A computer program for enzyme kinetics that combines model discrimination, parameter refinement and sequential experimental design

Rafael FRANCO, Maria Teresa GAVALDÀ and Enrique I. CANELA

Departament de Bioquímica, Facultat de Química, Universitat de Barcelona, Martí i Franques 1, 08028 Barcelona, Catalunya, Spain

A method of model discrimination and parameter estimation in enzyme kinetics is proposed. The experimental design and analysis of the model are carried out simultaneously and the stopping rule for experimentation is deduced by the experimenter when the probabilities *a posteriori* indicate that one model is clearly superior to the rest. A FORTRAN77 program specifically developed for joint designs is given. The method is very powerful, as indicated by its usefulness in the discrimination between models. For example, it has been successfully applied to three cases of enzyme kinetics (a single-substrate Michaelian reaction with product inhibition, a single-substrate complex reaction and a two-substrate reaction). By using this method the most probable model and the estimates of the parameters can be obtained in one experimental session. The FORTRAN77 program is deposited as Supplementary Publication SUP 50134 (19 pages) at the British Library (Lending Division), Boston Spa, Wetherby, West Yorkshire LS23 7BQ, U.K., from whom copies can be obtained on the terms indicated in Biochem. J. (1986) 233, 5.

INTRODUCTION

Currently, the kineticist has two main basic areas of interest: elucidation of the kinetic mechanism of the enzymes and estimation of the unknown parameter values with the maximum of accuracy and precision.

At present, the pursuit of these goals, in enzyme kinetics, has been independently attempted by using optimal experimental designs both for refinement of the parameters according to the Box–Lucas (1959) criterion (Duggleby, 1979, 1981; Endrenyi, 1981; Currie, 1982; Canela, 1985) and discrimination among rival models (Reich, 1970; Bartfai & Mannervik, 1972; Mannervik, 1981, 1982). With these methods, several initial rates are taken; after these measurements the data obtained are analysed, and from the results the planning of a further stage begins, and so on. However, these measurements are taken on different days and, hence, though the design of the experiment is optimal, the experimental error may introduce bias into the parameter values.

Likewise, attempts have been made in order to estimate parameter values without a sufficient discrimination among models. Thus, in the literature, single models for many oligomeric enzymes have been described. Although they could have more complex behaviour, a Michaelis-like behaviour can be found when an insufficient range of substrate concentration is employed (Hill *et al.*, 1977; Bardsley *et al.*, 1980a).

In the present paper a program written in FORTRAN77 that combines the determination of the true model and the refinement of the estimates of parameter values, by using the criterion of **D** optimality, is described. Likewise, a useful experimental strategy that permits positive results in only a single experimental session if the kineticist has a computer at his disposal is stated.

THEORY

The kinetic models of the enzymes can be described as $v = f(s, \theta)$, where v , s and θ denote velocities, substrate concentrations and parameters respectively. These models are non-linear in some parameters, i.e. the derivative of v with respect to parameter contains it. Therefore the dispersion matrix of the estimates, $D(\theta)$, depends on the estimated values of θ .

When the regression analysis of the model is performed, the solution of the iterative process, in matricial notation, is:

$$\hat{\theta} - \theta = M^{-1} \cdot Y \quad (1)$$

being:

$$M = F \cdot W \cdot F' \quad (2)$$

where the prime ' denotes transposed matrix, and:

$$Y = F \cdot R \quad (3)$$

where **F** denotes a matrix of the form:

$$F_{ij} = \partial f(S_j, \theta) / \partial \theta_i \quad (4)$$

and **W** is the matrix

$$W_{ij} = w_i \delta_{ij} \quad (5)$$

where $\delta_{ij} = 1$ if $i = j$, or otherwise = 0 (Kronecker delta), where:

$$R_i = v_i - f_i(S_i, \theta) \quad (6)$$

but:

$$M^{-1} = D(\theta) \quad (7)$$

only if (Fedorov, 1972):

$$\sum_{i=1}^n \left| \frac{\partial f_i(S_i, \theta)}{\partial \theta_j} \cdot \frac{\partial P_i(S_i, \theta)}{\partial \theta_k} \right|_{\theta = \hat{\theta}} \gg \sum_{i=1}^n \left| w_i^{\frac{1}{2}} \cdot \frac{\partial^2 f_i(S_i, \theta)}{\partial \theta_j \cdot \partial \theta_k} \right|_{\theta = \hat{\theta}} \quad (8)$$

However, if the number of determinations is small this inequality is not satisfied [for review of regression analysis see Draper & Smith (1981) or Watts (1981)].

Optimal designs for estimating parameter values or for discriminating among rival models consist of a set of several points that minimizes (or maximizes) a given function (in general it minimizes a magnitude related with the dispersion matrix of the parameter estimates) (Fedorov, 1972). However, as indicated above, $D(\theta)$ depends upon the parameter values. This leads to inaccuracy in the parameter values, which leads to an unsuitable design.

