Journal of Bioscience and Bioengineering
VOL. 107 No. 1, 84–89, 2009

ELSEVIER

www.elsevier.com/locate/jbiosc

Maximum entropy decomposition of flux distribution at steady state to elementary modes

Quanyu Zhao and Hiroyuki Kurata*

Department of Bioscience and Bioinformatics, Kyushu Institute of Technology, 680-4 Kawazu, Iizuka, Fukuoka, 820-8502, Japan

Received 29 July 2008; accepted 6 September 2008

Enzyme Control Flux (ECF) is a method of correlating enzyme activity and flux distribution. The advantage of ECF is that the measurement integrates proteome data with metabolic flux analysis through Elementary Modes (EMs). But there are a few methods of effectively determining the Elementary Mode Coefficient (EMC) in cases where no objective biological function is available. Therefore, we proposed a new algorithm implementing the maximum entropy principle (MEP) as an objective function for estimating the EMC. To demonstrate the feasibility of using the MEP in this way, we compared it with Linear Programming and Quadratic Programming for modeling the metabolic networks of Chinese Hamster Ovary, *Escherichia coli*, and *Saccharomyces cerevisiae* cells. The use of the MEP presents the most plausible distribution of EMCs in the absence of any biological hypotheses describing the physiological state of cells, thereby enhancing the prediction accuracy of the flux distribution in various mutants.

[**Key words:** Maximum entropy principle; Metabolic flux analysis; Enzyme control flux; Elementary mode; Linear programming; Quadratic programming]

A cell can be thought of as a highly efficient factory producing biomass, energy, secondary metabolites, and other macromolecules. The biochemical synthesis of target compounds can be analyzed and optimized by metabolic pathway analysis or Flux Balance Analysis (FBA) (1). FBA optimizes a special objective function to predict flux distributions under the constraints of a stoichiometric matrix for metabolic networks (2, 3). Metabolic pathway analysis is performed based on Elementary Modes (EMs) or extreme pathways (4). The development of high-throughput technologies has increased the types of data available for such analyses. These include gene expression, enzyme activity, flux distribution, and the intracellular concentration of metabolites (5). Integrating these data will help describe the mechanisms controlling various cellular behaviors (6, 7).

EMs include all of the possible and non-decomposable pathways involved in any given metabolic network. The determined flux distribution is the non-negative linear combination of these irreversible EMs. Elementary Mode Coefficients (EMCs) are considered as the contribution that various EMs have on diverse physiological states (8, 9). The Enzyme Control Flux (ECF) was the first means of correlating enzyme activities and flux distributions (10). ECF allows for the integration of proteome data with metabolic flux analysis through EMs.

Each set of EMCs must be determined to estimate any given flux distribution by ECF analysis. However, the solution is complex, because the many different EMs vary with the objective functions used. Certain algorithms can be used to estimate the EMCs from the experimental flux distribution. With the α-spectrum method (11), every EMC is maximized and minimized to represent the available ranges of each EMC, thus the EMC is not exclusively determined. The solutions obtained for each maximum and minimum EMC are averaged to get the statistical mean value using linear programming denoted as ECFLP (Enzyme Control Flux' Linear Programming) (10).

The problem is that the ECFLP has neither a biological nor a theoretical background, despite its use for EMC estimation. Therefore, other methods have been attempted. While it is difficult to obtain solutions when negative values are provided for the irreversible modes, the Moore–Penrose generalized inverse was used for Poolman's algorithm (12, 13). When the Linear Programming (LP) method was used, the maximum biomass or specific metabolite formation was selected as the objective function (8). The objective functions relate to the optimum physiological states, but these are still not known for many organisms. It is possible that Quadratic Programming (QP) could optimize EMCs by defining the objective function as the minimal norm of the EMCs, but there is neither a physical nor a biological background behind this method (9).

The maximum entropy principle (MEP) is derived from Shannon's information theory and is widely used in physics, chemistry, and bioinformatics for gene expression (14) and sequence analysis (15, 16). However, the MEP has rarely been implemented in metabolic flux analysis. Here, we demonstrate the feasibility of using the MEP and a

* Corresponding author. Tel./fax: +81 948 29 7828.
E-mail address: kurata@bio.kyutech.ac.jp (H. Kurata).

1389-1723/$ - see front matter © 2008, The Society for Biotechnology, Japan. All rights reserved.
doi:10.1016/j.jbiosc.2008.09.011

new nonlinear programming method to optimize EMCs. We used this method to predict (I) the intracellular flux distribution caused by external fluxes and (II) the flux distribution in cellular mutants by integrating enzyme activities with the ECF.

MATERIALS AND METHODS

Shannon's MEP for elementary mode analysis For elementary mode analysis of metabolic networks, a flux distribution is generally denoted as follows:

$$\mathbf{v} = \mathbf{P} \cdot \lambda, \tag{1}$$

where, **P** is the elementary mode matrix in which the rows represent the reactions, and the columns correspond to the elementary modes; λ is the EMC vector and ***v*** is the flux vector. **P** is directly derived from the associated stoichiometric matrix (4).

