Continuous-Time Markov Chain–Based Flux Analysis in Metabolism

YUNZHANG HUO and PING JI

ABSTRACT

Metabolic flux analysis (MFA), a key technology in bioinformatics, is an effective way of analyzing the entire metabolic system by measuring fluxes. Many existing MFA approaches are based on differential equations, which are complicated to be solved mathematically. So MFA requires some simple approaches to investigate metabolism further. In this article, we applied continuous-time Markov chain to MFA, called MMFA approach, and transformed the MFA problem into a set of quadratic equations by analyzing the transition probability of each carbon atom in the entire metabolic system. Unlike the other methods, MMFA analyzes the metabolic model only through the transition probability. This approach is very generic and it could be applied to any metabolic system if all the reaction mechanisms in the system are known. The results of the MMFA approach were compared with several chemical reaction equilibrium constants from early experiments by taking pentose phosphate pathway as an example.

Key words: continuous-time Markov chain, flux analysis, metabolism

1. INTRODUCTION

THE METABOLIC NETWORK CAN BE ANALYZED from various angles, and metabolic flux analysis (MFA) is the most popular one (Bailey, 1991; Stephanopoulos et al., 1998). MFA is in fact a computational model of the intracellular metabolic network (Fischer et al., 2004). The idea of MFA is to build a model containing flows and metabolites. Considering that the inflow is equal to the outflow when the metabolic process reaches its steady state, differential equations can be established to analyze the concentration of each metabolite at the steady state (Selivanov et al., 2004). In addition, the minimum error between analytical data obtained from the model and actual measurement data at the steady state can be considered as an indicator for assessment. MFA can be simulated by the Monte Carlo method (Kadirkamanathan et al., 2006; Schellenberger et al., 2012), Bayesian estimation simulation (Jayawardhana et al., 2008), and Markov process (Zhang, 2013), or a hybrid method, such as Markov chain Monte Carlo (Kadirkamanathan et al., 2006). Analytical data can be obtained through computer programming (Chen and Ji, 2011) and then analyzed using Lagrange multiplier methods (Wiechert et al., 1997; Xu et al., 2013) and finally processed with MFA.

In order to figure out the reaction mechanism in the metabolic process, generally, we use the method of ^{13}C isotope labeling. From the perspective of experimental analysis, MFA is carried out according to the relative concentration of metabolites containing ^{13}C , which can be obtained through hydrogen-1 nuclear

magnetic resonance (1H NMR) spectroscopy experiment. Theoretically, a model can be built to simulate the metabolic pathway of metabolites containing ^{13}C (Benevolensky et al., 1994; Schellenberger et al., 2012). The ^{13}C marking method can be expressed by Monte Carlo random simulation (Schellenberger et al., 2012). A binary representation was used to express the reaction mechanism of metabolites containing ^{13}C by Selivanov et al. (2004). Then, experimental and theoretical analysis can be compared to calculate the relative error that serves as the assessment indicator of the model (Gombert et al., 2001).

In this article, chemical dynamics, bioinformatics, and mathematical statistics are combined to carry out MFA. On the basis of chemical dynamics, we use the simple mathematical relations between the instantaneous rate (ν) of the reaction and the concentration of the reaction products (A and B) as follows:

$$\nu = k[A]^\alpha[B]^\beta \quad (1)$$

where k is a constant indicating the chemical reaction rate that can be measured by NMR experiments (Milstein and Bremermann, 1979; Greenfield et al., 1988) or calculated by establishing correspondent differential equations. Furthermore, the metabolic process can be seen as the transition process of carbon atoms that conforms to the Markov process, that is, the no memory property (Maruyama, 1955). Therefore, we can establish a Markov model to carry out MFA (MMFA), and thus obtain the relative concentrations of various metabolites when the metabolic process reaches the steady state.

For the purpose of verification, we select the pentose phosphate pathway to calculate the relative concentrations of individual metabolites at the steady state. From the result, we calculate the equilibrium constants of some reversible reaction in the metabolic process, and compare them with the measured data.

