

Development of a computer program for analysis of enzyme kinetics by progress curve fitting

A.J.J. Straathof *

Delft University of Technology, Kluyver Laboratory for Biotechnology, Julianalaan 67, NL-2628 BC Delft, Netherlands

Abstract

In order to facilitate the study of enzyme kinetics by progress curve analysis, a convenient Windows 95 program was developed. For describing a set of progress curves, the user should select an enzymatic mechanism. This is guided by a menu. Reversible and irreversible reactions up to two substrates and products are supported. The program will automatically use the corresponding rate equation and fit the parameters of that equation at request by numerical integration of the batch differential equations. Degradation of reactants, enzyme inactivation and inhibition phenomena can easily be incorporated if required.

The program was tested for finding a rate equation that would describe a set of nine progress curves that was generated for a hypothetical enzymatic reaction according to an ordered bi–uni mechanism. The mechanistically correct equation fitted the progress curves best and gave good estimates for the five parameters involved. If there was Gaussian noise of 2% standard deviation superimposed on the simulated curves, the correct model still fitted satisfactorily, but it became impossible to discriminate it from some related incorrect models. © 2001 Elsevier Science B.V. All rights reserved.

Keywords: Kinetics; Progress curve fitting; Inhibition; Parameter estimation

1. Introduction

The optimization of an enzymatic process by a modeling approach requires a rate equation of the enzymatic reaction including its parameter values. The rate equation should be valid for the whole range of concentrations of the substrates, products and enzyme in a process. In many cases, the rate equation will contain a significant number of parameters, even when only the reactant concentrations are varied while fixing other potential variables, such as

pH, temperature and solvent type (as in this paper). Then, large experimental and computational efforts are required for a proper prediction of the reaction rates. The development of enzymatic processes could be greatly facilitated if the methods for obtaining enzymatic kinetic models would be simplified.

There are two types of enzymatic kinetic experiments [1]: transient (or pre-steady state) methods, which are of special interest if detailed mechanistic information is required; and steady-state methods, which are easier, but may lead to less detailed mechanistic information. Steady-state reaction rates can be analyzed in two manners [1]: initial rate analysis, using the traditional Lineweaver–Burk plots for example; and progress curve analysis.

* Tel.: +31-15-2782330; fax: +31-15-2782355.

E-mail address: straathof@stm.tudelft.nl (A.J.J. Straathof).

When one is interested in only a limited subset of the kinetic properties, e.g. in many mechanistic studies, initial rate analysis is preferred, because it is easier. However, if one requires the rate equation including the parameters for reaction equilibria, enzyme inactivation, product inhibition and background decomposition of substrate and product, then progress curve analysis becomes attractive. Progress curve analysis has been reviewed by Duggleby [2], and some recent works include Refs. [3–8]. A relatively modest set of progress curves may contain more information than a large set of initial rates, thus reducing experimental effort. To extract this potential information from the progress curves can be a major problem. First, rate equations that are expected to describe the experiments with a minimum number of parameters have to be derived. Usually, non-linear equations will be involved. Then, these equations have to be incorporated in computer programs that are able to simultaneously fit multiple parameters of multiple equations with multiple variables on multiple curves. Commercial software packages that are able to carry out most types of fits usually show severe limitations in this respect.

To facilitate progress curve analysis for non-specialists, the Windows program Encora is being developed. It is based on a powerful fitting algorithm that was described earlier [8]. In this paper, the background of the present program will be explained, and a study of the potential of progress curve analysis using this program will be presented.

2. Selection strategy for mechanism and rate equations

The approach described here assumes that one desires to use a rate equation that is based on a plausible enzymatic mechanism. (Otherwise, some kind of polynomial equation may be used to fit the progress curves, requiring no special software.) The main advantage of using a rate equation that is based on the correct mechanism of the enzymatic reaction is that the equation will have the correct structure. Predictions outside the range of concentrations that were used for parameter estimation are not specula-

tive and the physical meaning of the parameters will be clear.

Several selection steps towards a plausible mechanism, and hence, a rate equation will be formulated subsequently.