Shannon's Entropy (I) is defined as:

$$I = -\sum_{i=1}^{k} \rho_i \log \rho_i \tag{2}$$

where, ρ_i is the probability and $\sum_{i=1}^{k} \rho_i = 1$; i is the number of random events and k is the total number of random events.

The central problem is defining the probability of random events to apply Shannon's Entropy to the MEP. EMs include all possible pathways and each EM could be regarded as a random event. Next, a suitable calculation assessing the probability of EM should be made. Each EM excluding internal loops has an uptake reaction. From Eq. (1), the flux of a substrate uptake reaction should be calculated as, $\frac{\sum_i p_{\text{substrate uptake},\, i} \cdot \lambda_i}{v_{\text{substrate uptake}}} = 1$.

Based on Eqs. 1 and 2, the probability of the i-th EM in Shannon's entropy is provided as follows:

$$\rho_i = \frac{1}{v_{\text{substrate uptake}}} p_{\text{substrate uptake},\, i} \cdot \lambda_i, \tag{3}$$

where, $p_{\text{substrate uptake},\, i}$ is the element for the substrate uptake reaction in the i-th EM in matrix **P**, and $v_{\text{substrate uptake}}$ is the flux of substrate uptake.

To apply the maximum entropy principle to the optimization of EMCs, we defined a new algorithm as,

$$Max - \sum_{i=1}^{ne} \rho_i \cdot \ln \rho_i \left(p_{\text{substrate uptake},\, i} \neq 0\right), \tag{4}$$

$$\text{subject to } P_d \cdot \lambda = v_d \tag{5}$$

$$\lambda > 0 \tag{6}$$

where, ne is the total number of EMs. The low boundary of the EMC was set to 10^{-9}. $\boldsymbol{v}_d$ is the vector whose fluxes are to be determined and P_d is the EM sub-matrix that consists of rows corresponding to the determined fluxes.

All the metabolic network models were constructed using CADLIVE (17) and the EMs were calculated by CellNetAnalyzer (18). The nonlinear optimization was performed in Matlab (Mathworks Inc., Natick, MA, USA) using the function fmincon.

Other objective functions We used LP, QP, and ECFLP as control methods to estimate the value of the EMCs (λ_{optimal}). The objective functions for these are provided below under the constraint equations (5,6). The objective function for LP (8) is defined as follows:

$$Max\, v_{\text{biomass}} = \sum_{i=1}^{ne} p_{\text{biomass},\, i} \cdot \lambda_i, \tag{7}$$

where, v_{biomass} is the flux for biomass formation, $p_{\text{biomass},i}$ is the element in the biomass formation reaction in the i-th EM.

TABLE 1. Intracellular flux prediction errors using LP, QP, and MEP

Models	Conditions	QP	LP	MEP	ECFLP	Ref.
CHO	μ=0.69 d^{-1}	1.14	–	0.64	0.69	(19)
	Anaerobic	11.01	33.03	2.95	15.24	(20)
E. coli	Aerobic	47.63	49.62	17.98	25.98	
	μ=0.15 h^{-1}	42.19	50.65	20.66	14.92	(21)
S. cerevisiae	μ=0.30 h^{-1}	28.60	34.67	6.85	10.30	
	μ=0.40 h^{-1}	22.70	14.31	9.34	5.92	
B. subtilis	μ=0.42 h^{-1}	17.59	17.25	13.84	20.15	(22)
	Wild-type					

μ is the specific growth rate for CHO, *B. subtilis*, and *S. cerevisiae*.

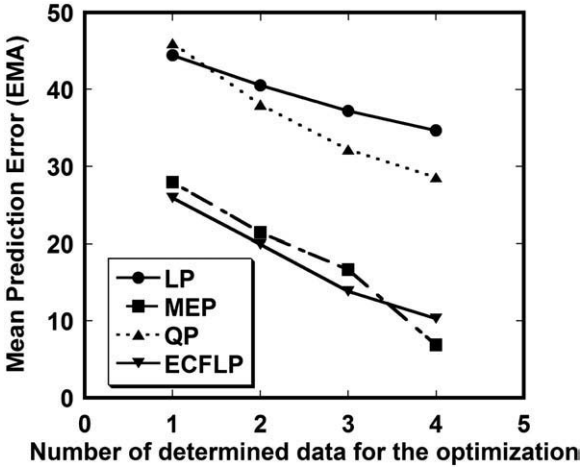

FIG. 1. Mean prediction errors with respect to the number of determined data for *S. cerevisiae* (specific growth rate, μ=0.30 h^{-1}), as optimized by the Maximum Entropy Principle (MEP), Quadratic Programming (QP), Linear Programming (LP, where the objective function was for maximum biomass formation), and Linear Programming for the ECF (ECFLP).

The objective function of QP (9) is given as follows:

$$Min \sum_{i=1}^{ne} \lambda_i^2, \tag{8}$$

The objective function of ECFLP (10) is provided as follows:

$$Max/Min\ \lambda_i\ (i = 1, 2, \ldots, ne) \tag{9}$$

The optimized EMCs are defined as the mean of the maximum and minimum EMC values.