2. METHODS

Consider a simple metabolic system that contains m metabolites denoted as A_i ($i=1, 2, \dots, m$), and N_i represents the number of carbon atoms in the carbon backbone of A_i . First, suppose that an initial metabolite A_1 is selected. In fact, we can select any metabolite A_i ($i=1, 2, \dots, m$). Besides, suppose that A_1 has n carbon atoms, that is, $N_1=n$ ($n \geq 1$). Our problem is to derive the relative concentration for each metabolite A_i ($i=1, 2, \dots, m$) compared with the initial concentration of A_1 when the entire metabolic system reaches its steady state. We assume that the initial concentration of A_1 is defined as 1, while the initial concentration of A_i ($i=2, \dots, m$) is defined as 0. The metabolic process can be regarded as a transfer process for carbon atoms existed in A_1 , so our problem can be transformed to study the distribution of each metabolite when every carbon atom in A_1 reaches the steady state.

In order to solve the above problem, we first assign a serial number for each of the carbon atoms in A_1 referred to as $1, 2, \dots, n$. $X^h(t)$ represents the state of the h th carbon atom at time t , where the state space is $M=\{1, \dots, m\}$. For example, $X^h(t)=a$ ($a \in M$) indicates that the h th carbon atom of A_1 is in A_a at time t . $P_{ab}^h(t, t')$ stands for the transition probability for the h th carbon atom from A_a to A_b after a time interval t' , that is,

$$P_{ab}^h(t, t') = Pr(X^h(t+t')=b|X^h(t)=a) \quad (2)$$

The transition matrix for the h th carbon atom of A_i is expressed as follows:

$$P^h(t, t') = (P_{ab}^h(t, t'))_{m \times m}$$

Thus, we can obtain the density matrix Q by

$$Q^h(t) = \frac{dP^h(t, t')}{dt'} = \left(\frac{dP_{ab}^h(t, t')}{dt'} \right)_{m \times m} := (q_{ab}^h(t))_{m \times m} \quad (3)$$

If $a \neq b$, for the continuous-time Markov chain, we have

$$\begin{aligned} q_{ab}^h(t) &= \frac{dP_{ab}^h(t, t')}{dt'} \\ &= \lim_{t' \rightarrow 0} \frac{P_{ab}^h(t, t')}{t'} \\ &= \frac{\nu_{a \rightarrow b}^h(t)}{[A_a(t)]} \end{aligned} \quad (4)$$

where $[A_a(t)]$ represents the relative concentration of A_a at time t .

If $a=b$, then

$$q_{ab}^h(t) = - \sum_{b=1, b \neq a}^m q_{ab}^h(t) \quad (5)$$

We assume that there are g_h reactions in the transfer process from A_a to A_b of the h th carbon atom. When combined with Equation 1,

$$\nu_{a \rightarrow b}^h(t) = \sum_{w=1}^{g_h} k_w [A]^{\alpha} [B]^{\beta} \quad (6)$$

where k_w represents the reaction rate constant of the w th ($w=1, 2, \dots, g_h$) reaction in Equation 1.

For other reactions without the transfer process of the h th carbon atom, then

$$\nu_{a \rightarrow b}^h(t) = 0 \quad (7)$$

All reactions are assumed to be noncomplex here.

Assuming that π_j^h is the probability that the h th carbon atom existed in A_i transfers to A_j when the system reaches the steady state. So the steady-state equation is

$$(\pi_1^h, \pi_2^h, \dots, \pi_m^h) Q^h(\infty) = 0, \text{ and } \sum_{j=1}^m \pi_j^h = 1 \quad h=1, 2, \dots, n \quad (8)$$