2.1. Stoichiometry

First, the stoichiometry of the enzymatic reaction has to be determined. Usually, this is trivial, because the stoichiometry is obvious from the E.C.-classification of the enzyme:

- | | |
|--------------------|---------------------------|
| 1. Oxidoreductases | bi–bi, ter–bi (or other) |
| 2. Transferases | bi–bi |
| 3. Hydrolases | bi–bi |
| 4. Lyases | bi–uni |
| 5. Isomerases | uni–uni |
| 6. Ligases | ter–ter (or more complex) |

An important additional stoichiometry is uni–bi. Uni–bi reactions occur when lyases are used in the reverse mode.

By covering kinetic mechanisms in which up to two substrates (A and B) and two products (P and Q) are involved in a single reaction, most biotransformations can be described. Ligases, which convert more than two substrates into more than two products, are hardly ever used for biotransformations. Oxidoreductases, however, are widely used, and a third substrate may be involved. Then, apparent kinetic parameters may be obtained for the two substrates if the concentration of the third substrate is not varied. An analogous situation applies to a reaction with a third product. Thus, a computer program that allows the selection of one or two substrates and one or two products should be adequate in almost all cases.

2.2. Order of reactants

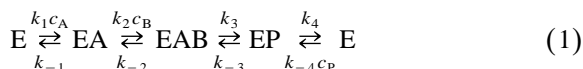
In general, by studying the steady state kinetics, the details of a mechanism cannot be detected, only the order (sequence) in which the substrates enter and products leave the active site. Consequently, this order determines which rate equation will be valid. For uni–uni reactions, there are no options for the order. Uni–bi, bi–uni and bi–bi reactions can either

be ordered or at random, and when they are ordered, it may matter which substrate or product will be first. The convention to use is that A is the first and B the second substrate to enter the enzyme, and P is the first and Q the second product to leave the enzyme [9]. For ordered bi–bi reactions, one additional option will be considered here: the first product can leave either before or after the second substrate has entered the active site (ping-pong and ternary complex mechanism, respectively). It seems reasonable to assume that when the substrates enter at random, products will also leave at random. Thus, a progress curve-fitting program should allow for these selections.

2.3. Rate-limiting step

There are two methods to derive rate equations from a reaction mechanism as selected according to the preceding sections, relying on the pseudo-steady-state assumption and the pseudo equilibrium assumption [10], respectively.

According to the pseudo-equilibrium assumption, the association/dissociation of any Michaelis complex is at equilibrium, and the reaction rate is determined by the breaking and making of covalent bonds. For example, for an ordered bi–uni reaction, Eq. 1, step 3 will be rate-limiting and at steady state, and steps 1, 2 and 4 will be at equilibrium.



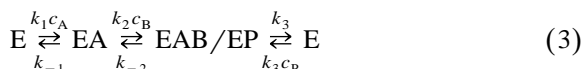
This leads to relatively simple rate equations, with three types of parameters: dissociation constants of the Michaelis complexes of A, B and P (indicated here by K_{mA} , K_{mB} and K_{mP}), a catalytic constant (k_{cat}^f) that equals the rate constant of the forward rate-determining step, and the equilibrium constant (K_{eq}).

$$-r_A = \frac{k_{cat}^f}{K_{mA} K_{mB}} (c_A c_B - c_P / K_{eq}) \quad (2)$$

$$1 + \frac{c_A}{K_{mA}} + \frac{c_A}{K_{mA}} \frac{c_B}{K_{mB}} + \frac{c_P}{K_{mP}}$$

Alternatively, the pseudo-steady-state assumption can be worked out. Then, any step may be a rate-

limiting step, and none is supposed to be at equilibrium. A step in which an isomerization occurs (i.e. no substrates enter and no products leave the enzyme) does not change the structure of a pseudo-steady-state rate equation. For simplicity, isomerization steps should be cancelled from the mechanism. Thus, for an ordered bi–uni reaction, the EAB and EP states can be lumped into a EAB/EP state:



Despite this reduction in the number of steps, the rate equations become less simple and contain more parameters, now also involving some inhibition constants (K_i 's). For the preceding mechanism,

$$-r_A = \frac{k_{cat}^f}{K_{iA} K_{mB}} (c_A c_B - c_P / K_{eq}) \quad (4)$$

$$1 + \frac{c_A}{K_{iA}} + \frac{K_{mA}}{K_{iA}} \frac{c_B}{K_{mB}} + \frac{c_A}{K_{iA}} \frac{c_B}{K_{mB}} + \frac{c_P}{K_{mP}} + \frac{c_B}{K_{iB}} \frac{c_P}{K_{mP}}$$

There may be constraints between the parameters that reduce the number of independent parameters [11], and in a progress curve-fitting program, these constraints must be taken into account. Still, there will usually be less parameters for the pseudo-equilibrium approach than for the pseudo-steady-state approach. The results of using either approach for progress curve fitting will be shown in a later section.

For random mechanisms, the pseudo-steady-state assumption is seldom used because it leads to extremely complex rate equations, involving numerous parameters. The constraints between the parameters have not been derived yet. Presently, a progress curve-fitting program can only include rate equations based on the assumption that in random mechanisms, bond breaking/making is rate-limiting. However, this is not an important limitation, because random mechanisms seem to occur much less frequently than ordered mechanisms. The rate equation thus obtained for a random bi–uni mechanism, differs from the

ordered equation (Eq. 2) only by an additional K'_{mB} -term, originating from the alternative pathway.

$$-r_A = \frac{k_{cat}^f}{K_{mA} K_{mB}} (c_A c_B - c_P / K_{eq})$$

$$= \frac{c_E}{1 + \frac{c_A}{K_{mA}} + \frac{c_A}{K_{mA}} \frac{c_B}{K_{mB}} + \frac{c_P}{K_{mP}} + \frac{c_B}{K'_{mB}}}$$
(5)

2.4. Irreversibility

When a reaction is irreversible, at least one of the microscopic steps involved in the mechanism will be irreversible. The rate equations may depend on which step is irreversible. For example, consider a uni–uni reaction via an EA/EP state. When the second step is irreversible, or when both steps are irreversible, the normal Michaelis–Menten equation is valid, but when only the first step is irreversible, this equation will include a product inhibition term. For each type of mechanism, each step may be either reversible or irreversible. These options must all be supported by a progress curve-fitting program, which leads to a large expansion of the number of rate equations that must be incorporated. However, in many cases, the same equations are valid.

In every case, where the step just before a single bond making/breaking step is irreversible, it was found that the equation for zero order kinetics in substrate is obtained. Then there is only one parameter: k_{cat}^f . This parameter equals the kinetic constant of the bond-making/breaking step (k_2 for uni-substrate or k_3 for bi-substrate reactions).

3. Implementation in a computer program

According to the aforementioned model selection strategy, the models of Table 1 were incorporated in the progress curve-fitting program Encora 1.2, which is downloadable from the Internet [12]. The total number of models is 204, but many end up in the same rate equation, e.g. the equation for the 0th order reaction. Still, the number of different equations is about 100.

Additional kinetic phenomena have also been incorporated. For substrate inhibition or inhibition by an exogenous inhibitor, additional denominator terms, corresponding to competitive or uncompetitive inhibition complexes, can be included in the rate equation. Note that the pseudo-steady-state assumption leads to equations in which product inhibition is intrinsically present, e.g. in Eq. 4. Allosteric effects have not been incorporated.

Chemical background reactions during enzymatic reactions are not uncommon. Each reactant can be

Table 1
Overview of enzymatic kinetic models incorporated in Encora

Stoichiometry	Order	Rate-limitation	Number of cases with different (ir)reversibility
Uni–uni	Ordered	Unknown	4
	Ordered	Known	8 (4 0th order)
Uni–bi	Ordered	Unknown	8
	Ordered	Known	16 (8 0th order)
	Random	Known	16 (8 0th order)
Bi–uni	Ordered	Unknown	8
	Ordered	Known	16 (8 0th order)
	Random	Known	16 (8 0th order)
Bi–bi	Ping-pong	Unknown	16
	Ping-pong	Known	16
	Ordered ternary complex	Unknown	16
	Ordered ternary complex	Known	32 (16 0th order)
	Random ternary complex	Known	32 (16 0th order)

selected to decompose by an (irreversible) first or second order rate equation. A special case is that the decomposition reaction has the same stoichiometry as the main enzymatic reaction. If the main reaction is reversible, this decomposition reaction must be reversible to the same degree because of thermodynamic restrictions. The program will use the equilibrium constant of the enzymatic reaction also for the chemical reaction.