EMA-based prediction of intracellular fluxes from the external ones To characterize the feasibility of using our MEP method, the internal flux distributions are predicted by extracellular fluxes. Using the objective functions, the EMCs for a metabolic network are optimized or estimated only from the determined fluxes. The fluxes are provided as follows:

$$\mathbf{v}_{prediction} = \mathbf{P} \cdot \lambda_{optimal} \tag{10}$$

where, λ_{optimal} is the optimized EMC vector and $\boldsymbol{v}_{\text{prediction}}$ is the estimated flux vector.

The prediction accuracy of these objective functions was evaluated by the prediction error:

$$\text{Prediction error(EMA)} = \sqrt{\frac{1}{n}\sum_{i=1}^{n} \left(v_{\text{prediction},i} - v_{\text{exp},i}\right)^2}, \tag{11}$$

where, n is the number of fluxes in each model, and $v_{\text{prediction},i}$ and $v_{\text{exp},i}$ are the predicted and experimental fluxes for the i-th reaction, respectively. We compared the MEP method with the other algorithms, where the objective functions are the maximum biomass formation (LP) (8), the minimal norm of EMCs (QP) (9), and ECFLP (10).

The external fluxes of glucose, lactate, glutamine, and alanine were used for the optimization of CHO cells (19). Several intracellular fluxes in CHO cells were determined and others were calculated by metabolic flux analysis. The fluxes were determined by ^{13}C tracer experiments in other models (20–21). The fluxes of glucose and acetate, and glucose, acetate, ethanol, lactate and succinate were used for the aerobic and anaerobic conditions of *E. coli*, respectively (20). The determined fluxes for *S. cerevisiae* included glucose, ethanol, acetate, and glycerol (21). The uptake flux of glucose and the external fluxes of riboflavin, acetoin, acetate, and lactate were used for the EMC optimization of *Bacillus subtilis* (22). Details of the network data are shown in Supplementary data.

ECF-predicted flux distribution for mutants ECF integrates enzyme activity data into the EMCs to estimate the flux distribution of mutants. Details of the algorithm are described elsewhere (10). Briefly, the EMCs for the wild-type cells were optimized by LP, QP, MEP, and ECFLP (Eqs. 4–9). Then, the EMCs of mutants were calculated as follows:

$$\lambda_i^{\text{mutant}} = \gamma \cdot \lambda_i^{w} \prod_{j=1}^{ne} a_{j,i} \tag{12}$$

where, λ_i^w is the i-th EMC for wild-type, which corresponds to the i-th element of λ_{optimal} provided by Eqs. 4–9, and $\lambda_i^{\text{mutant}}$ is that for a mutant. γ is the parameter to adjust the flux of uptake reaction of a mutant to the determined value for wild-type. $\alpha_{j,i}$

is the parameter for the enzyme activity of the j-th reaction, as defined by the following equation.

$$\alpha_{j,i} = \begin{cases} a_j & (\text{if } p_{j,i} \neq 0) \\ 1 & (\text{if } p_{j,i} = 0) \end{cases} \quad (13)$$

where, a_j is the relative enzyme activity of the mutant to the wild-type for the j-th reaction. $p_{j,i}$ is the element for the j-th reaction and i-th EM in matrix **P**. The flux distribution of the mutant could be predicted as follows:

$$\mathbf{v}_{\text{prediction}}^{\text{mutant}} = \mathbf{P} \cdot \lambda^{\text{mutant}} \quad (14)$$

where, $\lambda^{\text{mutant}} = (\lambda_i^{\text{mutant}})$ and $\boldsymbol{v}_{\text{prediction}}^{\text{mutant}} = (v_{\text{prediction},i}^{\text{mutant}})$. The prediction errors of the estimated fluxes for mutants could be calculated by the following equation.

$$\text{Prediction error (ECF)} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(v_{\text{prediction},i}^{\text{mutant}} - v_{\text{exp},i}^{\text{mutant}}\right)^2}, \quad (15)$$

where, $v_{\text{prediction,i}}^{\text{mutant}}$ and $v_{\text{exp},i}^{\text{mutant}}$ are the predicted and experimental fluxes of mutants for the i-th reaction, respectively.

The ECF algorithm was applied to estimate the EMCs from the fluxes determined for wild-type *E. coli* (10, 23, 24) and *B. subtilis* (22), where the enzyme activity profiles were measured. Use of ECF predicted the flux distributions of the *pyk*, *ppc*, *fnr*, and *cra* mutants of *E. coli*, as well as those for the *als* over-expressing and *pta* knockout mutants of *B. subtilis*. Details of the network data are shown in Supplementary data.

RESULTS

Feasibility of using the MEP to estimate flux distributions by elementary mode analysis External fluxes were used to predict intracellular fluxes for three cellular metabolic network models (CHO, *E. coli*, *S. cerevisiae*, and *B. subtilis*) using the LP, QP, ECFLP, and MEP-based algorithms. Determining intracellular fluxes from observed external fluxes is difficult (25). Gas Chromatography–Mass Spectrometry (GC–MS) or Nuclear Magnetic Resonance (NMR) is commonly used to derive experimental data in the absence of a ^{13}C tracer (26). The prediction accuracy of the four algorithms used in our work is shown in Table 1. The prediction error of the MEP or ECFLP algorithm is lower than observed when using LP or QP. These data indicated that using the MEP was a feasible alternative.