The entire metabolic process in fact is the transition of carbon atoms existed in A_1 , so that at the steady state, the total number of carbon atoms that have transferred to A_j is

$$[A_1(0)] V_0 (\pi_j^1 + \pi_j^2 + \dots + \pi_j^n) \quad j=1, 2, \dots, m \quad (9)$$

where V_0 stands for the entire volume (assumed to be unchanged during the metabolic process). Since each metabolite A_j contains N_j carbon atoms, the total number of metabolites at the steady state is

$$\frac{[A_1(0)] V_0 (\pi_j^1 + \pi_j^2 + \dots + \pi_j^n)}{N_j} \quad (10)$$

Therefore, the relative concentration of A_j at the steady state can be derived as

$$[A_j(\infty)] = \frac{[A_1(0)] V_0 (\pi_j^1 + \pi_j^2 + \dots + \pi_j^n)}{V_0 N_j [A_1(0)]} = \frac{(\pi_j^1 + \pi_j^2 + \dots + \pi_j^n)}{N_j} \quad (11)$$

By combining Equations 4–8 and 11, a set of quadratic equations can be obtained. Then we can calculate π_j^h , which is again plugged into Equation 11 so that the relative concentration for each metabolite compared with the initial concentration of A_i at the steady state can be obtained (see Supplementary Material at www.liebertonline.com).

FIG. 1. The process of the pentose phosphate pathway simulated in the model. Reaction 1 is Ru5P-R5P; reaction 2 is Ru5P-X5P; reaction 3 is R5P+X5P-S7P+GAP; reaction 4 is S7P-DHAP+E4P; reaction 5 is X5P+E4P-GAP+F6P; reaction 6 is F6P-DHAP+GAP; reaction 7 is DHAP-GAP. Details of the reaction are shown in reaction mechanism. DHAP, dihydroacetone phosphate; E4P, erythrose 4-phosphate; F6P, fructose 6-phosphate; GAP, glyceraldehyde 3-phosphate; R5P, ribose 5-phosphate; Ru5P, ribulose 5-phosphate; S7P, sedoheptulose 7-phosphate; X5P, xylulose 5-phosphate.

TABLE 1. THE MECHANISMS OF THE SEVEN REACTIONS

1.			abcde Ru5P	$\xrightleftharpoons[k_2]{k_1}$	abcde R5P		
2.			abcde Ru5P	$\xrightleftharpoons[k_4]{k_3}$	abcde X5P		
3.	abcde R5P	+	ABCDE X5P	$\xrightleftharpoons[k_6]{k_5}$	ABabcde S7P	+	CDE GAP
4.			abcdefg S7P	$\xrightleftharpoons[k_8]{k_7}$	abc DHAP	+	defg E4P
5.	abcde X5P	+	ABCD E4P	$\xrightleftharpoons[k_{10}]{k_9}$	cde GAP	+	abABCD F6P
6.			abcdef F6P	$\xrightleftharpoons[k_{12}]{k_{11}}$	abc DHAP	+	def GAP
7.			abc DHAP	$\xrightleftharpoons[k_{14}]{k_{13}}$	abc GAP		

3. APPLICATION

In this section, we apply the proposed MMFA approach to flux distribution analysis in the pentose phosphate pathway.

First, we can draw a schematic diagram of the pentose phosphate pathway (Fig. 1). From Figure 1, we can easily discover that the entire system contains eight metabolites and seven reactions. The eight metabolites are represented as A_1 (Ru5P), A_2 (R5P), A_3 (X5P), A_4 (S7P), A_5 (GAP), A_6 (F6p), A_7 (E4P), and A_8 (DHAP). The mechanisms of these seven reactions are listed in Table 1.

In Table 1, k_w is a constant indicating the chemical reaction rate, shown in Table 2 (Milstein and Bremermann, 1979).

In this pathway, by the analysis of the carbon atom's reaction mechanism in the entire process, we find that the first carbon atom (i.e., "a" in Table 1) in A_1 can transfer back to any position (i.e., a–e) in A_1 . So the transfer process of each carbon atom is equivalent; this means $Q^1 = Q^2 = Q^3 = Q^4 = Q^5$. Then we can only calculate the value of Q^1 . After simulating the transfer process of the first carbon atom, we find that it is hard to obtain the accurate probability of which product contains the first carbon atom if there are two or more products in the reaction. In fact, in the entire pathway, the rate of each reaction is different. So in this article, we distribute the probability by the ratio of products' factors in the reaction, that is, 1:1 in Table 1 obviously.

