Incorporating Qualitative Knowledge in Enzyme Kinetic Models Using Fuzzy Logic

Bogju Lee,[1] John Yen,[1] Linyu Yang,[1] James C. Liao[2]

[1]Department of Computer Science, Texas A&M University, College Station, Texas
[2]Department of Chemical Engineering, 5531 Boelter Hall, University of California at Los Angeles, Los Angeles, CA 90095-1592; telephone: (310)-825-1656; fax: 310-206-1642; e-mail: liaoj@ucla.edu

Received 12 February 1998; accepted 18 September 1998

Abstract: Modeling of metabolic pathway dynamics requires detailed kinetic equations at the enzyme level. In particular, the kinetic equations must account for metabolite effectors that contribute significantly to the pathway regulation in vivo. Unfortunately, most kinetic rate laws available in the literature do not consider all the effectors simultaneously, and much kinetic information exists in a qualitative or semiquantitative form. In this article, we present a strategy to incorporate such information into the kinetic equation. This strategy uses fuzzy logic-based factors to modify algebraic rate laws that account for partial kinetic characteristics. The parameters introduced by the fuzzy factors are then optimized by use of a hybrid of simplex and genetic algorithms. The resulting model provides a flexible form that can simulate various kinetic behaviors. Such kinetic models are suitable for pathway modeling without complete enzyme mechanisms. Three enzymes in *Escherichia coli* central metabolism are used as examples: phosphoenolpyruvate carboxylase; phosphoenolpyruvate carboxykinase; and pyruvate kinase I. Results show that, with fuzzy logic-augmented models, the kinetic data can be much better described. In particular, complex behavior, such as allosteric inhibition, can be captured using fuzzy rules. The resulting models, even though they do not provide additional physical meaning in enzyme mechanisms, allow the model to incorporate semiquantitative information in metabolic pathway models. *Biotechnol Bioeng* **62**: 722–729, 1999.
Keywords: fuzzy logic; modeling; enzyme kinetics

INTRODUCTION

Because of the rapid progress in molecular biology, the underlying mechanisms of many biological processes have been elucidated at the molecular level. These molecular mechanisms can be combined to explain system behavior, often in an intuitive manner. However, intuitive reasoning becomes unsatisfactory as one demands more detailed explanation of system behavior. Therefore, mathematical modeling of biological systems is increasingly important for understanding complex behavior at the systems level. Numerous attempts have been reported to simulate or predict system behavior based on individual mechanisms (Achs and Garfinkel, 1977; Heinrich and Rapoport, 1974; Lee and Bailey, 1984; Liao et al. 1988; Shuler and Domach, 1983). In general, modeling of biochemical systems involves efforts at two levels: (1) a component level involving description of each molecular operation; and (2) a system level involving interactions among each component. For metabolic systems, the components are the enzymes, which interact with each other according to the stoichiometry and enzyme kinetics. Once the kinetic rate laws are known, they can be used in pathway models, which take the following form:

$$\frac{d\mathbf{X}}{dt} = \mathbf{SV} \tag{1}$$

$\mathbf{X}$ is the vector of metabolite concentrations, $\mathbf{S}$ is the stoichiometric matrix, and $\mathbf{V}$ is the vector of enzyme kinetic rate laws. In metabolic systems, the stoichiometry is generally well characterized. However, the enzyme kinetic rate laws are often incomplete, such that key characteristics (i.e., activation or inhibition) are missed, and thus the pathway model [Eq. (1)] becomes unrealistic even at the qualitative level. Therefore, it is important to develop enzyme kinetic models that capture at least the qualitative characteristics of metabolite effects on the enzymes.

Despite tremendous progress in understanding enzyme actions in the past few decades, most enzyme kinetic studies do not aim at developing kinetic rate expressions for the purpose of pathway modeling. Therefore, kinetic equations for enzymes are often incomplete, and most kinetic data are not used to develop quantitative rate expressions. Furthermore, qualitative or semiquantitative information is common. For example, the effect of ATP on *Escherichia coli* phosphoenolpyruvate carboxykinase (PCK) has been shown to be biphasic: it accelerates the reaction at low concentrations, whereas it inhibits at high concentrations (Wright and Sanwal, 1969). Although it is possible to develop mechanistic or empirical equations for describing the experimental data, there is no simple and systematic way to capture such

Correspondence to: J. Liao
Contract grant sponsor: National Science Foundation; contract grant number: BES-9511737

 CCC 0006-3592/99/060722-08

characteristics for the purpose of pathway modeling. This article addresses this problem by using fuzzy logic-augmented models and a hybrid of combinatorial and directional optimization algorithms.

Our goal is to provide a general approach to incorporate qualitative or semiquantitative information into enzyme kinetic rate laws. Three enzymes in *E. coli* central metabolism are used as examples: phosphoenolpyruvate carboxylase (PPC); PCK; and pyruvate kinase I (PYKI). The resulting models would capture the key characteristics of enzyme kinetics, and thus will be suitable for use in pathway models [Eq. (1)]. It should be emphasized that these rate laws are meant to be descriptive, and no mechanistic information will be gained by this approach. The fuzzy logic approach is certainly not the only methodology suitable for this purpose. It is chosen because of its conceptual simplicity and generality, as has been demonstrated repeatedly in multiple fields, including medicine (Sproule et al., 1997; Wirsam and Uthus, 1996), agriculture (Cassel-Gintz et al., 1997), and biotechnology (Kennedy and Spooner, 1996; Mukherjee et al., 1998;). Compared with other approximation techniques (e.g., piecewise linear approximation, splines, etc.), a fuzzy model offers two advantages. First, it is more flexible in providing a smooth approximation to a complex nonlinear relationship. Second, it explicitly describes qualitative knowledge. In this work, a full integration of fuzzy modeling and mathematical modeling is accomplished.