The immediate question is: how can the kineticist design the experiment to obtain positive results? In order to answer this question the first point to consider is that in a preliminary experiment the parameter values can be inaccurate owing to the lack of optimal design (linear, $S_i = S_0 + a \cdot i$, for $i = 1$ to n , or geometric, $S_i = S_0 \cdot a^{i-1}$ for $i = 1$ to n , n being the number of substrate concentrations and a the progression ratio) and errors associated with the determinations. The second point to consider is that a discrete design obtained from previous results would be suboptimal and lead to incorrect results.

Generally, the designs for optimization in enzyme kinetics are obtained by selecting a set of points from a preliminary experiment, carrying out the experiment and obtaining estimates of the parameters by regression analysis, then designing another experiment, and so on, until the results of two consecutive well-designed experiments agree.

This system has a major problem: it would never finish. The question is: why does this occur? The answer is simple: the errors associated with the measure are randomly distributed and, hence, they lead to different results. Thus errors vary from day to day and the results also. Then, the new question is: how is this problem solved? The solution is to carry out an experiment planned with a sequential design. In this design a set of determinations is performed, and then the data are fitted to the corresponding models. From the information obtained, a new point is chosen and a new measure is performed. The set of data, including this new measure, is fitted to the corresponding equation. The process is repeated until sufficient or no further improvement can be achieved.

However, before refining the estimates of the parameters it is necessary to be sure of the model to which the investigated enzyme conforms. This is a controversial question, and a number of interesting publications about the theory of enzyme models and about the problem of finding the real model from a data set are available (Bardsley & Childs, 1975; Bardsley & Crabbe, 1976; Childs & Bardsley, 1976; Bardsley, 1976, 1977a,b; Whitehead, 1976, 1978; Bardsley *et al.*, 1980b; Solano-Muñoz, 1981). Also, the difficulty of achieving this aim has been stated (Burguillo *et al.*, 1983; Del Arco *et al.*, 1984). For example, many mechanistic models of enzyme kinetics do not agree with the real model, although they explain the behaviour of the enzymes in the studied range of substrate concentrations. These differences are due to the difficulty encountered in fitting the data to a complex model, and the unavailability of powerful computers and of suitable techniques for fitting the data. From a theoretical point of view oligomeric enzymes have complex rate equations. In many cases they follow Michaelian kinetics, but some of them have a Michaelian-

like behaviour only when a narrow range of substrate concentrations is used. Thus it seems logical that, having the suitable tools, the kineticist should study a wider range of substrate concentrations and test other models of a higher degree, and not merely 'forcing' the data to conform the simplest one. Thus, upon studying the behaviour of an enzyme, the first idea should be to consider different possibilities assigning to each model identical probability *a priori*.

Assume that a given model is correct; the probability density fraction of m parameters equals (Fedorov, 1972):

$$p(\theta) = \exp[-\frac{1}{2}S - \frac{1}{2}(\hat{\theta} - \theta)' \cdot D^{-1}(\hat{\theta} - \theta)] \quad (9)$$

being:

$$S = R' \cdot W \cdot R \quad (10)$$

It should be noted that on diminishing the value of $|D|$ the likelihood of the parameter values increases.

Now, considering h different possible models (denoted by H_j , $j = 1, 2, \dots, h$), with equal probability *a priori* and taking into account that:

$$\sum_{i=1}^h \int p_i(\theta) \cdot d\theta = 1 \quad (11)$$

the probability of each model is obtained by (Fedorov, 1972):

$$p(H_j) = \frac{(2\pi)^{-m_j/2} \cdot |D_j|^{-\frac{1}{2}} \cdot \exp(-\frac{1}{2}S_j)}{\sum_{i=1}^h (2\pi)^{-m_k/2} \cdot |D_k|^{-\frac{1}{2}} \cdot \exp(-\frac{1}{2}S_k)} \quad (12)$$

It is obvious that if the value of $|D_j|$ diminishes but the values of $|D_k|$ increase for $k = 1, \dots, h$ and $k \neq j$, the value of $p(H_j)$ increases.

From the above it follows that performing determinations in the suitable zone should increase $|D_j|$, both if the purpose is estimating parameter values and if the purpose is also discriminating among rival models.

The question now is: how is the point selected for diminishing the value of $|D_j|$? The answer is in the following expression (Fedorov, 1972):

$$|D_{n+1}| = |D_n| / [1 + \lambda(x) \cdot d(x)] \quad (13)$$

being $\lambda(x)$ the efficiency function. The efficiency function is the function that transforms the absolute error into relative error. In this manner, a comparison between the measurements at different points is possible. In general, it can be substituted by the expected weight. $d(x)$ corresponds to the dispersion of the new selected point, defined as $F'(x) \cdot D \cdot F(x)$ [$F(x)$ corresponds to F applied to the new point x].

It is observed that, when $d(x)$ achieves the maximum value, maximum decrement of $|D|$ during performance of the new determination is obtained.