A number of the measurable external fluxes are critical for the optimization of EMCs, but in most cases the measurable external fluxes are a limiting factor. We examined how the amount of available flux data affected the optimization of EMCs for *S. cerevisiae*. Using a specific growth rate (μ) of 0.30 h^{-1}, we compared the prediction capacity for four objective functions, LP, QP, ECFLP, and MEP. The uptake reaction is the reference value for the metabolic model so that it could be fixed for all of the calculations. For example, the uptake flux of glucose was fixed, as well as others that were randomly selected from the external fluxes.

If two experimental data values were examined for the optimizations, these were chosen from acetate, ethanol, and glycerol. There were three possible outcomes in this situation. As shown in Fig. 1, the prediction error decreased with an increase in the number of external fluxes used for the optimization. The predicted mean errors using ECFLP and MEP are much lower than those for LP and QP. Therefore, we determined that ECFLP and MEP are effective for EMC determination.

The optimized EMC profiles using LP, QP, ECFLP, and MEP for *S. cerevisiae* (μ=0.30 h^{-1}) are shown in Fig. 2. There were several dominant EMs whose EMCs were large in all four cases. The dominant

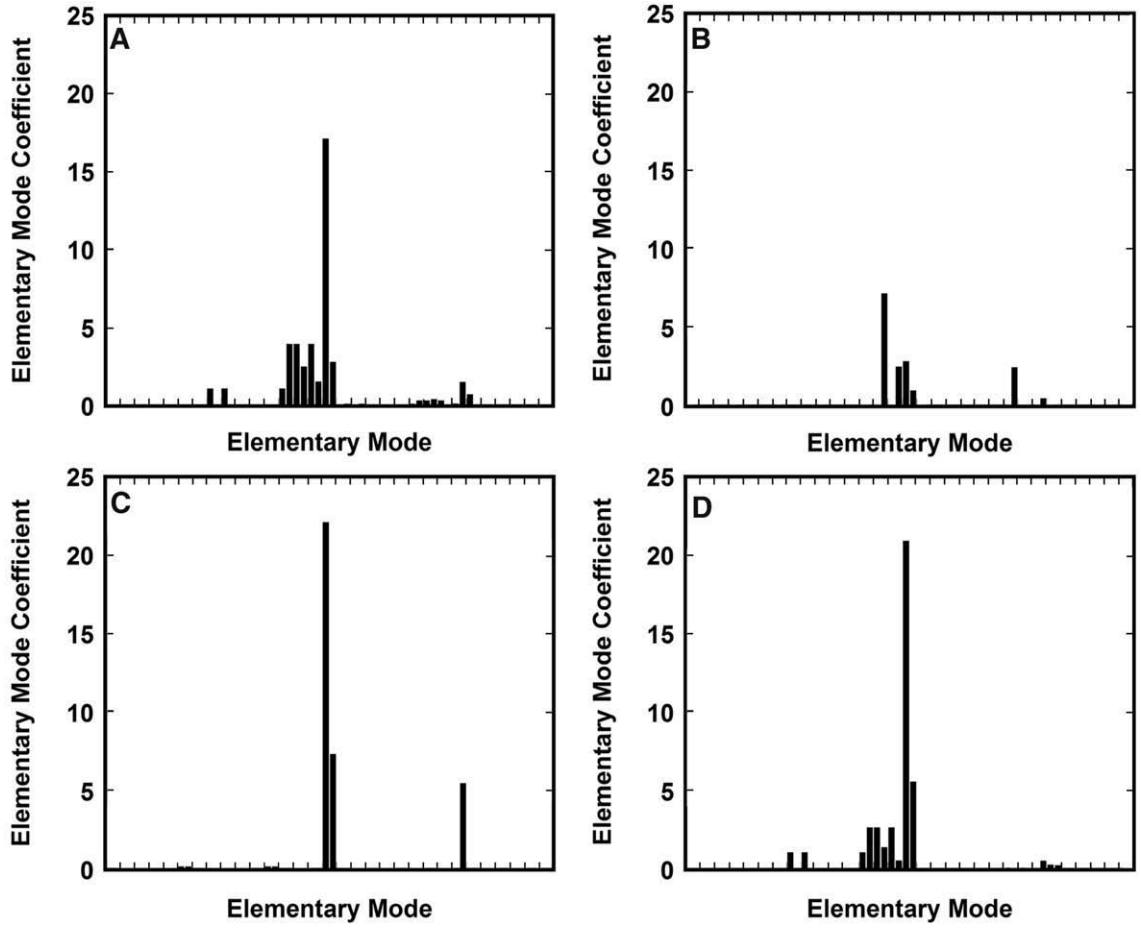

FIG. 2. The Elementary Mode Coefficients (EMCs) for *S. cerevisiae* (specific growth rate, μ=0.30 h^{-1}), as optimized by the Maximum Entropy Principle (MEP), Quadratic Programming (QP), Linear Programming (LP, where the objective function was for maximum biomass formation), and Linear Programming for the ECF (ECFLP). (A), MEP; (B), QP; (C), LP; (D), ECFLP.