METHODS

Fuzzy Logic-Based Modeling

A fuzzy logic-based model uses a set of fuzzy if–then rules to capture the functional mapping relationship between a set of input variables and an output variable. The rule's ''if'' part (i.e., the antecedent) describes a fuzzy subregion and the rule's ''then'' part (i.e., the consequent) describes a local model for the region. There are two major types of fuzzy rule-based model for function approximation: the Mamdani model (Mamdani, 1974,1976), and the Takagi–Sugeno–Kang (TSK) model (Sugeno and Kang, 1988; Takagi and Sugeno, 1985). Interpolative reasoning is used to combine the output of multiple fuzzy rules whose antecedents partially overlap. This process, which has also been referred to as ''fuzzy inference'' or ''approximate reasoning'' in the literature, is analogous to linear interpolation. Details of this reasoning process can be found in the literature (Yen et al. 1995).

Parameter Estimation Using the Hybrid Simplex and Genetic Algorithm

Once the model is formulated, the parameters involved need to be determined based on experimental data. We used a hybrid of simplex and genetic algorithm (GA) by introducing the simplex method as an additional operator in the GA (Yen et al., 1998). GAS are global search and optimization techniques modeled from natural genetics, exploring search space by incorporating a set of candidate solutions in parallel (Holland, 1975). The main benefit of GAs is that they are less likely to be trapped in local optima, due to their parallelism and randomness. The main disadvantage of GAs are their high computational cost, because they typically converge slowly. The simplex method (Nelder and Mead, 1965; Spendley et al., 1962) is a local search technique that uses the current data set to determine the promising direction of search.

During the reproduction step of each iteration, the hybrid approach applies the simplex method to a top percentage of the population to produce new candidate solutions in the next generation. The rest of the new population are generated using the GA reproduction scheme (i.e., selection, crossover, and mutation). The hybrid method outperformed the GA in terms of the speed of convergence and the quality of solution (Yen et al., 1998).

RESULTS

Phosphoenolpyruvate Carboxylase

E. coli PPC catalyzes the carboxylation of phosphoenolpyruvate (PEP) to form oxaloacetate (OAA):

$$CO_2 + PEP \rightarrow OAA + P_i$$

The experimental data (Fig. 1) show the following characteristics (Izui et al. 1981): (1) the reaction rate exhibits a hyperbolic function of PEP concentration; (2) without any activator, the reaction proceeds at a very low rate; (3) acetyl-CoA (ACoA) is a very potent activator; and (4) fructose 1,6-diphosphate (FDP) exhibits no activation alone, but it produces a strong synergistic activation with ACoA. Despite such qualitative and quantitative information, no kinetic models were developed to capture these features.

Without considering the detailed molecular mechanism, we attempted to fit these data using an algebraic model based on the hyperbolic relationship between the reaction rate (V_{ppc}) and PEP:

$$V_{ppc} = V_m \frac{[PEP]}{K_m + [PEP]} \tag{2}$$

where [ ] indicates the concenatration. Two activators, ACoA and FDP, modulate V_m by the following equation:

$$V_m = \frac{K_1 + K_2[ACoA] + K_3[FDP] + K_4[ACoA][FDP]}{1 + K_5[ACoA] + K_6[FDP]} \tag{3}$$

In addition to K_m in Eq. (2), K_1 through K_6 are parameters to be estimated by data fitting. These equations were chosen because it exhibits the characteristics of the data, and were used to demonstrate an intuitive representation of the data. By using the hybrid GA simplex method, the parameters were optimized to take the following values: $K_m = 0.3231$

