Use of the *F* test for determining the degree of enzyme-kinetic and ligand-binding data

A Monte Carlo simulation study

Francisco J. BURGUILLO, Andrew J. WRIGHT and William G. BARDSLEY
*Department of Obstetrics and Gynaecology, University of Manchester, St. Mary's Hospital,
Manchester M13 0JH, U.K.*

(Received 26 July 1982/Accepted 26 November 1982)

1. Initial-rate data were simulated for 13 representative enzyme mechanisms with the use of several distributions of rate constants in order to locate conditions leading to $v([S])$ curves in physiological ranges of substrate concentration. 2. In all, 420 sets of such $v([S])$ curves were generated with the use of several choices of substrate concentration range (two, three or four orders of magnitude), number of experimental points (10, 15 or 20), error on v (5–10%) and standard deviation on v (5–9%) in order to simulate experimental results in a number of possible ways. 3. Curve-fitting was carried out to rational functions of degree 1:1, 2:2, ..., 5:5 until there was no statistically significant decrease in the sum of weighted squared residuals as judged by the *F* test at 95% and 99% confidence levels. 4. It was checked whether the non-linear regression program had located a good minimum in the sum of squares by also fitting the data with the correct values of parameters as starting estimates. 5. A similar procedure was adopted with 110 sets of binding data simulated for 11 models, and the *F* test was used to see if fractional-saturation data generated by a binding polynomial of order n could be adequately fitted by one of order m , $m < n$. 6. From the 530 simulations the *F* test was successful in fixing the correct degree with a probability of 0.62 at the 95% confidence level, but this fell with increase in degree as follows: 1:1 (0.98), 2:2 (0.71), 3:3 (0.43) and 4:4 (0.34), the first numbers being the degree of the rate equation and those in parentheses referring to the 95% confidence level. 7. It made little difference whether the 95% or the 99% confidence level was consulted, as there were very few borderline cases. 8. The chance of detecting deviations from Michaelis–Menten kinetics, i.e. terms of at least second-order in a rate equation of degree $n:n$, $n > 1$, was estimated to be about 0.8. 9. The probability of the *F* test leading to a spurious result due to error in the data was found to be about 0.04. 10. The probability with which 4:4 mechanisms can lead to $v([S])$ plots with no, one, two or three turning points was computed, and it was established that there is a small but finite chance that the increase in degree that occurs in some mechanisms when $ES \rightleftharpoons EP$ interconversions are explicitly allowed for can be detected by the *F* test.

We have suggested that steady-state kinetic and equilibrium binding data do not usually conform to the Michaelis–Menten equation, but can be fitted satisfactorily by rational functions of degree 2:2, 3:3 or 4:4 (Bardsley & Childs, 1975; Bardsley, 1976; Hill *et al.*, 1977; Bardsley *et al.*, 1980a). We have further computed the probability that several enzyme mechanisms can lead to the various types of non-linear double-reciprocal, Scatchard or Hanes plots that are possible with 2:2 (Solano-Muñoz *et al.*, 1981a,b) or 3:3 functions (Wardell *et al.*,

1982). However, these simulation studies have left unanswered the extremely important question of the probability that such deviations from Michaelis–Menten kinetics could be recognized. Stated formally: what is the chance that rate data generated by possible enzyme mechanisms will be sufficiently distinct from 1:1 data to be recognized as such, and what are the chances that an experimentalist could then determine the degree of rate equations objectively? It is perfectly obvious that sometimes statistical tests will be successful and sometimes they

will fail, and this can be for numerous reasons. In the present paper we attempt to get a more definitive assessment of the probability of success than 'sometimes' in the special circumstances of a determined investigator working with a favourable experimental situation.

There is only one way known at present to make such an estimate, and that is to use a computer to generate rate constants according to some specified distribution, calculate $v([S])$ data, add random errors to simulate reality, fit the resulting artificial data by using a minimization technique and then apply statistical tests. However, all of these procedures are problematical and controversial, so, before we can proceed to present our results, it is necessary to justify systematically all the steps that we have adopted.

The statistical analysis of enzyme-kinetic data has attracted a lot of attention (Wilkinson, 1961; Cleland, 1967; Haarhof, 1969; Reich, 1970; Otaway & Apps, 1972; Mannervik, 1975; Lam *et al.*, 1977; Feraudi *et al.*, 1977a,b; Garfinkel *et al.*, 1977; Matthews & Alben, 1978; Endrenyi, 1981). However, the subject developed in the present paper complements these studies by investigating in more detail an area first explored in an important pioneering paper by Pettersson & Pettersson (1970). Their approach assumes that the concentration of only one substrate is varied at a time, that the enzyme does not aggregate and that $[S] \gg [E_0]$, when the rate equation takes the form:

$$v/[E_0] = \sum_{i=1}^n \alpha_i [S]^i / \sum_{i=0}^n \beta_i [S]^i \quad (1)$$

$$= \sum_{i=1}^n \alpha_i [S]^i / \left(1 + \sum_{i=1}^n b_i [S]^i \right) \quad (2)$$