EMCs were dependent on the objective function. The EMC profiles of ECFLP and MEP are similar, but different from those derived using LP and QP.

We compared the predicted with the experimental flux to further investigate the accuracy of each objective function. As shown in Fig. 3, ECFLP and MEP predicted each flux more accurately than did LP and QP. This supported our hypothesis that MEP and ECFLP are useful optimization methods. The MEP and ECFLP estimated flux distributions and EMC profiles were similar. The largest prediction error was for CO_2 excretion by LP and QP. This error was suppressed in our MEP and ECFLP analysis. CO_2 is mainly produced by the tricarboxylic acid cycle in our cellular models.

Application of the MEP to the ECF Provided that there is sufficient enzyme activity in both wild-type and mutant cells, the ECF can be used to estimate the flux distribution in the mutants. We used the ECF to predict the flux distribution of four *E. coli* mutants and one mutant of *B. subtilis* using different objective functions as shown in Figs. 4 and 5. The prediction error obtained by LP was the lowest for the *ppc* mutant (Fig. 5), suggesting that LP was most effective if the exact biological objective function was provided. For the other mutants, where the exact objective functions were not available, ECFLP and MEP predicted the flux distributions more accurately than both LP and QP. This finding supported the concept that MEP and ECFLP are effective optimization methods of the ECF where no biological objective function is available. These data also suggest that the experimental flux distribution was similar to the flux distribution estimated by MEP.

DISCUSSION

Generally, an additional hypothesis or an objective function is necessary to estimate the EMC as shown in equations (4, 7, 8, 9). With LP, an objective function is derived based on the concept that an organism can reach a predicted state, such as maximum biomass or ATP production (8). But it is difficult to deduce the rules for these set situations because objective functions can vary between different organisms and physiological conditions. In addition, mutants often reach a suboptimal state after gene deletion, thus maximum biomass formation, is not always a good choice for metabolic flux analysis (27). QP, in contrast, employs the minimum normal solution for EMCs, is suitable for systems whose objective function cannot be defined in biological terms. However, it should be noted that a sufficient number of measurable fluxes are required to fit the actual flux distribution of the system. Significantly, while Wlaschin's algorithm could not predict all of the intracellular fluxes (25), a linear relation was observed between EM combustion entropies and the EM family weight factors as estimated by the combustion thermodynamics of biochemical molecules.

To overcome these challenges, we proposed to demonstrate that the MEP, which maximizes Shannon's entropy, could effectively be applied to elementary mode and ECF analyses. Shannon's entropy is called the information entropy and is a measure of system complexity. The most possible distribution of random events is obtained by maximizing it. In biological sciences, this function has been widely used for analyses involving genes, RNA, and protein in living systems