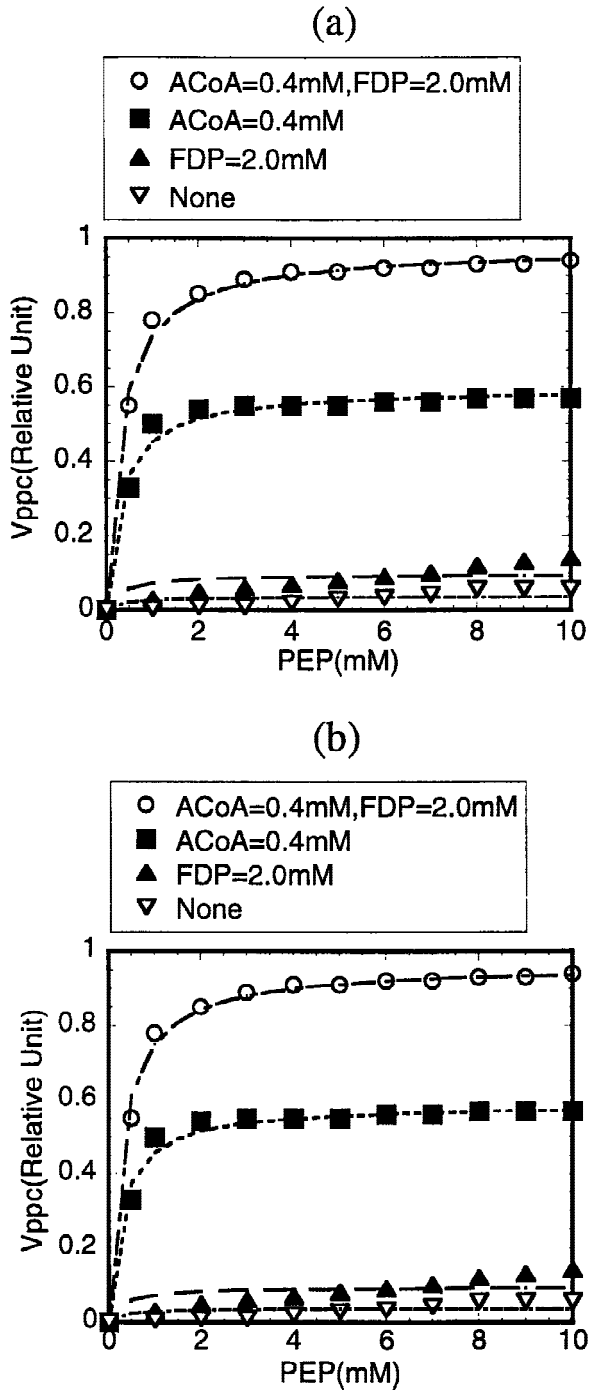

Figure 1. Kinetic data (symbols) (from Izui et al., 1981) and model fitting (lines) for PPC using: (a) the algebraic model [Eq. (2)] and (b) the fuzzy logic-augmented model [Eq. (4)].

mM, K_1 = 0.03176, K_2 = 1.2878 mM^{-1}, K_3 = 0.05425 mM^{-1}, K_4 = 0.8139 mM^{-1}, K_5 = 0.0939 mM^{-1}, and K_6 = 0.2693 mM^{-1}. Figure 1a shows that the model fits the data reasonably well. Despite the satisfactory fitting, the ad hoc approach is not generally applicable to other cases. In more complex data, the algebraic form of Eq. (3) may not be easily obtained.

To develop a general approach for data representation, we used the fuzzy logic approach and incorporated a situation-dependent scaling factor, α_{ppc}, to modify V_m:

$$V_{ppc} = \alpha_{ppc} V_m \frac{[PEP]}{K_m + [PEP]} \quad (4)$$

α_{ppc} captures the various activation effects of ACoA and FDP by the following fuzzy rules:

R_1^{ppc}: If [ACoA] is LOW and [FDP] is LOW, then $\alpha_{ppc} = c_1$

R_2^{ppc}: If [ACoA] is LOW and [FDP] is HIGH, then $\alpha_{ppc} = c_2$

R_1^{ppc}: If [ACoA] is HIGH and [FDP] is LOW, then $\alpha_{ppc} = c_3$

R_1^{ppc}: If [ACoA] is HIGH and [FDP] is LOW, then $\alpha_{ppc} = c_4$

where c_1, c_2, c_3, and c_4 are parameters to be optimized. The membership functions of the fuzzy sets (i.e., LOW and HIGH), shown in Figure 2, were chosen based on experimental data in Figure 1. Notice that the above rules (a special case of the TSK model) are conceptually clear and can be readily generalized. In this model, six parameters need to be optimized (as opposed to seven in the algebraic model). Again, simplex–GA was used to obtain the following optimized values: V_m = 0.987, K_m = 0.259 mM, c_1 = 0.037, c_2 = 0.095, c_3 = 0.594, and c_4 = 0.973. The performance of this model with the identified parameters (Fig. 1b) is as good as the algebraic model (Fig. 1a) for the

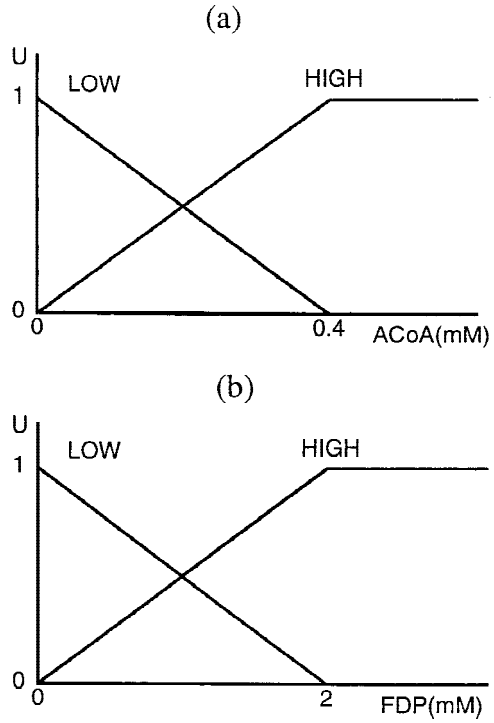

Figure 2. Membership functions of the fuzzy sets for PPC modeling.

purpose of pathway modeling. They all catch the essential characteristics of the data. In this case, the fuzzy logic-augmented model used fewer parameters than the algebraic model; furthermore, it is conceptually clear and potentially generalizable.

Phosphoenolpyruvate Carboxykinase

PCK is a gluoneogenic enzyme found in many organisms. The *E. coli* PCK catalyzes the following reaction:

$$\mathrm{OAA} + \mathrm{ATP} \rightarrow \mathrm{PEP} + \mathrm{ADP} + \mathrm{CO_2}$$

Experimental data (Wright and Sanwal, 1969) shown in Figure 3 show the following characteristics: (1) the rate of reaction is increased by ATP, but high ATP concentration inhibits the reaction; and (2) other metabolites (i.e., OAA, ADP, and PEP) increase the reaction rate in a hyperbolic manner.