TABLE 2. EXACT VALUE OF k_w

k_w	Exact value
k_1	11.52
k_2	26.72
k_3	9.67
k_4	6.91
k_5	5.26
k_6	4.46
k_7	0.0017
k_8	22.85
k_9	267.78
k_{10}	21.27
k_{11}	0.000280
k_{12}	1.83
k_{13}	0.066
k_{14}	1.38

TABLE 3. EACH CHEMICAL REACTION EQUILIBRIUM CONSTANT K OF THE SECOND, SIXTH, AND SEVENTH REACTIONS

	<i>Concentration ratio</i>	<i>Exact value</i>	<i>References</i>
K_1	[X5P]/[Ru5P]	1.4	Casazza and Veech (1986)
K_2	[GAP]/[DHAP]	0.047	Connett et al. (1985)
K_3	[GAP]/[DHAP]/[F6P]	0.086	Veech et al. (1969)

TABLE 4. THE ACTUAL ERROR AND RELATIVE ERROR BETWEEN THE K_1 'S ESTIMATION VALUE AND EXACT VALUE WHEN U CHANGES FROM 10 TO 50 MMOL/L WITH THE DISTANCE VALUE OF 5 MMOL/L

<i>Value of u (mol/L)</i>	<i>Estimation value</i>	<i>Exact value</i>	<i>Actual error</i>	<i>Relative error (%)</i>
0.01	1.399351	1.4	0.000649	0.0464
0.015	1.399334	1.4	0.000666	0.0476
0.02	1.39932	1.4	0.00068	0.0486
0.025	1.399309	1.4	0.000691	0.0494
0.03	1.3993	1.4	0.0007	0.0500
0.035	1.399309	1.4	0.000691	0.0494
0.04	1.399284	1.4	0.000716	0.0511
0.045	1.399277	1.4	0.000723	0.0516
0.05	1.399271	1.4	0.000729	0.0521
Average	1.3993061	1.4	0.0006939	0.0496

To test the effect of the MMFA approach, we evaluate the relative concentration through the reaction equilibrium constant (K) in Section 4.

4. RESULTS

Using MATLAB, we input an initial concentration of A_1 and we got a group of probabilities when the system reached the steady state. Also we obtained the related concentration of each metabolism at the steady state. To test the approach, we selected the reaction equilibrium constant K to be the assessment indicator because each reaction in the pentose phosphate pathway is a reversible one.

TABLE 5. THE ACTUAL ERROR AND RELATIVE ERROR BETWEEN THE K_2 'S ESTIMATION VALUE AND EXACT VALUE WHEN U CHANGES FROM 10 TO 50 MMOL/L WITH THE DISTANCE VALUE OF 5 MMOL/L

<i>Value of u (mol/L)</i>	<i>Estimation value</i>	<i>Exact value</i>	<i>Actual error</i>	<i>Relative error (%)</i>
0.01	0.048174	0.047	0.001174	2.50
0.015	0.048364	0.047	0.001364	2.90
0.02	0.048555	0.047	0.001555	3.31
0.025	0.048746	0.047	0.001746	3.71
0.03	0.048937	0.047	0.001937	4.12
0.035	0.048746	0.047	0.001746	3.71
0.04	0.049321	0.047	0.002321	4.94
0.045	0.049515	0.047	0.002515	5.35
0.05	0.049709	0.047	0.002709	5.76
Average	0.048896	0.047	0.001896	4.03

TABLE 6. THE ACTUAL ERROR AND RELATIVE ERROR BETWEEN THE K_3 'S ESTIMATION VALUE AND EXACT VALUE WHEN u CHANGES FROM 10 TO 50 MMOL/L WITH THE DISTANCE VALUE OF 5 MMOL/L