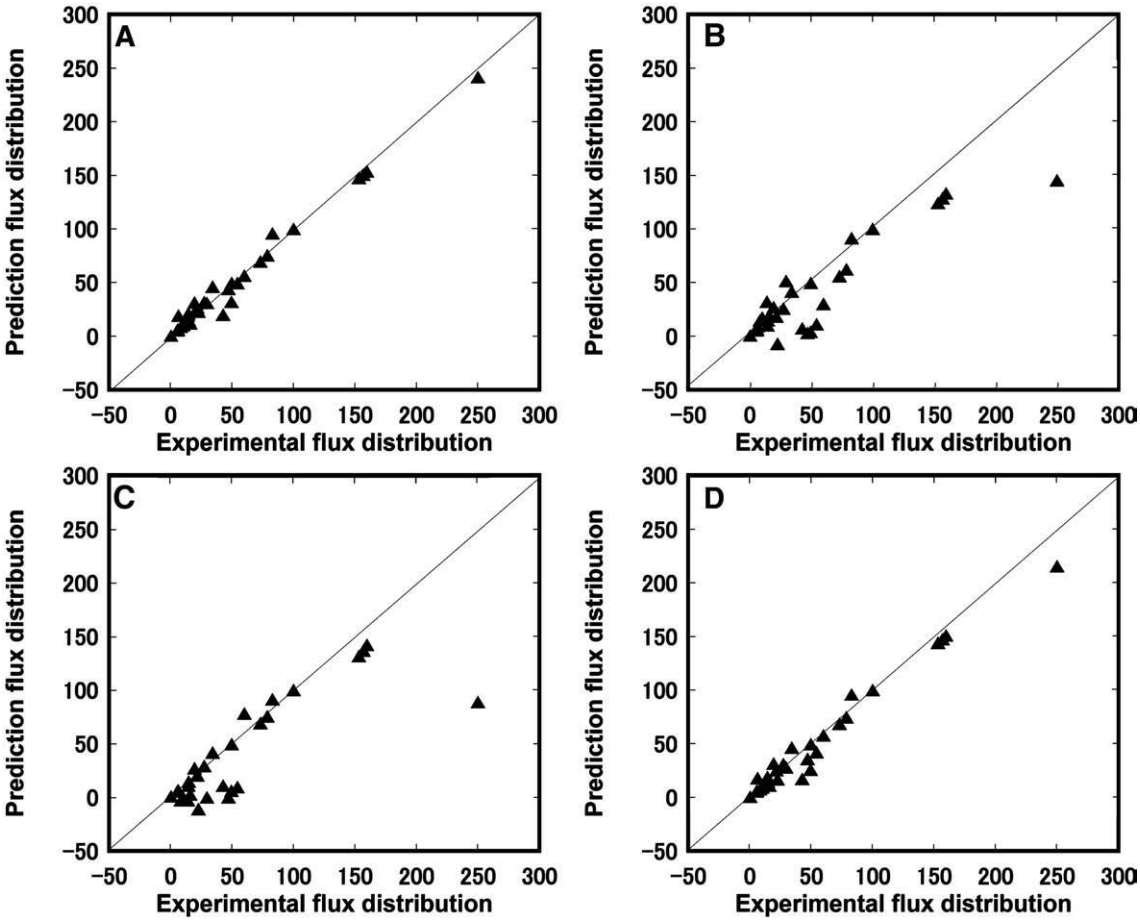

FIG. 3. The predicted fluxes versus experimental fluxes for *S. cerevisiae* (specific growth rate, μ=0.30 h^{-1}). The fluxes were optimized using the Maximum Entropy Principle (MEP), Quadratic Programming (QP), Linear Programming (LP, where the objective function was for maximum biomass formation) and Linear Programming to obtain the ECF (ECFLP). (A), MEP; (B), QP; (C), LP; (D), ECFLP.

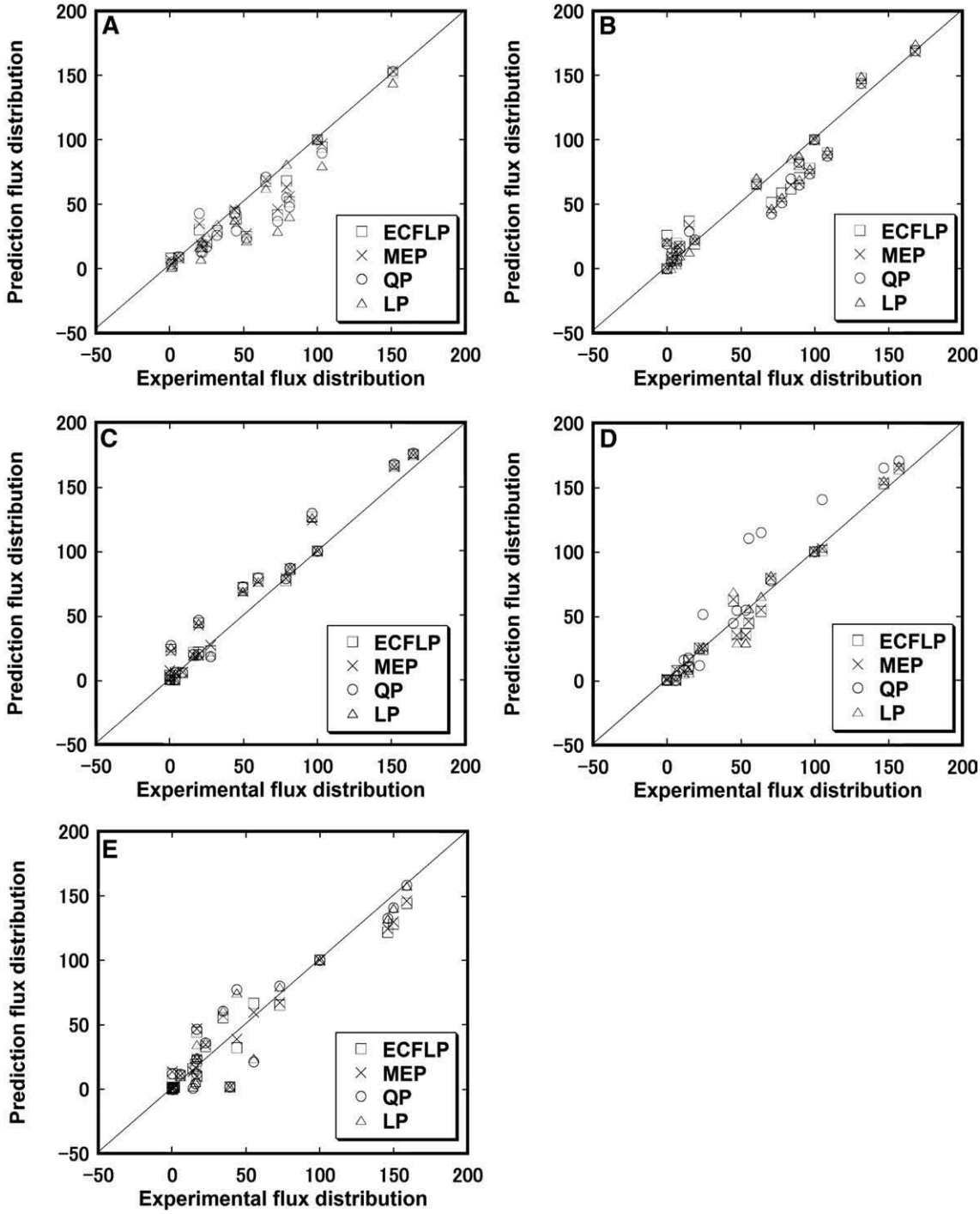

FIG. 4. Flux distributions for *pykF*(A), *ppc* (B), *fnr* (C), and *cra* (D) in *E. coli* mutants, as well as *als* over-expressing and *pta*-deletion mutants of *B. subtilis* (E) with ECF calculations using the objective functions for ECFLP, MEP, QP, and LP.