We first use an algebraic equation to capture the hyperbolic behavior of the kinetics.

$$v_{\mathrm{pck}} = \frac{K_7[\mathrm{ATP}][\mathrm{OAA}] - K_8[\mathrm{ADP}][\mathrm{PEP}]}{1 + K_1[\mathrm{ATP}] + K_2[\mathrm{OAA}] + K_3[\mathrm{ATP}][\mathrm{OAA}] + K_4[\mathrm{PEP}] + K_5[\mathrm{ADP}] + K_6[\mathrm{PEP}][\mathrm{ADP}]} \tag{5}$$

This form does not describe the inhibition effect of ATP at the high concentration range. If all the data are used in parameter optimization, the fitting is poor (Fig. 3). However, if only the OAA, ADP, and PEP data are used for parameter estimation, then the above equation fit these data very well (K_1 = 1016.594 mM^{-1}, K_2 = 962.644 mM^{-1}, K_3 = 626.338 mM^{-2}, K_4 = 79.769 mM^{-1}, K_5 = 887.975 mM^{-1}, K_6 = 155.380 mM^{-2}, K_7 = 38.744 mM^{-2} , and K_8 = 335.25 mM^{-2}). Instead of modifying the algebraic equation to represent the ATP inhibition effect, we used a situation-dependent factor, α_{pck}, in the following equation:

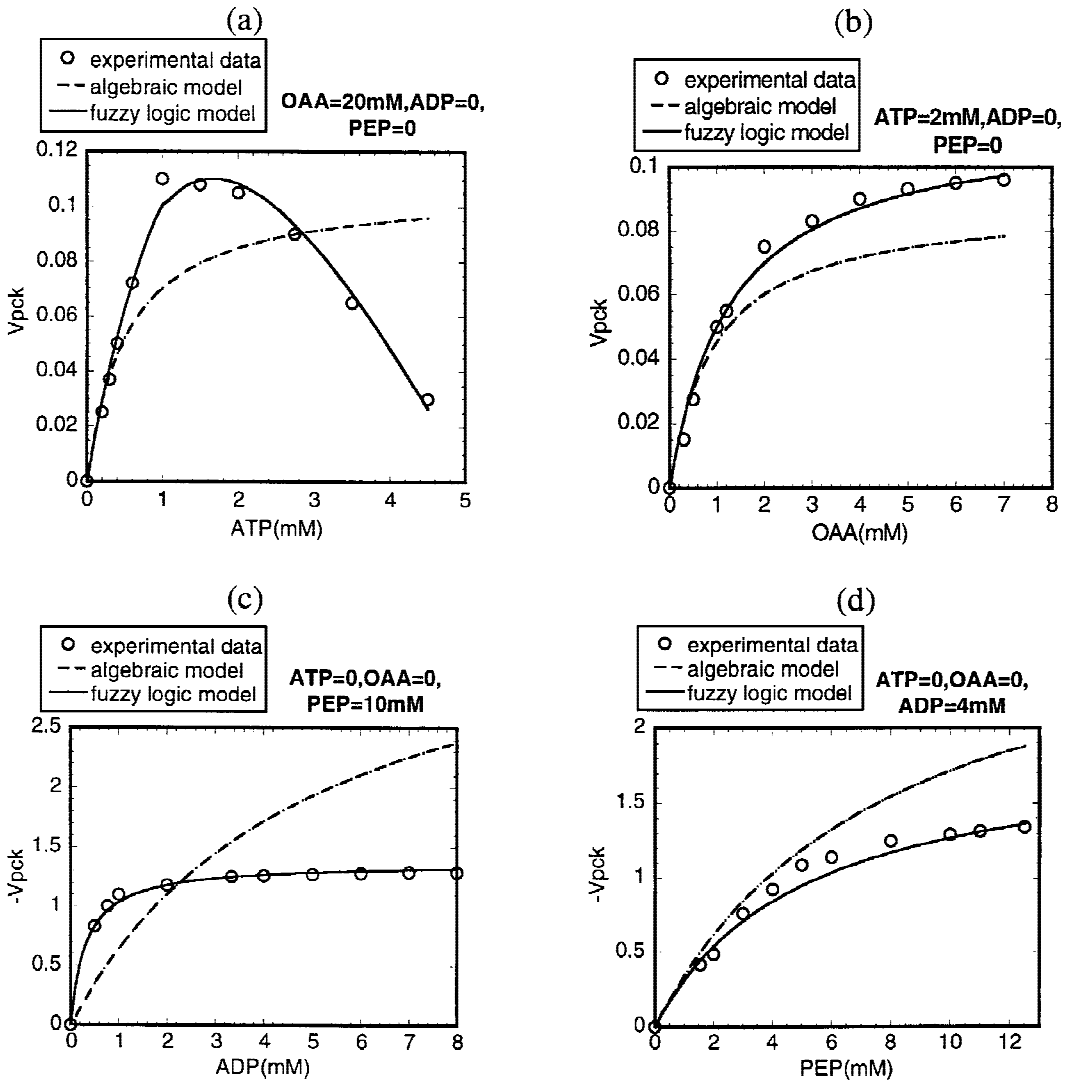

Figure 3. Experimental data (Wright and Sanwal, 1969) and model fitting for PCK kinetics using the algebraic [Eq. (5)] and fuzzy logic model [Eq. (6)].

$$v_{pck} = \alpha_{pck} \frac{K_7[ATP][OAA] - K_8[ADP][PEP]}{1 + K_1[ATP] + K_2[OAA] + K_3[ATP][OAA] + K_4[PEP] + K_5[ADP] + K_6[PEP][ADP]} \quad (6)$$

where α_{pck} is modeled according to the following rules:

R_1^{pck}: If [ATP] is LOW, then $\alpha_{pck} = 1.0$

R_2^{pck}: If [ATP] is HIGH, then $\alpha_{pck} = 1.0 - b[ATP]$

where b is a constant to be optimized. Because ATP concentration starts to inhibit the PCK reaction from around 1.0, as shown in the experimental data (Fig. 3) we divided the ATP input space into two partially overlapping subregions: ''LOW ATP'' and ''HIGH ATP.'' The membership functions for the ATP fuzzy sets in the fuzzy model are shown in Figure 4. The fuzzy model behaves as follows. The output of the fuzzy model (i.e., α_{pck}) is 1.0 when ATP concentration belongs completely to LOW (i.e., ATP $\leqslant$0.9). When ATP concentration belongs to HIGH (i.e., ATP $\geqslant$1.1), the output of the fuzzy model decreases as ATP concentration increases, which is represented by a linear equation (i.e., $\alpha_{pck} = 1.0 - b$ [ATP]). If the ATP concentration is in the middle area between completely LOW and completely HIGH (i.e., 0.9 < ATP < 1.1), the scaling factor is an interpolation of the two aforementioned effects:

$$\alpha_{pck} = 1 - \left(\frac{[ATP] - 0.9}{0.2}\right) b[ATP] \quad (7)$$

Although the fuzzy rule is linear, α_{pck} becomes a second-order polynomial equation of [ATP] in the transition region, which gives a smooth transition between the two regions. Note that nonlinear features of enzyme kinetics can be captured by properly choosing the linear fuzzy rules.

Because the values of K_4, K_5, K_6, and K_8 in the augmented model [Eq. (6)] are not affected by ATP concentrations, they can be obtained directly from the partially optimized values using Eq. (5), based on OAA, ADP, and PEP data. We then can use OAA and ATP data to optimize parameters related to ATP and OAA, providing the following results: $K_1 = 91145.54$ mM^{-1}, $K_2 = 66866.11$ mM^{-1}, $K_3 = 36753.48$ mM^{-2}, $K_7 = 11005.78$ mM^{-2}, $b = 0.24$ mM^{-1}.

The performance of this model with the identified parameters is shown in Figure 3. Table I compares the root-mean-square errors of modeling V_{pck} using the algebraic model and the fuzzy logic-augmented model. The fuzzy logic-augmented model fits the experimental data much better than the algebraic model. The former correctly represents the ATP inhibition effect, whereas the latter does not.

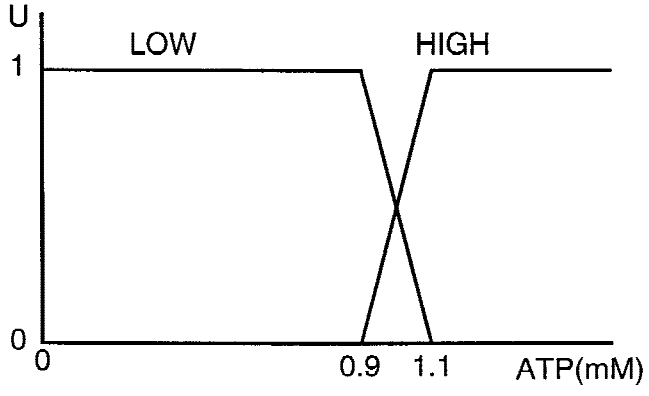

Figure 4. The membership function of fuzzy sets for PCK modeling.

Pyruvate Kinase I

PYKI in *E. coli* is an allosteric enzyme, which shows non-hyperbolic kinetic behavior under different concentrations of metabolites (Fig. 5) (Waygood and Sanwal, 1974). It catalyzes the following reaction:

$$PEP + ADP \rightarrow Pyruvate + ATP$$

The reaction is activated by FDP and PEP with a qualitative change from a sigmoidal kinetic to hyperbolic behavior (Fig. 5). The data also indicate an inhibitory effect by ACoA and ATP. However, the data are not sufficient to show the quantitative behavior of the activators and inhibitors. Only the trends of their effects are shown.

Because of the complex behavior, deriving a mechanistic model to account for all the effectors is difficult. The Monod–Wyman–Changeux (MWC) model (Monod et al., 1965) can only describe the basic behavior with only one varying metabolite. It was modified to account for some aspects of PEP and FDP effects:

$$V_{pykI} = \frac{[PEP](1 + [PEP])^3 + LC(1 + C[PEP])^3 + [FDP](1 + [FDP])^3 + L_f C_f (1 + C_f[FDP])^3}{(1 + [PEP]^4 + L(1 + C[PEP])^4 + (1 + [FDP])^4 + L_f(1 + C_f[FDP])^4} \quad (8)$$

where L, L_f, C, and C_f are to be determined. These parameters change the kinetic curve from sigmoidal to hyperbolic. They are influenced by activators and inhibitors and should be treated as functions of these metabolites. However, it is difficult to identify the appropriate structure of these functions. Moreover, the need to represent ACoA and ATP inhibition in addition to FDP and PEP activation is likely to lead to a structure that is too complex for purposes of pathway modeling.

Table I. Root-mean-square errors of modeling PCK.