Enzyme inactivation is important in many cases, and the number of mechanisms describing it is quite large. When the degradation can be described by a complicated function only, it is doubtful if progress curve analysis is suitable for the determination of the kinetic parameters of the enzymatic reaction. Therefore, the possibilities have been restricted to three types: degradation that is first order in enzyme concentration, second order in enzyme concentration, and first order in enzyme concentration and also in concentration of one of the reactants.

In some cases, a single enzyme may catalyze different reactions simultaneously, such as a kinetically controlled synthesis with concomitant undesired hydrolysis of substrate and product that cannot be prevented. Other examples include asymmetric syntheses, where the undesired enantiomer is formed to some extent, and kinetic resolution. Presently, only a single case of a simultaneous reaction is incorporated in Encora. This is an asymmetric synthesis by a reversible bi-uni reaction, leading to both enantiomers of the product. The enantiomeric ratio E becomes an additional parameter to be fitted in this case.

Progress curve fitting does not seem suitable for analyzing the kinetics of a mixture of enzymes. The enzymes should be studied one at a time.

4. The fitting procedure

4.1. Initiation

When the model has been selected, the fitting procedure has to be initiated. The program lists the parameters of the selected model equation(s). The user has to enter initial values, and has to indicate which parameters should be optimized. Some param-

eters (the equilibrium constant, and the decomposition and enzyme inactivation constants) can best be determined independently, so that the number of parameters to be optimized is as low as possible. The program has not been designed to work properly if the latter number exceeds six. It would be difficult to find the correct parameter values in such cases.

The values given as initial estimates of the parameters determine, to a large extent, the success of the fitting procedure. This is because the rate equations usually are non-linear, leading to local minima in the parameter space. When the number of parameters is large, the absolute minimum cannot be found easily. The chances to find the absolute minimum increase if the fitting procedure is carried out several times starting from very different initial estimates. Also, the step size in the iteration may be decreased or increased to find maxima in the parameter space that are otherwise missed. However, there is no known procedure that guarantees that the absolute minimum is found. When lack of good initial parameter estimates delays the fitting procedure, one may perform an initial rate analysis first.

4.2. Data handling

Encora requires a data set that includes the concentrations of all compounds involved. From these concentrations, the “relative concentration of P” is calculated. This is the absolute difference between the actual and initial concentration of P, subsequently divided by the initial concentration of A (in forward reactions) or P (in reverse reactions). In this way, the whole range from 0 to 1 can be covered for each progress curve, even when P is used as a product inhibitor or (in a reverse reaction) is in excess to Q. Hence, each curve may potentially get the same weight during fitting without having to use a different scaling for different curves in a plot of the curves.

4.3. Iteration

The fitting procedure itself is similar as in Ref. [8]. Every progress curve is simulated by numerically integrating the batch differential equations using a fourth order Runge–Kutta routine. Constraints are

Table 2
Initial concentrations used to simulate progress curves

Curve no.	c_{A0} (mol/l)	c_{B0} (mol/l)	c_{P0} (mol/l)	c_{E0} (mol/l)
1	0.1	0.1	0	0.0001
2	0.05	0.3	0	0.0001
3	0.3	0.05	0	0.0001
4	0.02	1	0	0.00005
5	1	1	0	0.0002
6	0	0	0.02	0.0001
7	0	0	0.1	0.0001
8	0	0	1	0.0001
9	0	2	0.5	0.00002

taken into account [11]. By comparing the relative concentration of P for the simulated and experimental data, a sum of squared residuals (SSres) is obtained for each progress curve. This is divided by the number of measurements of that curve, thus obtaining an estimate of the measurement variance. The total of these values is minimized, new estimates of the parameter set being generated by the Simplex-like algorithm of Nelder and Mead [13]. The stop criterion is that SSres does not change anymore or that the maximum number of iterations is reached.