This is of small interest when the goal is also the discrimination: this expression indicates the form of diminishing the value of $|D_j|$ but says nothing about the remaining $|D|$ values.

From this it is necessary to consider the generalized dispersion defined as (Fedorov, 1972):

$$D(x+1) = \sum_{j=1}^p (H_j) \cdot d_j(x) + \sum_{j=1}^p (H_j) \cdot [f_j(x, \theta) - \sum_{k=1}^p (H_k) \cdot f_k(x, \theta)]^2 \quad (14)$$

The first term in the right-hand side of the expression (14)

Table 1. One-substrate-one-product mechanism with product mixed inhibition

Model 1, model 2, model 3 and model 4 correspond to competitive, uncompetitive, mixed and non-competitive inhibition mechanisms respectively. Velocity indicates the experimental velocities (perturbed exact velocities), $|D|$ denotes the determinant of the dispersion matrix of the estimates, and probability the probability *a posteriori* of each model.

Concn. of substrate (M)		Velocity ($\mu\text{mol/min}$)		Model 1		Model 2		Model 3		Model 4	
A	B			$ D $	Probability	$ D $	Probability	$ D $	Probability	$ D $	Probability
1.00×10^{-3}	0	0.974									
5.00×10^{-4}	0	0.960									
1.00×10^{-4}	0	0.700									
5.00×10^{-5}	0	0.543									
1.00×10^{-5}	0	0.193									
5.00×10^{-6}	0	0.114									
1.00×10^{-3}	1.00×10^{-4}	0.330									
5.00×10^{-4}	1.00×10^{-4}	0.357									
1.00×10^{-4}	1.00×10^{-4}	0.278									
5.00×10^{-5}	1.00×10^{-4}	0.230									
1.00×10^{-5}	1.00×10^{-4}	0.135									
5.00×10^{-6}	1.00×10^{-4}	0.070									
1.00×10^{-3}	2.50×10^{-4}	0.198									
5.00×10^{-4}	2.50×10^{-4}	0.171									
1.00×10^{-4}	2.50×10^{-4}	0.158									
5.00×10^{-5}	2.50×10^{-4}	0.150									
1.00×10^{-5}	2.50×10^{-4}	0.109									
5.00×10^{-6}	2.50×10^{-4}	0.051									
1.00×10^{-3}	4.00×10^{-4}	0.110									
5.00×10^{-4}	4.00×10^{-4}	0.100									
1.00×10^{-4}	4.00×10^{-4}	0.114									
5.00×10^{-5}	4.00×10^{-4}	0.109									
1.00×10^{-5}	4.00×10^{-4}	0.061									
5.00×10^{-6}	4.00×10^{-4}	0.048									
6.11×10^{-4}	1.10×10^{-7}	0.984									
4.43×10^{-5}	5.11×10^{-7}	0.480									
1.00×10^{-3}	3.40×10^{-6}	0.936									
5.00×10^{-4}	6.06×10^{-6}	0.898									
1.00×10^{-3}	1.13×10^{-6}	0.968									
7.37×10^{-5}	2.66×10^{-6}	0.588									
8.80×10^{-4}	1.15×10^{-6}	0.972									
5.48×10^{-4}	1.67×10^{-7}	0.951									
9.88×10^{-4}	1.75×10^{-6}	0.957									
2.36×10^{-5}	8.63×10^{-7}	0.308									
8.22×10^{-5}	6.87×10^{-7}	0.632									
1.00×10^{-3}	5.97×10^{-6}	0.471									
3.93×10^{-5}	2.93×10^{-6}	0.453									
8.09×10^{-4}	1.25×10^{-6}	0.980									
1.00×10^{-3}	5.15×10^{-6}	0.512									
1.55×10^{-5}	2.65×10^{-6}	0.036									
1.35×10^{-5}	2.91×10^{-6}	0.038									
1.86×10^{-5}	1.67×10^{-6}	0.053									
2.71×10^{-5}	8.13×10^{-6}	0.144									
4.23×10^{-5}	9.78×10^{-6}	0.199									
3.31×10^{-5}	6.78×10^{-6}	0.195									
6.23×10^{-5}	2.53×10^{-6}	0.567									
1.00×10^{-3}	5.24×10^{-6}	0.479									
8.69×10^{-4}	1.79×10^{-6}	0.958									
4.23×10^{-5}	7.00×10^{-6}	0.207									
1.00×10^{-3}	9.89×10^{-6}	0.984									
				7.30×10^{-11}	0.000						
				4.45×10^{-11}	0.248						
				1.85×10^{-11}	0.299						
				8.52×10^{-13}	0.394						
				7.35×10^{-13}	0.389						
				4.13×10^{-13}	0.387						
				3.88×10^{-13}	0.383						
				3.03×10^{-13}	0.419						
				2.28×10^{-13}	0.416						
				2.06×10^{-13}	0.413						
				1.98×10^{-13}	0.411						
				1.88×10^{-13}	0.425						
				1.37×10^{-13}	0.445						
				1.23×10^{-13}	0.473						
				1.13×10^{-13}	0.463						
				9.14×10^{-13}	0.475						
				7.71×10^{-13}	0.498						
				7.67×10^{-13}	0.140						
				7.64×10^{-13}	0.050						
				7.59×10^{-13}	0.005						
				7.07×10^{-13}	0.000						
						5.90×10^{-13}	0.752	4.98×10^{-11}		4.98×10^{-11}	0.000
						2.84×10^{-13}	0.701	2.01×10^{-11}		2.01×10^{-11}	0.000
						1.78×10^{-13}	0.606				
						1.51×10^{-13}	0.611				
						8.43×10^{-13}	0.613				
						7.96×10^{-13}	0.617				
						6.95×10^{-13}	0.581				
						5.17×10^{-13}	0.584				
						4.64×10^{-13}	0.587				
						4.42×10^{-13}	0.589				
						4.41×10^{-13}	0.575				
						3.53×10^{-13}	0.552				
						3.52×10^{-13}	0.527				
						3.09×10^{-13}	0.537				
						2.64×10^{-13}	0.525				
						2.47×10^{-13}	0.502				
						6.76×10^{-14}	0.859				
						4.57×10^{-14}	0.950				
						2.72×10^{-14}	0.995				
						8.73×10^{-16}	1.000				
						8.08×10^{-16}	1.000				
						5.94×10^{-16}	1.000				
						5.18×10^{-16}	1.000				
						5.00×10^{-16}	1.000				
						4.61×10^{-16}	1.000				
						3.75×10^{-16}	1.000				
						3.09×10^{-16}	1.000				