Data set	Algebraic model [Eq (5)]	Fuzzy logic model [Eq. (6)]
Fig. 3a	0.02234	0.00419
Fig. 3b	0.01369	0.00315
Fig. 3c	0.65955	0.02690
Fig. 3d	0.29632	0.06835

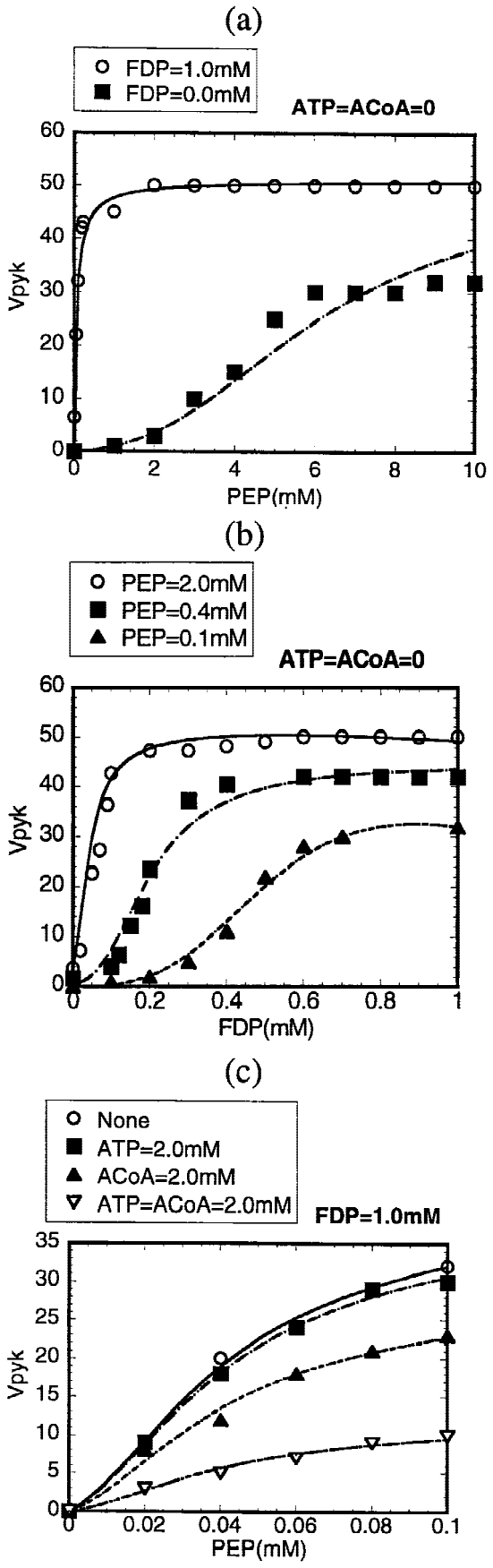

Figure 5. Experimental data (Waygood and Sanwal, 1974) and model fitting for PYKI.

We thus use fuzzy logic to represent these unknown functions, in which only the trends are revealed by the data. Eq. (8) is then modified to be:

$$V_{\text{pykI}} = \alpha_1(1-\alpha_2)\frac{[\text{PEP}](1+[\text{PEP}])^3 + LC(1+\text{C}[\text{PEP}])^3 + [\text{FDP}](1+[\text{FDP}])^3 + L_f C_f(1+C_f[\text{FDP}])^3}{(1+[\text{PEP}])^4 + L(1+C[\text{PEP}])^4 + (1+[\text{FDP}])^4 + L_f(1+C_f[\text{FDP}])^4} \quad (9)$$

where α_1 and α_2 represent the activation and inhibition effects, respectively, and they are determined by the following fuzzy rules:

$R_1^{\alpha 1}$: If [FDP] is LOW and [PEP] is LOW, then $\alpha_1 = a_1$

$R_2^{\alpha 1}$: If [FDP] is LOW and [PEP] is HIGH, then $\alpha_1 = a_2$

$R_3^{\alpha 1}$: If [FDP] is HIGH and [PEP] is LOW, then $\alpha_1 = a_3$

$R_4^{\alpha 1}$: If [FDP] is HIGH and [PEP] is HIGH, then $\alpha_1 = a_4$

Similarly, we model α_2 using fuzzy if–then rules with two inputs that determine the inhibition effects by ATP and ACoA:

$R_1^{\alpha 2}$: If [ATP] is LOW and [AcoA] is LOW, then $\alpha_2 = a_5$

$R_2^{\alpha 2}$: If [ATP] is LOW and [AcoA] is HIGH, then $\alpha_2 = a_6$

$R_3^{\alpha 2}$: If [ATP] is HIGH and [AcoA] is LOW, then $\alpha_2 = a_7$

$R_4^{\alpha 2}$: If [ATP] is HIGH and [AcoA] is HIGH, then $\alpha_2 = a_8$

However, changing only α_1 and α_2 will not be sufficient because they do not affect the qualitative behavior (i.e., the shape of the kinetic curve). The qualitative change in kinetic behavior by FDP and PEP activation is achieved by using fuzzy if–then rules for C, L, C_f, and L_f:

R_1: If [FDP] is LOW and [PEP] is LOW, then $C = C_1$, $L = L_1$, $C_f = C_{f1}$, $L_f = L_{f1}$

R_2: If [FDP] is LOW and [PEP] is HIGH, then $C = C_2$, $L = L_2$, $C_f = C_{f2}$, $L_f = L_{f2}$

R_3: If [FDP] is HIGH and [PEP] is LOW, then $C = C_3$, $L = L_3$, $C_f = C_{f3}$, $L_f = L_{f3}$

R_4: If [FDP] is HIGH and [PEP] is HIGH, then $C = C_4$, $L = L_4$, $C_f = C_{f4}$, $L_f = L_{f4}$

The membership functions for FDP, PEP, ATP, and ACoA

fuzzy sets in the fuzzy model are shown in Figure 6, which were constructed from the experimental data (Fig. 5).