Value of u (mol/L)	Estimation value	Exact value	Actual error	Relative error (%)
0.01	0.097444	0.086	0.011444	13.31
0.015	0.086965	0.086	0.000965	1.12
0.02	0.081333	0.086	0.004667	5.43
0.025	0.077802	0.086	0.008198	9.53
0.03	0.075388	0.086	0.010612	12.34
0.035	0.077802	0.086	0.008198	9.53
0.04	0.072343	0.086	0.013657	15.88
0.045	0.071343	0.086	0.014657	17.04
0.05	0.070562	0.086	0.015438	17.95
Average	0.078998	0.086	0.009761	11.35

The pentose phosphate pathway has been studied by many researchers (Kleijn et al., 2006; Luo et al., 2007; Zhao et al., 2008). Although it is simple and small, the most reactants and products in the pentose phosphate pathway are related with other metabolic systems. In this article, we selected the second, sixth, and seventh reactions to test the approach. Each K of these three reactions is listed in Table 3.

In MATLAB, we verified the initial concentration (defined as u) from 10 to 50 mmol/L and the distance value is 5 mmol/L, and the results are shown in Tables 4–6.

The results are also shown in Figure 2, where solid lines represent the estimation values of K_1 , K_2 , and K_3 , while dotted lines are the exact values. From Figure 2a, the relative error of K_1 between the estimation value and the exact value is extremely small, so the estimation value is quite accurate. Also, the estimation value of K_2 is very close to its exact value, and its relative error is less than 6%, as seen in Table 5 and Figure 2b. The relative error of K_3 is a little more than K_1 and K_2 ; however, the estimation value is very close to its exact value. The relative error of K_3 is a little bit large because the corresponding metabolite F6P can also be produced by G6P in the entire metabolic system, but G6P is not included in the pentose phosphate pathway. In all, we can conclude that the results of this approach are quite reliable and accurate.

FIG. 2. (a) The estimation value of K_1 when w changes from 10 to 50 mmol/L with the distance value of 5 mmol/L. (b) The estimation value of K_2 when u changes from 10 to 50 mmol/L with the distance value of 5 mmol/L. (c) The estimation value of K_3 when u changes from 10 to 50 mmol/L with the distance value of 5 mmol/L.

5. CONCLUSION

MFA is a powerful technique for determining intracellular metabolic fluxes in living cells. MFA has been found its widespread use in metabolic engineering, systems biology, biomedical research, and biotechnology.

In this article, a brand new approach, which combines MFA and continuous-time Markov chain, has been put forward to analyze metabolic flux in the metabolic system. On the basis of the study of the pentose phosphate pathway discussed in the Application section, this approach calculated the steady-state concentration by the distribution of each carbon atom. We obtained the reaction equilibrium constant for analysis and comparison, and found that this approach is feasible and the results are accurate. The approach proposed in this article is generic, which can be applied to any metabolic systems.

ACKNOWLEDGMENTS

The work described in this article was supported by a grant from The Hong Kong Polytechnic University (Project G-UA71).

AUTHOR DISCLOSURE STATEMENT

The authors declare that no competing financial interests exist.