corresponds to the weighted dispersion of the experiment; it should be noted that it is the sum of the probability of each model multiplied by its dispersion. The second term in the right hand side is the expectation of the squared differences between the velocity calculated from the model and the expectation of the velocity (obtained by weighting the velocity from each model with its probability). The increment of this term leads to the maximum increment of the sum of squares of the regression.

It should be noted that when the probability of one model is practically 1 (the remaining are negligibles) the generalized dispersion becomes $d(x)$.

However, the values of $D(x+1)$ must be weighted, and therefore:

$$D(x+1) \cdot w = \text{objective function} \quad (15)$$

for maximizing.

In accordance with the above, we have developed a program useful for enzyme kinetics that computes the values of concentrations that maximize the weighted generalized dispersion both for discrimination and refinement of parameter values.

COMPUTATIONAL METHOD

The programs used in this paper have been written in FORTRAN77 and implemented on a VAX 11/750 computer and on a IBM 3083 computer. All the programs were developed in this laboratory. The program used double precision in the VAX 11/750 computer and extended precision in the IBM 3083 computer. The reason for using extended precision was not for increasing the accuracy but for avoiding computation errors (overflow or underflow) due to the excessively high or low values of some numbers. If it is available, we recommend the use of G floating point form of numerical representation suitable for VAX 11/750 computers. This allows work with numbers ranging between 5.6×10^{-309} and 9.0×10^{307} . In this form the real $\times 8$ numbers have a degree of precision of 15 decimal digits instead the typical value of 16.

The program consists of a main program with an input section in which the user supplies the computer with the data [substrate concentrations, velocities, initial estimates of the parameters for non-linear regression (one set of parameters for each considered model), the number of models and the parameters of the model]. The main program calls the subroutine AJUST, which calls for each model the subroutine REGRE, which carries out the fit of the data by non-linear regression. This subroutine is based on the program published by Canela (1984), based on Marquardt's (1963) algorithm. When the regression for the whole of the models is ended, the control is transferred to the main program, which computes the probabilities of the models. Then the program calls the subroutine DISPUN, which computes the point that gives the maximum value of the product generalized dispersion by weight. The last part is the output, in which analysis regression with the confidence intervals, the determinants of the dispersion matrices of the parameters and the calculated point for analysis *a posteriori* are presented.

Obviously, with a set of minor modifications the Jacobians, complete matrices etc. would be obtained.

Fig. 1. Variation of the accuracy of values of parameters with the number of determinations

(a) One-substrate-one-product mechanism with product mixed inhibition. (b) Sequential bi bi mechanism. (c) 2:2 one-substrate mechanism.

RESULTS

Three cases of enzyme kinetic mechanisms were stimulated to show the performances of the propounded method.

Table 2. Sequential bi bi mechanism

Model 1 and model 2 correspond to ping-pong bi bi and sequential bi bi mechanisms respectively. Velocity indicates the experimental velocities (perturbed exact velocities), $|D|$ denotes the determinant of the dispersion matrix of the estimates, and probability the probability *a posteriori* of each model.