5. Results and discussion

To demonstrate the potential use of Encora, a realistic but complicated kinetic problem was studied as an example. A hypothetical enzymatic bi–uni reaction according to the mechanism of Eq. 1 was taken as the starting point. For simplicity, it was assumed that $k_1 = 100 \text{ mol/l}^{-1} \text{ s}^{-1}$ and that each other rate constant in this mechanism has the value of $k = 10 \text{ s}^{-1}$ or $10 \text{ mol/l}^{-1} \text{ s}^{-1}$. Hence, one might

expect that the equations based on the assumption that the bond breaking/making step (step 3 in the mechanism) is rate-limiting, or that B can directly bind to the free enzyme, could have problems to describe the progress curves of such a reaction. This was tested as follows.

Using the eight rates corresponding to the eight rate constants and using the batch mass balances for A, B, P, E, EA, EAB and EP, numerical simulations of the progress of the formation (or breakdown) of P were performed, starting with different initial concentrations of the reactants as shown in Table 2. These simulations were carried out with very small integration time steps in order to deal with the steepness of the differential equations.

The simulations were used to generate nine data files consisting of 50 equidistant data points per progress curve. Each data point consists of a reaction time, the corresponding concentrations of A, B and P, and the enzyme concentration (which was assumed to remain constant).

In real experimental data, noise will be present which may trouble a parameter fitting procedure. To obtain a second set of nine curves, Gaussian-distributed noise was generated and added to the simulated concentration values of the reactants in such a way that the standard deviation was 2% of the exact concentrations.

The data files thus obtained are included in the downloadable version of Encora [12]. They were used for parameter estimation with different kinetic models as explained subsequently. The progress curves showed that the reaction proceeds to equilibrium. In such a case, to reduce the number of parameters to be fitted, one should try to obtain the equilibrium constant from independent equilibration experiments. Assuming this was accomplished, the value of the equilibrium constant was fixed at $K_{eq} =$

Table 3
SSres-values for different kinetic models, fitted to progress curves with and without noise

Model no.	Bi–uni order	Rate-limitation	SSres without noise	SSres with 2% noise
1	Ordered	Unknown	$8.79 \cdot 10^{-7}$	0.00148
2	Ordered incorrectly (B first)	Unknown	$6.97 \cdot 10^{-5}$	0.00124
3	Ordered	Step 3	0.00134	0.00175
4	Ordered incorrectly (B first)	Step 3	0.03580	0.03662
5	Random order	Step 3	$1.16 \cdot 10^{-5}$	0.00275

Table 4

Theoretical and fitted values of the steady-state parameters of model 1. Only five of the eight parameters were independently fitted, three were found via constraints [11]

Method	k_{cat}^f (s^{-1})	k_{cat}^r ($1 \text{ mol}^{-1} \text{ s}^{-1}$)	K_{mA} (mol/l)	K_{mB} (mol/l)	K_{mP} (mol/l)	K_{iA} (mol/l)	K_{iB} (mol/l)	K_{iP} (mol/l)
Theoretical	3.33	2.50	0.0333	1.000	0.750	0.100	2.00	0.500
Fit without noise	3.43	2.50	0.0348	1.035	0.749	0.0993	2.00	0.507
Fit with 2% noise	3.12	3.22	0.0253	0.879	0.952	0.105	2.10	0.549
Fit with noise after $t = 0$	3.34	2.73	0.0353	1.030	0.822	0.0975	1.60	0.462

10 for all fits. Fitting typically required a few minutes per case.

Table 1 lists only three different models for a bi–uni reaction, but if the order is not known, there are additional wrong-order options for the ordered models, so the total number of models that can be used for fitting is five. Table 3 shows the sum of SSres-values that were obtained for fitting to these five models.

Surprisingly, only model 4, which assumes the incorrect order and a rate limitation, did not lead to a good fit of the data. The other fits were visually good. In the absence of noise, the low SSres-value for model 1 reflects that this is the best model; it assumes the correct order and no rate limitation. However, in the presence of noise, the performance of models 1, 2, 3 and 5 is in the same range, and discrimination is not possible. The difficulty to discriminate between the models is caused by the large number of parameters that is fitted (four for models 3 and 4; and five for the others). In general, the difference between the denominators of the model equations (Eqs. 2, 4 and 5) is small. In the present case, the numerical values of the constants in the denominator are such that discrimination is possible only when very accurate data are available. Perhaps, model discrimination would also be possible for the noise-containing data with a much larger amount of data.

