **4**, R54.

Sauer,U. *et al.* (1998) Metabolic capacity of *Bacillus subtilis* for the production of purine nucleosides, riboflavin, and folic acid. *Biotechnol. Bioeng.*, **59**, 227–238.

Sauer,U. *et al.* (1997) Metabolic fluxes in riboflavin-producing *Bacillus subtilis*. *Nat. Biotechnol.*, **15**, 448–452.

Sauer,U. *et al.* (1996) Physiology and metabolic fluxes of wild-type and riboflavin-producing *Bacillus subtilis*. *Appl. Environ. Microbiol.*, **62**, 3687–3696.

Savageau,M.A. (1969a) Biochemical systems analysis. I. Some mathematical properties of the rate law for the component enzymatic reactions. *J. Theor. Biol.*, **25**, 365–369.

Savageau,M.A. (1969b) Biochemical systems analysis. II. The steady-state solutions for an n-pool system using a power-law approximation. *J. Theor. Biol.*, **25**, 370–379.

Savinell,J.M. and Palsson,B.O. (1992) Network analysis of intermediary metabolism using linear optimization. I. Development of mathematical formalism. *J. Theor. Biol.*, **154**, 421–454.

Schaff,J. and Loew,L.M. (1999) The virtual cell. *Pac. Symp. Biocomput.*, 228–239.

Schilling,C.H. *et al.* (2002) Genome-scale metabolic model of *Helicobacter pylori* 26695. *J. Bacteriol.*, **184**, 4582–4593.

Schilling,C.H. *et al.* (2000) Combining pathway analysis with flux balance analysis for the comprehensive study of metabolic systems. *Biotechnol. Bioeng.*, **71**, 286–306.

Schilling,C.H. and Palsson,B.O. (2000) Assessment of the metabolic capabilities of *Haemophilus influenzae* Rd through a genome-scale pathway analysis. *J. Theor. Biol.*, **203**, 249–283.

Schilling,C.H. *et al.* (1999) Metabolic pathway analysis: basic concepts and scientific applications in the post-genomic era. *Biotechnol. Prog.*, **15**, 296–303.

Schuster,S. *et al.* (1999) Detection of elementary flux modes in biochemical networks: a promising tool for pathway analysis and metabolic engineering. *Trends Biotechnol.*, **17**, 53–60.

Stephanopoulos,G. (1999) Metabolic fluxes and metabolic engineering. *Metab. Eng.*, **1**, 1–11.

Stephanopoulos,G. and Vallino,J.J. (1991) Network rigidity and metabolic engineering in metabolite overproduction. *Science*, **252**, 1675–1681.

Stephanopoulos,G.N., Aristidou,A.A. and Nielsen,J. (1998) *Metabolic Engineering*. San Diego, Academic Press.

Stewart,W.E. *et al.* (1992) Parameter estimation from multiresponse data. *AIChE J.*, **38**, 641–650.

Stewart,W.E. *et al.* (1996) Model discrimination and criticism with single-response data. *AIChE J.*, **42**, 3055–3062.

Stewart,W.E. *et al*. (1998) Discrimination and goodness of fit of multiresponse mechanistic models. *AIChE J*., **44**, 1404–1412.

Stewart,W.E. and Sørenson,J.P. (1981) Bayesian estimation of common parameters from multiresponse data with missing observations. *Technometrics*, **23**, 131–141.

Vallino,J.J. and Stephanopoulos,G. (2000) Metabolic flux distributions in *Corynebacterium glutamicum* during growth and lysine overproduction. Reprinted from *Biotechnology and Bioengineering*, 41, 633–646 (1993). *Biotechnol. Bioeng*., **67**, 872–885.

Varma,A. and Palsson,B.O. (1993) Metabolic capabilities of *Escherichia-coli* 2. Optimal-growth patterns. *J. Theor. Biol*., **165**, 503–522.

Varma,A. and Palsson,B.O. (1994) Metabolic flux balancing—basic concepts, scientific and practical use. *Biotechnology*, **12**, 994–998.

Vo,T.D. *et al*. (2004) Reconstruction and functional characterization of the human mitochondrial metabolic network based on proteomic and biochemical data. *J. Biol. Chem*., **279**, 39532–39540.

Zimmer,C. (2005) *Building a Virtual Microbe, Gene by Gene*. The New York Times. New York.