Concn. of substrate (M)		(Velocity) ($\mu\text{mol/min}$)	Model 1		Model 2	
A	B		$ D $	Probability	$ D $	Probability
1.00×10^{-3}	1.00×10^{-3}	1.024				
5.00×10^{-4}	1.00×10^{-3}	0.916				
1.00×10^{-4}	1.00×10^{-3}	0.568				
1.00×10^{-5}	1.00×10^{-3}	0.098				
5.00×10^{-6}	1.00×10^{-3}	0.070				
1.00×10^{-3}	5.00×10^{-4}	0.968				
5.00×10^{-4}	5.00×10^{-4}	0.906				
1.00×10^{-4}	5.00×10^{-4}	0.549				
1.00×10^{-5}	5.00×10^{-4}	0.121				
5.00×10^{-6}	5.00×10^{-4}	0.053				
1.00×10^{-3}	1.00×10^{-4}	0.808				
5.00×10^{-4}	1.00×10^{-4}	0.761				
1.00×10^{-4}	1.00×10^{-4}	0.466				
1.00×10^{-5}	1.00×10^{-4}	0.102				
5.00×10^{-6}	1.00×10^{-4}	0.048				
1.00×10^{-3}	1.00×10^{-5}	0.272				
5.00×10^{-4}	1.00×10^{-5}	0.277				
1.00×10^{-4}	1.00×10^{-5}	0.218				
1.00×10^{-5}	1.00×10^{-5}	0.069				
5.00×10^{-6}	1.00×10^{-5}	0.044				
1.00×10^{-3}	5.00×10^{-6}	0.161				
5.00×10^{-4}	5.00×10^{-6}	0.155				
1.00×10^{-4}	5.00×10^{-6}	0.126				
1.00×10^{-5}	5.00×10^{-6}	0.050				
5.00×10^{-6}	5.00×10^{-6}	0.021	4.00×10^{-13}	0.007	1.66×10^{-14}	0.993
1.00×10^{-3}	2.13×10^{-5}	0.452	3.25×10^{-13}	0.006	1.28×10^{-14}	0.994
8.45×10^{-4}	2.58×10^{-5}	0.497	2.61×10^{-13}	0.006	1.01×10^{-14}	0.994
9.23×10^{-4}	2.92×10^{-5}	0.530	1.37×10^{-13}	0.005	5.11×10^{-15}	0.995
1.00×10^{-3}	1.00×10^{-3}	1.001	1.13×10^{-13}	0.005	4.20×10^{-15}	0.995
1.00×10^{-3}	1.00×10^{-3}	0.996	1.01×10^{-13}	0.004	3.71×10^{-15}	0.996
7.88×10^{-5}	7.83×10^{-4}	0.505	8.83×10^{-14}	0.003	3.11×10^{-15}	0.997
1.00×10^{-3}	1.00×10^{-3}	0.999	5.37×10^{-14}	0.000	1.86×10^{-15}	1.000
4.71×10^{-5}	1.58×10^{-5}	0.203	4.98×10^{-14}	0.000	1.16×10^{-15}	1.000
8.42×10^{-5}	1.00×10^{-3}	0.509			9.94×10^{-16}	1.000
2.61×10^{-5}	9.13×10^{-3}	0.122			9.00×10^{-16}	1.000
1.00×10^{-3}	3.48×10^{-5}	0.565			8.46×10^{-16}	1.000
7.33×10^{-5}	1.00×10^{-3}	0.479			7.78×10^{-16}	1.000
1.00×10^{-3}	3.00×10^{-5}	0.419			7.15×10^{-16}	1.000
2.65×10^{-5}	1.23×10^{-5}	0.135			6.04×10^{-16}	1.000
8.83×10^{-5}	9.95×10^{-4}	0.515			4.52×10^{-16}	1.000
1.00×10^{-3}	2.61×10^{-5}	0.510			2.55×10^{-16}	1.000
1.00×10^{-3}	1.00×10^{-3}	0.996			2.39×10^{-16}	1.000
1.00×10^{-3}	1.00×10^{-3}	1.014			2.25×10^{-16}	1.000
1.00×10^{-3}	1.00×10^{-3}	0.995			2.11×10^{-16}	1.000
4.79×10^{-5}	1.29×10^{-5}	0.198			2.03×10^{-16}	1.000
1.00×10^{-3}	1.00×10^{-3}	1.011			1.84×10^{-16}	1.000
8.09×10^{-3}	1.00×10^{-3}	0.515			1.56×10^{-16}	1.000
1.99×10^{-5}	1.61×10^{-5}	0.117			1.43×10^{-16}	1.000
1.00×10^{-3}	2.15×10^{-5}	1.007			1.34×10^{-16}	1.000
7.50×10^{-5}	2.15×10^{-5}	0.272			9.69×10^{-17}	1.000

One-substrate-one-product mechanism with product mixed inhibition

The possible mechanism considered with equal probability *a priori* were: competitive inhibition, uncompetitive inhibition, non-competitive inhibition and mixed inhibition.