The hybrid GA–simplex method was used to identify the parameter set in three steps. In each step, we used a subset of the experimental data to concentrate on optimizing parameters related to the specific effect demonstrated by the data. This strategy significantly reduces the number of parameters that need to be optimized in each step. First, we use data associated with the condition ATP = ACoA = 0 to optimize all the parameters except those related to ATP–ACoA inhibition. We then use the data in Figure 5c to optimize a_5 through a_8 while fixing other parameters. The last step is to fine-tune model parameters for repairing model deficiency resulting from an unbalanced distribution of training data. The data in Figure 5a and b are highly unbalanced in terms of PEP concentration. Most of them are for high PEP concentration (i.e., $\geqslant$0.4 m*M*); only a small portion of the data is for low PEP concentration. Consequently, the parameter optimization process is easily dominated by data at the high PEP concentration range, which causes the GA to converge to a local minimum that does not fit well at low PEP range (i.e., data for PEP = 0.1m*M*). To address this deficiency, the third step uses the data of low PEP concentration to optimize only parameters related to ''PEP is LOW'' (i.e., fixing all other parameters). The final parameter set identified is as follows (with units consistent with data in Fig. 5): $\alpha_1 = 95.59$, $\alpha_2 = 56.61$, $\alpha_3 = 45.78$, $\alpha_4 = 50.97$, $\alpha_5 = 0.0$, $\alpha_6 = 0.29$, $\alpha_7 = 0.042$, $\alpha_8 = 0.70$, $L_1 = 0.0$, $L_2 = 264618$, $L_3 = 5890509$, $L_4 = 2263077$, $c_1 = 0.0$, $c_2 = 0.31$, $c_3 = 24.96$, $c_4 = 14.61$, $L_{f1} = 88577785$, $L_{f2} = 8811417$, $L_{f3} = 4059350$, $L_{f4} = 3440$, $c_{f1} = 0.0$, $c_{f2} = 0.0001$, $c_{f3} = 0.49$, and $c_{f4} = 0.0001$.

The performance of the model with the identified parameters is shown in Figure 5, which shows that the fuzzy logic-augmented model fits all the experimental data very well.

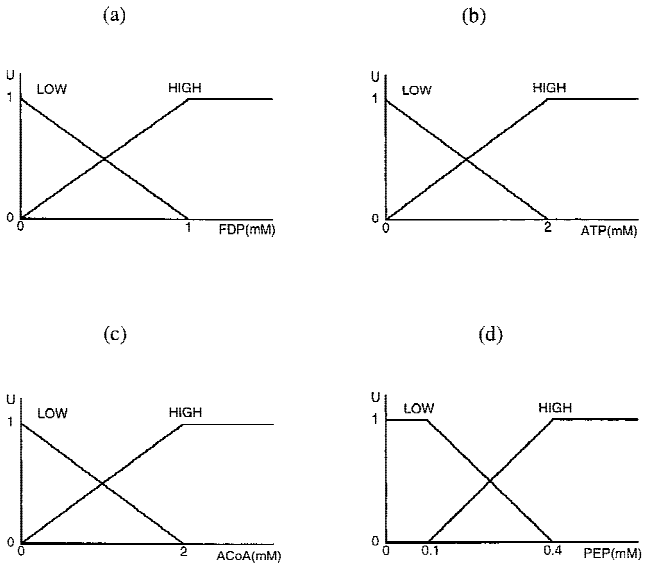

Figure 6. Membership functions of the fuzzy sets for PYKI modeling.

DISCUSSION

This article presents a fuzzy logic-based approach to modify algebraic kinetic equations to account for qualitative information beyond the scope of the existing model. The fuzzy parameter can be used to modify either V_m or any other parameters in a kinetic rate expression. In the cases of PPC and PCK, modification of V_m was sufficient to fit the data for various effectors. In the case of PYKI, however, one has to modify six parameters (α_1, α_2, c, L, c_f, and L_f) by fuzzy factors to account for the allosteric effects of FDP, ATP, and ACoA. Although the fuzzy logic-augmented model provides no additional mechanistic insight, it is instrumental in modeling pathway regulation, which is strongly dependent on the effect of various allosteric effectors. One of the major advantages of this approach is conceptually simple and can be generalized to most situations. Fuzzy logic is a tool for approximating a complex surface without a complicated mathematical form. Obviously, there are other tools that allow the approximation of a complex surface, such as nonparametric regression (Hardle, 1989), splines, and radial basis functions (Poggio and Girosi, 1990; Powell 1985). In fact, fuzzy logic modeling shares a common property with these techniques: they all approximate a complex surface by combining multiple local models (Yen et al., 1996). However, fuzzy logic is unique in that it uses linguistic terms to describe local models. This important feature establishes a bridge between qualitative knowledge and numerical models. Without such a bridge, qualitative information cannot be explicitly incorporated into an algebraic model of enzyme kinetics.

We chose to use the TSK fuzzy model instead of the Mamdani model for its simplicity and its compactness. The TSK model was developed to reduce the total number of rules required by the Mamdani model. Consequently, the number of rules in a TSK model is typically less than that in a Mamdani model, assuming that they approximate the same function to about the same accuracy. The simplicity and the compactness of the TSK model further simplifies the optimization problem that estimates the model parameters.