REFERENCES

- Bailey, J.E. 1991. Toward a science of metabolic engineering. *Science* 252, 1668–1675.
- Benevolensky, S.S.V., Clifton, D., and Fraenkel, O.D.G. 1994. The effect of increased phosphoglucose isomerase on glucos metabolism in *Saccharomyces cerevisiae*. *J. Biol. Chem.* 269, 4878–4882.
- Casazza, J.P., and Veech, R.L. 1986. The interdependence of glycolytic and pentose cycle intermediates in ad libitum fed rats. *J. Biol. Chem.* 261, 690–698.
- Chen, C.S.H., and Ji, P. 2011. Decomposing flux distributions into elementary flux modes in genome-scale metabolic networks. *Bioinformatics* 27, 2256–2262.
- Connett, R.J. 1985. In vivo glycolytic equilibria in dog gracilis muscle. *J. Biol. Chem.* 260, 3314–3320.
- Fischer, E., Zamboni, N., and Sauer, U. 2004. High-throughput metabolic flux analysis based on gas chromatography-mass spectrometry derived ^{13}C constraints. *Anal. Biochem.* 325, 308–316.
- Gombert, A.K., dos Santos, M.M., Christensen, B., and Nielsen, J. 2001. Network identification and flux quantification in the central metabolism of *Saccharomyces cerevisiae* under different conditions of glucose repression. *J. Bacteriol.* 183, 1441–1451.
- Greenfield, N.J., McKenzie, M.A., Adebodun, F., et al. 1988. Metabolism of D-glucose in a wall-less mutant of *Neurospora crassa* examined by ^{13}C and ^{31}P nuclear magnetic resonances: effects of insulin. *Biochemistry* 27, 8526–8533.
- Jayawardhana, B., Kell, D.B., and Rattray, M. 2008. Bayesian inference of the sites of perturbations in metabolic pathways via Markov chain Monte Carlo. *Bioinformatics* 24, 1191–1197.
- Kadirkamanathan, V., Yang, J., Billings, S.A., and Wright, P.C. 2006. Markov Chain Monte Carlo Algorithm based metabolic flux distribution analysis on *Corynebacterium glutamicum*. *Bioinformatics* 22, 2681–2687.
- Kleijn, R.J., van Winden, W.A., Ras, C., et al. 2006. ^{13}C -labeled gluconate tracing as a direct and accurate method for determining the pentose phosphate pathway split ratio in *Penicillium chrysogenum*. *Appl. Environ. Microbiol.* 72, 4743–4754.
- Luo, B., Groenke, K., Takors, R., et al. 2007. Simultaneous determination of multiple intracellular metabolites in glycolysis, pentose phosphate pathway and tricarboxylic acid cycle by liquid chromatography–mass spectrometry. *J. Chromatogr. A* 1147, 153–164.
- Maruyama, G. 1955. Continuous Markov processes and stochastic equations. *Rend. Circ. Mat. Palermo* 4, 48–90.
- Milstein, J., and Bremermann, H.J. 1979. Parameter identification of the Calvin photosynthesis cycle. *J. Math. Biology* 7, 99–116.
- Schellenberger, J.D., Zielinski, C., Choi, W., et al. 2012. Predicting outcomes of steady-state ^{13}C isotope tracing experiments using Monte Carlo sampling. *BMC Syst. Biol.* 6, 9.

- Selivanov, V.A., Puigjaner, J., Sillero, A., et al. 2004. An optimized algorithm for flux estimation from isotopomer distribution in glucose metabolites. *Bioinformatics* 20, 3387–3397.
- Stephanopoulos, G.N., Aristidou, A.A., and Nielsen, J. 1998. *Metabolic Engineering*. Academic Press, San Diego, CA.
- Veech, R.L., Raijman, L., Dalziel, K., and Krebs, H.A. 1969. Disequilibrium in the triose phosphate isomerase system in rat liver. *Biochem. J.* 115, 837–842.
- Wiechert, W., Siefke, C., de Graaf, A.A., and Marx, A. 1997. Bidirectional reaction steps in metabolic networks: II. Flux estimation and statistical analysis. *Biotechnol. Bioeng.* 55, 118–135.
- Xu, Z.X., Sun, X., and Sun, J.B. 2013. Construction and analysis of the model of energy metabolism in *E. coli*. *PLoS ONE* 8, e55137.
- Zhang, Y. 2013. A dynamic Bayesian Markov model for phasing and characterizing haplotypes in next-generation sequencing. *Bioinformatics* 10, 1–8.
- Zhao, Z., Kuijvenhoven, K., Ras, C., et al. 2008. Isotopic non-stationary ^{13}C gluconate tracer method for accurate determination of the pentose phosphate pathway split-ratio in *Penicillium chrysogenum*. *Metab. Eng.* 10, 178–186.

Address correspondence to:

Mr. Yunzhang Huo
Department of Industrial and Systems Engineering
The Hong Kong Polytechnic University
Hung Hom
Hong Kong

E-mail: hyzhang1989@gmail.com