(28). Our data indicated that using the MEP accurately estimated the flux distribution when compared with other existing methods. This finding indicated that the MEP is a suitable objective function for estimating a flux distribution. This result may also suggest that the experimental flux distribution approximates the probability distribution provided by the MEP in a random system. When thinking of extremes, if the most likely probability distribution is conserved in mutants or any set of environmental conditions, the MEP might be a universal principle for metabolic flux analysis. Our future work will examine this intriguing concept.

We observed that EMCs and flux distributions optimized by MEP and ECFLP were very similar, but quite different from those predicted using LP and QP. We believe this similarity results from the fact that both the MEP and ECFLP are based purely on statistics and have no biological bias. The MEP derives the most probable distribution of EMCs based on Shannon's information theory. ECFLP, on the other hand, explores all possible EMC vectors, calculating their mean as a final solution for each EMC. However, note that ECFLP is not theoretical but empirical. Although we observed that the prediction error obtained using the MEP was almost the same as that for ECFLP,

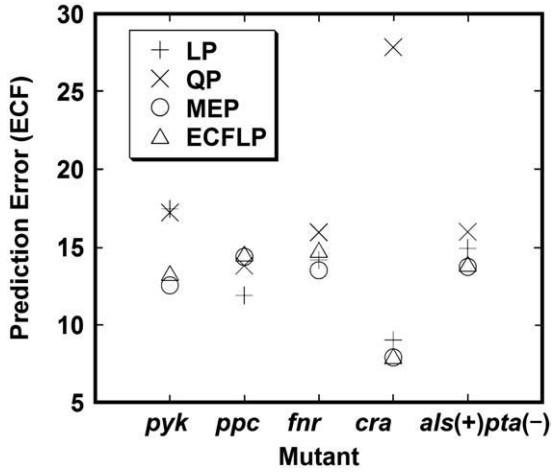

FIG. 5. ECF prediction errors using objective functions for ECFLP, MEP, QP, and LP to describe the *pyk*, *ppc*, *fnr*, and *cra* mutants of *E. coli*, as well as the *als* over-expression and *pta* deletion mutants of *B. subtilis*.

we consider the value obtained using the MEP to be superior because the MEP derivation is based on a straight forward formula and a theoretically sound background (Shannon's information theory). Because of its capacity for theoretical prediction, the MEP can predict the most likely probability distribution for metabolic fluxes but ECFLP cannot.

Shannon's information entropy is a measure of the average information content missing in a system. Therefore, compared to LP, the MEP needs neither additional experimental information nor a biological objective function. Generally it is difficult to predict a suitable biological objective function for different mutants and growth conditions. But the MEP is a physical rule predicting random events without any need of additional biological hypothesis or objective functions. Therefore, it is reasonable to choose the MEP to optimize EMCs when no biological objective function is used.

ACKNOWLEDGMENTS

This work was supported by grants from the Ministry of Education, Science, Sports and Culture, Grant-in-Aid for Scientific Research (B), 2006, 18300098.

APPENDIX A. SUPPLEMENTARY DATA

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.jbiosc.2008.09.011.