The preliminary experiment was of 24 points: four series of six substrate concentrations ($[\text{substrate}] = 5, 10, 50, 100, 500$ and $1000 \mu\text{M}$) with that of the product held constant ($[\text{product}] = 0, 100, 250$ and $400 \mu\text{M}$). The allowed experimental concentration of substrate ranged between 1.00 mM and $5.00 \mu\text{M}$; the corresponding one for the product ranged between $400 \mu\text{M}$ and 0 . The exact parameter values were $V = 1.05 \mu\text{mol/min}$, $K_m = 50 \mu\text{M}$, $K_{ip} = 100 \mu\text{M}$ and $K_{ii} = 50 \mu\text{M}$. The exact velocities from

the equation were randomly perturbed ($\mu = 0$, $\delta = 2.5 \times \text{maximum experimental velocity}$) and five replicates for each point were made.

The results appear in Table 1. It can be observed that the probability 1 for the correct model was obtained after the 43th determination. Refinement was then carried out. After the 43th point the improvement of K_m , K_{ip} and K_{ii} values was hardly significant (Fig. 1a). The last parameters estimates were $V = 1.05 \mu\text{mol/min}$, $K_m = 51.2 \mu\text{M}$, $K_{ip} = 91.0 \mu\text{M}$ and $K_{ii} = 55.6 \mu\text{M}$.

Ordered bi bi mechanism

The possible mechanisms considered with equal probability *a priori* were: ordered and ping-pong.

The preliminary experiment was of 25 points: five

Table 3. 2:2 one-substrate mechanism

Model 1, model 2, model 3 and model 4 correspond to 1:1, 1:2, 2:2 and 2:3 one-substrate mechanisms respectively. Velocity indicates the experimental velocities (perturbed exact velocities), $|D|$ denotes the determinant of the dispersion matrix of the estimates, and probability the probability *a posteriori* of each model.

Concn. of substrate (M)	Velocity (μmol/min)	Model 1		Model 2		Model 3		Model 4	
		D	Probability	D	Probability	D	Probability	D	Probability
1.00 × 10 ⁻³	0.330	3.95 × 10 ⁻⁸	0.000	4.99 × 10 ⁻¹¹	0.490	1.58 × 10 ⁻¹¹	0.510	5.62 × 10 ⁻¹¹	0.000
5.55 × 10 ⁻⁴	0.465			2.79 × 10 ⁻¹¹	0.073	2.15 × 10 ⁻¹²	0.927	1.65 × 10 ⁶	0.000
3.08 × 10 ⁻⁴	0.694			2.25 × 10 ⁻¹¹	0.057	1.51 × 10 ⁻¹²	0.943		
1.71 × 10 ⁻⁴	0.902			1.87 × 10 ⁻¹¹	0.030	1.04 × 10 ⁻¹²	0.970		
9.49 × 10 ⁻⁵	0.996			1.53 × 10 ⁻¹¹	0.029	8.29 × 10 ⁻¹³	0.971		
5.27 × 10 ⁻⁵	0.921			1.25 × 10 ⁻¹¹	0.022	6.32 × 10 ⁻¹³	0.978		
2.92 × 10 ⁻⁵	0.658			8.50 × 10 ⁻¹²	0.117	6.41 × 10 ⁻¹³	0.883		
1.62 × 10 ⁻⁵	0.431			7.62 × 10 ⁻¹²	0.092	5.16 × 10 ⁻¹³	0.907		
9.01 × 10 ⁻⁶	0.288			6.35 × 10 ⁻¹²	0.107	4.52 × 10 ⁻¹³	0.893		
5.00 × 10 ⁻⁶	0.176			5.26 × 10 ⁻¹²	0.061	3.00 × 10 ⁻¹³	0.939		
1.00 × 10 ⁻³	0.330			4.84 × 10 ⁻¹²	0.061	2.75 × 10 ⁻¹³	0.939		
7.03 × 10 ⁻⁵	0.972			4.21 × 10 ⁻¹²	0.041	2.10 × 10 ⁻¹³	0.959		
7.07 × 10 ⁻⁵	0.995			3.45 × 10 ⁻¹²	0.025	1.52 × 10 ⁻¹³	0.973		
1.88 × 10 ⁻⁵	0.498			3.08 × 10 ⁻¹²	0.024	1.32 × 10 ⁻¹³	0.976		
1.00 × 10 ⁻³	0.319			2.83 × 10 ⁻¹²	0.018	1.12 × 10 ⁻¹³	0.982		
2.80 × 10 ⁻⁴	0.763			2.79 × 10 ⁻¹²	0.016	1.07 × 10 ⁻¹³	0.984		
7.57 × 10 ⁻⁵	1.001	2.57 × 10 ⁻¹²	0.018	1.02 × 10 ⁻¹³	0.981				
1.80 × 10 ⁻⁵	0.501	1.45 × 10 ⁻¹²	0.000	2.57 × 10 ⁻¹⁴	1.000				
1.00 × 10 ⁻³	0.327			2.33 × 10 ⁻¹⁴	1.000				
7.54 × 10 ⁻⁵	0.979			8.04 × 10 ⁻¹⁶	1.000				
1.78 × 10 ⁻⁵	0.447			7.43 × 10 ⁻¹⁶	1.000				
7.52 × 10 ⁻⁵	0.996								
2.06 × 10 ⁻⁵	0.518								
2.80 × 10 ⁻⁴	0.736								
1.00 × 10 ⁻³	0.337								
7.90 × 10 ⁻⁵	0.975								
1.00 × 10 ⁻³	0.354								
1.99 × 10 ⁻⁵	0.528								
7.97 × 10 ⁻⁵	1.002								
2.11 × 10 ⁻⁵	0.548								