Suppose the correct model is known from independent data. Then it is interesting to find out if progress curve analysis can lead to reliable parameter estimates for this model. The theoretical values of the steady-state parameters of model 1 were calculated by deriving the rate equation for the mechanism of Eq. 1 and entering the rate constants [14]. In

Table 4, these theoretical values are compared to the fitted values.

For the fits to the progress curves without noise, the parameter values were found back with deviations in the range of only 0–4%, which shows that the calculational procedures were carried out correctly. The highest errors are in k_{cat}^f and K_{mA} . These parameters are clearly correlated, suggesting that additional curves at low concentrations of A could improve the fits. For the fits to the progress curves with 2% noise, the deviation was 5–25%, which will be satisfactory in many cases. In an additional series of fits, the noise was removed from the data point at $t = 0$, while maintaining it in the other 49 data points of each curve. This led to a reduction of the deviations to 1–10%, except for the value of K_{iB} , which still deviated by 20%. Clearly, the accuracy of the initial values determines the overall accuracy to a large extent, confirming the previous analyses [2].

Upon removal of the noise from the initial values, the model discrimination procedure was still not satisfactory using the present data.

6. Conclusions

The computer program Encora [12] can be used for parameter estimation of enzymatic reactions using progress curve analysis. Forward and reverse reaction curves can be fitted simultaneously, and degradation reactions, enzyme inactivation and inhibition phenomena can be included in the models without having to study the kinetic equations involved.

A study of a hypothetical bi–uni reaction showed that good parameter estimates can be obtained by progress curve analysis with measurements that contain (except for the initial concentrations) 2% noise. Model discrimination, however, requires a much higher accuracy.

Acknowledgements

This work has partly been supported financially by the EU (BIO4 CT960112).

References

- [1] A. Cornish-Bowden, *Fundamentals of Enzyme Kinetics*, Portland Press, Colchester, 1995, revised edition.
- [2] R.G. Duggleby, *Methods Enzymol.* 249 (1995) 61.
- [3] M.R. Schiller, L.D. Holmes, E.A. Boeker, *Biochim. Biophys. Acta* 1297 (1996) 17.
- [4] P. Mendes, D.B. Kell, *Bioinformatics* 14 (1998) 869.
- [5] C.T. Goudar, J.R. Sonnad, R.G. Duggleby, *Biochim. Biophys. Acta* 1429 (1999) 377.
- [6] M. Bauer, H. Griengl, W. Steiner, *Biotechnol. Bioeng.* 62 (1999) 20.
- [7] A. Bódalo, J.L. Gómez, E. Gómez, J. Bastida, G. León, M.F. Máximo, A.M. Hidalgo, M.C. Montiel, *Enzyme Microb. Technol.* 24 (1999) 381.
- [8] J.L.L. Rakels, B. Romein, A.J.J. Straathof, J.J. Heijnen, *Biotechnol. Bioeng.* 43 (1994) 411.
- [9] W.W. Cleland, *Biochim. Biophys. Acta* 67 (1963) 104.
- [10] J.A. Roels, *Energetics and Kinetics in Biotechnology*, Elsevier, Amsterdam, 1983.
- [11] A.J.J. Straathof, J.J. Heijnen, *Biotechnol. Bioeng.* 52 (1996) 433.
- [12] R.J.W. Slats, J.W. Hennipman, B. Romein, A.J.J. Straathof, *Encora 1.2*, Delft University of Technology, The Netherlands (1999), Freely downloadable from <http://www.kluyver.stm.tudelft.nl/miscell/freewf.htm>.
- [13] D.M. Himmelblau, *Process Analysis by Statistical Methods*, Wiley, New York, 1990.
- [14] A.J.J. Straathof, J.J. Heijnen, *Biocatal. Biotransform.* 15 (1997) 29.