References

1. **Llaneras, F. and Picó, J.:** Stoichiometric modelling of cell metabolism, J. Biosci. Bioeng., **105**, 1–11 (2008).
2. **Shlomi, T., Eisenberg, Y., Sharan, R., and Ruppin, E.:** A genome-scale computational study of the interplay between transcriptional regulation and metabolism, Mol. Syst. Biol., **3**, 101 (2007).
3. **Herrgard, M. J., Lee, B. S., Portnoy, V., and Palsson, B. O.:** Integrated analysis of regulatory and metabolic networks reveals novel regulatory mechanisms in *Saccharomyces cerevisiae*, Genome Res., **16**, 627–635 (2006).
4. **Schuster, S., Dandekar, T., and Fell, D. A.:** Detection of elementary flux modes in biochemical networks: a promising tool for pathway analysis and metabolic engineering, Trends Biotechnol., **17**, 53–60 (1999).
5. **Ishii, N., Nakahigashi, K., Baba, T., Robert, M., Soga, T., Kanai, A., Hirasawa, T., Naba, M., Hirai, K., and Hoque, A., and other 19 authors:** Multiple high-throughput analyses monitor the response of *E. coli* to perturbations, Science, **316**, 593–597 (2007).
6. **Joyce, A. R. and Palsson, B. O.:** The model organism as a system: integrating 'omics' data sets, Nat. Rev., Mol. Cell Biol., **7**, 198–210 (2006).
7. **Cakir, T., Patil, K. R., Onsan, Z. I., Ulgen, K. O., Kirdar, B., and Nielsen, J.:** Integration of metabolome data with metabolic networks reveals reporter reactions, Mol. Syst. Biol., **2**, 50 (2006).
8. **Gayen, K. and Venkatesh, K. V.:** Analysis of optimal phenotypic space using elementary modes as applied to *Corynebacterium glutamicum*, BMC Bioinformatics, **7**, 445 (2006).
9. **Schwartz, J. M. and Kanehisa, M.:** A quadratic programming approach for decomposing steady-state metabolic flux distributions onto elementary modes, Bioinformatics, **21**, (Suppl. 2) ii 204–ii 205 (2005).
10. **Kurata, H., Zhao, Q. Y., Okuda, R., and Shimizu, K.:** Integration of enzyme activities into metabolic flux distributions by elementary mode analysis, BMC Syst. Biol., **1**, 31 (2007).
11. **Wiback, S. J., Mahadevan, R., and Palsson, B. O.:** Using metabolic flux data to further constrain the metabolic solution space and predict internal flux patterns: the *Escherichia coli* spectrum, Biotechnol. Bioeng., **86**, 317–331 (2004).
12. **Poolman, M. G., Venkatesh, K. V., Pidcock, M. K., and Fell, D. A.:** A method for the determination of flux in elementary modes, and its application to *Lactobacillus rhamnosus*, Biotechnol. Bioeng., **88**, 601–612 (2004).
13. **Schwartz, J. M. and Kanehisa, M.:** Quantitative elementary mode analysis ofmetabolic pathways: the example of yeast glycolysis, BMC Bioinformatics, **7**, 186 (2006).
14. **Lezon, T. R., Banavar, J. R., Cieplak, M., Maritan, A., and Fedoroff, N. V.:** Using the principle of entropy maximization to infer genetic interaction networks from gene expression patterns, Proc. Natl. Acad. Sci. U. S. A, **103**, 19033–19038 (2006).
15. **Capra, J. A. and Singh, M.:** Predicting functionally important residues from sequence conservation, Bioinformatics, **23**, 1875–1882 (2007).
16. **Martin, L. C., Gloor, G. B., Dunn, S. D., and Wahl, L. M.:** Using information theory to search for co-evolving residues in proteins, Bioinformatics, **21**, 4116–4124 (2005).
17. **Kurata, H., Matoba, N., and Shimizu, N.:** CADLIVE for constructing a large-scale biochemical network based on a simulation-directed notation and its application to yeast cell cycle, Nucleic Acids Res., **31**, 4071–4084 (2003).
18. **Klamt, S., Saez-Rodriguez, J., and Gilles, E. D.:** Structural and functional analysis of cellular networks with CellNetAnalyzer, BMC Syst. Biol., **1**, 2 (2007).
19. **Llaneras, F. and Pico, J.:** An interval approach for dealing with flux distributions and elementary modes activity patterns, J. Theor. Biol., **246**, 290–308 (2007).
20. **Burgard, A. P. and Maranas, C. D.:** Optimization-based framework for inferring and testing hypothesized metabolic objective functions, Biotechnol. Bioeng., **82**, 670–677 (2003).
21. **Frick, O. and Wittmann, C.:** Characterization of the metabolic shift between oxidative and fermentative growth in *Saccharomyces cerevisiae* by comparative 13C flux analysis, Microb. Cell Fact., **4**, 30 (2005).
22. **Zhu, Y. B., Chen, X., Chen, T., and Zhao, X. M.:** Enhancement of riboflavin production by overexpression of acetolactate synthase in a *pta* mutant of *Bacillus subtilis*, FEMS Microbiol. Lett., **266**, 224–230 (2007).
23. **Peng, L. F., Arauzo-Bravo, M. J., and Shimizu, K.:** Metabolic flux analysis for a *ppc* mutant *Escherichia coli* based on C-13-labelling experiments together with enzyme activity assays and intracellular metabolite measurements, FEMS Microbiol. Lett., **235**, 17–23 (2004).
24. **Siddiquee, K. A., Arauzo-Bravo, M. J., and Shimizu, K.:** Effect of a pyruvate kinase (*pykF*-gene) knockout mutation on the control of gene expression and metabolic fluxes in *Escherichia coli*, FEMS Microbiol. Lett., **235**, 25–33 (2004).
25. **Wlaschin, A. P., Trinh, C. T., Carlson, R., and Srienc, F.:** The fractional contributions of elementary modes to the metabolism of *Escherichia coli* and their estimation from reaction entropies, Metab. Eng., **8**, 338–352 (2006).
26. **Shirai, T., Matsuzaki, K., Kuzumoto, M., Nagahisa, K., Furusawa, C., Shioya, S., and Shimizu, H.:** Precise metabolic flux analysis of coryneform bacteria by gas chromatography-mass spectrometry and verification by nuclear magnetic resonance, J. Biosci. Bioeng., **102**, 413–424 (2006).
27. **Segre, D., Vitkup, D., and Church, G. M.:** Analysis of optimality in natural and perturbed metabolic networks, Proc. Natl. Acad. Sci. U. S. A, **99**, 15112–15117 (2002).
28. **Adami, C., Ofria, C., and Collier, T. C.:** Evolution of biological complexity, Proc. Natl. Acad. Sci. U. S. A, **97**, 4463–4468 (2000).