series of five concentrations of substrate A ([substrate A] = 5, 10, 100, 500 and 1000 μM) with that of substrate B held constant ([substrate B] = 5, 10, 100, 500 and 1000 μM). The allowed experimental concentrations ranged between 1.00 mM and 5.00 μM for both substrates. The exact parameters values were $V = 1.13 \mu\text{mol/min}$, $K_a = 30 \mu\text{M}$, $K_b = 100 \mu\text{M}$, and $K_{ab} = 0.4 \text{ nM}$. The exact velocities from the equation were randomly perturbed ($\mu = 0$, $\delta = 2.5 \times \text{maximum experimental velocity}$) and five replicates for each point were made.

The results appear in Table 2. It can be observed that the probability 1 for the correct model was obtained after the 33th determination. Refinement was then carried out. However, after the 33th point the improvement was almost insignificant (Fig. 1b). The last parameter estimates were $V = 1.13 \mu\text{mol/min}$, $K_a = 29.5 \mu\text{M}$, $K_b = 98.8 \mu\text{M}$ and $K_{ab} = 0.418 \text{ nM}$.

2:2 one-substrate mechanism

The possible mechanisms considered with equal probability *a priori* were: 1:1 one-substrate, 1:2 one-substrate, 2:2 one-substrate and 2:3 one-substrate. These mechanisms are defined by the formula:

$$v_j = \frac{\sum_{i=0}^n a_i \cdot S_j^i}{\sum_{i=0}^m b_i \cdot S_j^i} \quad (16)$$

The preliminary experiment was of ten points geometrically distributed between 1.00 mM and 5.00 μM . The allowed experimental concentration of substrate ranged between 1.00 mM and 5.00 μM . The exact parameters values were $a_1 = 3.30 \times 10^4 \mu\text{mol/min}$,

$a_2 = 1.14 \times 10^7 \mu\text{mol} \cdot \text{min}^{-1} \cdot \text{M}^{-1}$, $b_2 = 1.21 \times 10^8 \text{ M}^{-1}$, $b_1 = 1.20 \times 10^4$ and $b_0 = 1 \text{ M}$. The exact velocities from the equation were randomly perturbed ($\mu = 0$, $\delta = 5.0 \times \text{maximum experimental velocity}$) and five replicates for each point were made.

The results appear in Table 3. It can be observed that the probability 1 for the correct model was obtained after the 27th determination. Refinement was then carried out. After this point the increment in accuracy was hardly significant (Fig. 1c). The last parameter estimates were $a_1 = 3.28 \times 10^4 \mu\text{mol/min}$, $a_2 = 1.31 \times 10^7 \mu\text{mol} \cdot \text{min}^{-1} \cdot \text{M}^{-1}$, $b_2 = 1.23 \times 10^8 \text{ M}^{-1}$, $b_1 = 1.16 \times 10^4$ and $b_0 = 1.00 \text{ M}$.

DISCUSSION

With the availability of computers the results of the kinetic experiment are generally presented with confidence intervals. The necessity for evaluating these intervals has demonstrated that many experiments do not have the necessary requirements of a good fit. This fact has led to planning experiments with suitable designs to achieve these goals.

The problem of designing experiments with the aim of reducing the intervals of confidence is now well established and several programs are available (Metzler *et al.*, 1974; Duggleby, 1979; Canela, 1985). However, in enzyme kinetics, the problem of discriminating among several rival models is limited to discrimination between two models and applied to unweighted cases (Reich, 1970; Bartfai & Mannervik, 1972; Mannervik, 1981, 1982). Classical plots do not allow one to choose among typical models when non-Michaelian enzymes are studied, and therefore the selection depends upon the

Scheme 1. Proposed procedure for model discrimination and parameter estimation in enzyme kinetics

results of the regression analysis. Furthermore, the enzyme equations are non-linear in the parameters and so the use of these tests, for example the *F* test, would only be justified in problems where the conditions were appropriate, which implies that much caution should be observed (Burguillo *et al.*, 1983).

The above reasons demand an efficient tool for the investigation of enzyme mechanisms. In order to formulate a design is necessary to know the structure of the error. A preliminary experiment with many replicates for each point and covering the range of responses allows one to obtain it.