The fuzzy logic-based approach introduced a number of new parameters that have to be determined by data fitting. This task was made possible by using a hybrid of GA and simplex optimization. GA allows parallel search of a large parameter space without entrapment in a local minimum. The simplex method, on the other hand, greatly accelerates the convergence to a local minimum. By a proper mix of the two approaches, one can quickly converge into a local minimum and simultaneously search for other local valleys.

In summary, we have described an integration of four techniques for modeling metabolic pathways: fuzzy logic-based modeling; algebraic modeling of enzyme kinetics; use of the genetic algorithm; and use of the simplex method. The first two techniques were combined to achieve a more flexible model structure that can incorporate qualitative information explicitly into enzyme kinetics. The latter two

techniques were combined to obtain an efficient model parameter estimator that can avoid entrapment in local optima. Together, these four complementary techniques offer a promising approach for modeling enzyme kinetics when knowledge about their enzyme mechanisms is incomplete.

References

Achs MJ, Garfinkel, D. 1977. Computer simulation of rat heart metabolism after adding glucose to the perfusate. Am J Physiol 232:175–184.

Cassel-Gintz MA, Luedeke MKB, Petschel-Held G, Reusswig F, Ploechl M, Lammel G, Schellnuber HJ 1997. Fuzzy logic based global assessment of the marginality of agricultural land use. Climate Res 8: 135–150.

Hardle W. 1989. Applied nonparametric regression. Cambridge, UK: Cambridge University Press.

Heinrich R, Rapoport TA. 1974. A linear steady state treatment of enzymatic chains, general properties, control and effect strength. Eur J Biochem 42:89–95.

Holland JH. 1975. Adaptation in natural and artificial syatems. Ann Arbor, MI: University of Michigan Press.

Izui K, Taguchi M, Morikawa M, Katsuki H. 1981. Regulation of *Escherichia coli* phosphoenolpyruvate carboxylase by multiple effectors *in vivo.* ii. Kinetic studies with a reaction system containing physiological concentrations of ligands. J Biochem 90:1321–1331.

Kennedy MJ, Spooner NR. 1996. Using fuzzy logic to design fermentation media: A comparison to neural networks and factorial design. Biotechnol Techniq 10:47–52.

Lee SB, Bailey JE. 1984. A mathematical model for lambda dv plasmid replication: Analysis of copy number mutants. Plasmid 11:166–177.

Liao JC, Lightfoot EN, Jolly SO, Jacobson GK. 1988. Application of characteristic reaction paths: Rate-limiting capacity of phosphofructokinase in yeast fermentation. Biotechnol Bioeng 31:855–868.

Mamdani EH. 1974. Applications of fuzzy algorithms for simple dynamic plant. Proc IEEE 121:1585–1588.

Mamdani EH. 1976. Advances in linguistic synthesis of fuzzy controllers. Int J Man Machine Studies 8:669–678.

Monod J, Wyman J, Changeux JP. 1965. On the nature of allosteric transitions: a plausible model. J Mole Biol 12:88–118.

Nelder JA, Mead R. 1965. A simplex method for function minimization. Compu J 7:308–313.

Mukherjee S, Basu S, Chattopadhyay P. 1998. A fuzzy logic-based control system for compressed baker's yeast fermentation. J Food Sci Technol 35:164–168.

Poggio T, Girosi F. 1990. Networks for approximation and learning. Proc IEEE 78:1481–1497.

Powell MJD. 1985. Radial basis functions for multivariable interpolation: A review. In: IMA Conference on Algorithms for the Approximation of Functions and Data.

Shuler ML, Domach MM. 1983. Mathematical models of the growth of the individual cells. In: Blanch HW, Papoutsakis ET, Stephanopoulos G, editors. Foundations of biochemical engineering. Washington, DC: American Chemical Society. p 101.

Spendley W, Hext GR, Himsworth FR. 1962. Sequential application of simplex designs in optimization and evolutionary operation. Technometrics 4:441–461.

Sproule BA, Bazoon M, Shulman KI, Turksen IB, Naranjo CA. 1997. Fuzzy logic pharmacokinetic modeling: Application to lithium concentration prediction. Clin Pharmacol Ther 62:29–40.

Sugeno M, Kang GT. 1988. Structure identification of fuzzy model. Fuzzy Sets Syst 28:315–334.

Takagi T, Sugeno M. 1985. Fuzzy identification of systems and its applications to modeling and control. IEEE Trans Syst Man Cybernet 15: 116–132.

Waygood EB, Sanwal BD. 1974. The control of pyruvate kinases of *Escherichia coli.* J Biol Chem 249:265–274.

Wirsam B, Uthus EO. 1996. The use of fuzzy logic in nutrition. J Nutr 126(suppl):2337S–2341S.

Wright JA, Sanwal BD. 1969. Regulatory mechanisms involving nicotinamide adenine nucleotides as allosteric effectors. J Biol Chem 224: 1838–1845.

Yen J, Langari R, Zadeh LA. 1995. Industrial applications of fuzzy logic and intelligent systems. New York: IEEE Press.

Yen J, Langari R, Wang L. 1996. Principle components, b-splines, and fuzzy system reduction. Int J Uncertainty Fuzziness Knoledge-based Syst 4:561–572.

Yen J, Liao JC, Lee B, Randolph D. 1998. A hybrid approach to modeling metabolic systems using a genetic algorithm and simplex method. IEEE Trans Syst Man Cybernet B 28:173–191.