The next step is to carry out the experiment, which in an initial stage should cover the range of responses having the substrate concentrations geometrically distributed. A number of five replicates of each point seems suitable (Duggleby, 1979; Mannervik, 1981, 1982). Then the results should be analysed by non-linear-regression fitting of the data on the possible models. At this point, the choice of the model is the most important question. In enzyme kinetics it often happens that the number of observations is not great and/or the minimum of the sum of squares is badly defined. In these cases the classical procedure is inadvisable and it is better to use the procedure proposed by Fedorov (1972), which uses the calculation of a probability *a posteriori* function in the range of the parameters. By this method, a new point is selected in order to increase the probability of one among the set of possible models and to improve the accuracy of the parameters. From the expression (14) it can be deduced that, when increasing the probability of a model, the weight of the first term of this expression increases.

However, it should be noted that when unsuitable concentrations of substrates are used (by, for example, experimental constrictions, or other limitations), the problem cannot be solved, and it is possible only to know the probability of each model as a mathematical concept, although this would hardly be significant in the enzyme field.

CONCLUSIONS

The procedure suggested in this paper is summarized in Scheme 1, and for its application a FORTRAN77 program has been devised.

This research was partially supported by a Research Grant 0657-81 from Comisión Asesora de Investigación Científica y Técnica.

REFERENCES

- Bardsley, W. G. (1976) *Biochem. J.* **153**, 101–117
- Bardsley, W. G. (1977a) *J. Theor. Biol.* **65**, 281–316
- Bardsley, W. G. (1977b) *J. Theor. Biol.* **67**, 121–139
- Bardsley, W. G. & Childs, R. E. (1975) *Biochem. J.* **149**, 313–328
- Bardsley, W. G. & Crabbe, M. J. C. (1976) *Eur. J. Biochem.* **68**, 611–619
- Bardsley, W. G., Leff, P., Kavanagh, J. & Waight, R. D. (1980a) *Biochem. J.* **187**, 739–765
- Bardsley, W. G., Woolfson, R. & Mazat, J. P. (1980b) *J. Theor. Biol.* **85**, 247–284
- Bartfai, T. & Mannervik, B. (1972) *FEBS Lett.* **26**, 252–256
- Box, G. E. P. & Lucas, H. L. (1959) *Biometrika* **46**, 77–90
- Burguillo, F. J., Wright, A. J. & Bardsley, W. G. (1983) *Biochem. J.* **211**, 23–24
- Canela, E. I. (1984) *Int. J. Bio-Med. Comput.* **15**, 121–130
- Canela, E. I. (1985) *Int. J. Bio-Med. Comput.* **16**, 257–266
- Childs, R. E. & Bardsley, W. G. (1976) *J. Theor. Biol.* **63**, 1–18
- Currie, D. J. (1982) *Biometrics* **38**, 907–919
- Del Arco, A., Burguillo, F. J. & Bardsley, W. G. (1984) *Int. J. Biochem.* **16**, 189–194
- Draper, N. R. & Smith, H. (1981) *Applied Regression Analysis*, 2nd edn., pp. 458–629, John Wiley and Sons, New York
- Duggleby, R. G. (1979) *J. Theor. Biol.* **81**, 671–684
- Duggleby, R. G. (1981) in *Design and Analysis of Enzyme and Pharmacokinetics Experiments* (Endrenyi, L., ed.), pp. 169–179, Plenum Press, New York
- Endrenyi, L. (1981) in *Design and Analysis of Enzyme and Pharmacokinetics Experiments* (Endrenyi, L., ed.), pp. 137–167, Plenum Press, New York
- Fedorov, V. V. (1972) *Theory of Optimal Experiments*, Academic Press, New York
- Hill, C. M., Waight, R. D. & Bardsley, W. G. (1977) *Mol. Cell. Biochem.* **15**, 173–178
- Mannervik, B. (1981) in *Design and Analysis of Enzyme and Pharmacokinetics Experiments* (Endrenyi, L., ed.), pp. 235–270, Plenum Press, New York
- Mannervik, B. (1982) *Methods Enzymol.* **87**, 370–390
- Marquardt, D. W. (1963) *J. Soc. Ind. Appl. Math.* **11**, 431–441
- Metzler, C. M., Elfring, G. L. & McEwen, A. J. (1974) *A User Manual for NONLIN and Associated Programs*, Upjohn Co., Kalamazoo
- Rao, M. S. & Iyengar, S. S. (1984) in *Computer Modeling of Complex Model systems* (Iyengar, S. S. ed.), pp. 29–53, CRC Press, Boca Raton
- Reich, J. G. (1970) *FEBS Lett.* **9**, 245–251
- Solano-Muñoz, F., Bardsley, W. G. & Indge, K. J. (1981) *Biochem. J.* **195**, 589–601
- Watts, D. G. (1981) in *Design and Analysis of Enzyme and Pharmacokinetics Experiments* (Endrenyi, L., ed.), pp. 1–24, Plenum Press, New York
- Whitehead, E. P. (1976) *Eur. J. Biochem.* **59**, 449–456
- Whitehead, E. P. (1978) *Biochem. J.* **171**, 501–504
- Whitehead, E. P. (1979) *J. Theor. Biol.* **80**, 355–381