Pettersson & Pettersson (1970) argued that curve-fitting could be done systematically to degree 1:1, 2:2, 3:3 then 4:4 etc., with checking at each stage by the *F* test whether the resulting decrease in the sum of squares justified the increase in the number of parameters. Then a *t* test could be employed to assess the significance of estimated parameters. Although this sort of approach is the standard technique in most other branches of physical science, and is, quite rightly, finding increased use in enzyme kinetic studies, no one has yet estimated the value of such an approach. For instance, can the degree be overestimated? How does the success of the test depend on the number of points, orders of magnitude of the substrate concentration range or error structure of the data? Should the weighted or unweighted sum of squares be used? Can degree 4:4 ever be distinguished from 3:3?

Let us first make clear some ways in which the test could fail. First, it is possible for all the

resultants of numerator and denominator to vanish identically, in which case common factors cancel between numerator and denominator of eqn. (1) (Bardsley, 1977a; Whitehead, 1979), leading to the theoretical parent hyperbola:

$$\text{TPH} = \frac{\alpha_1 \alpha_n [S]}{\alpha_n \beta_0 + \alpha_1 \beta_n [S]}$$

(Bardsley, 1977b; Waight & Bardsley, 1977). We regard this as rather unlikely, for reasons outlined elsewhere (Solano-Muñoz *et al.*, 1981a,b), but at low concentrations eqn. (1) becomes the osculating hyperbola at zero (Bardsley & Waight, 1978a) defined by:

$$\text{OH}(0) = \frac{\alpha_1^2 [S]}{\alpha_1 \beta_0 - (\alpha_2 \beta_0 - \alpha_1 \beta_1) [S]}$$

whereas at high substrate concentrations it approaches:

$$\text{OH}(\infty) = \frac{\alpha_n^2 [S]}{(\alpha_n \beta_{n-1} - \alpha_{n-1} \beta_n) + \alpha_n \beta_n [S]}$$

and, since these are the linear regions of double-reciprocal, Scatchard or Hanes plots, it is obvious that obtaining steady-state data at only low or only high substrate concentrations will always lead to a failure to differentiate $n:n$ from 1:1 rate equations. Even if there are no common factors between numerator and denominator and the data can be obtained over a very wide concentration range, there still remains the chance that a particular set of rate constants have evolved for an enzyme so that, even though the rate equation be $n:n$, no statistical test could ever detect it. It is the probability of this happening that we have tried to estimate in the present paper, and we have not considered the failure of statistical tests that must be expected if fewer than ten points are collected either by design or default over a range of less than two orders of magnitude in the substrate concentration.

Theoretical

Kinetic and binding models simulated

The enzyme-kinetic and ligand-binding models used in this simulation are shown as nodal diagrams in Fig. 1. Mechanisms (i)–(iv) are sequential models in which an enzyme with n sites adds substrate (nk_{+1}) to give a species ES_1 , this being identical whichever site is occupied. ES_1 can then add substrate $[(n-1)k_{+2}]$, giving ES_2 , or release substrate (k_{-1}) or product (k_{p1}), to give back E. Finally, ES_{n-1} adds substrate (k_{+n}) producing ES_n , and this can release substrate (nk_{-n}) or product (nk_{pn}) to give ES_{n-1} . In this way a family of models of degree $n:n$ can be built up, but the models are very restrictive in

that all active sites behave identically at each stage of substrate occupancy. Model (v) is the one-site/two-state model, (vi) is the implicit modifier mechanism and (vii) is the Random Bi Bi model; all of these are 2:2, and the rate equations have been published (Bardsley *et al.*, 1980a; Solano-Muñoz *et al.*, 1981a,b). Mechanisms (viii)–(xi) are 3:3 models that have been simulated before (Wardell *et al.*, 1982). They are: (viii) the implicit homosubstrate modifier (Bardsley *et al.*, 1980a), (ix) the two-site implicit homosubstrate modifier, (x) alkaline phosphatase (Waight *et al.*, 1977) and (xi) peroxidase ([B] varied) (Childs & Bardsley, 1975). Two further 4:4 mechanisms are: (xii) the peroxidase ([A] varied) and (xiii) explicit version of the homosubstrate modifier mechanism (Bardsley *et al.*, 1980a). Binding models (xiv)–(xvii) have the binding polynomial:

$$N = 1 + \binom{n}{1} K_1 x + \binom{n}{2} K_1 K_2 x^2 + \dots + K_1 K_2 \dots K_n x^n$$

(Wyman, 1972) and the saturation function is simply:

$$y = \frac{1}{n} \cdot \frac{d(\ln N)}{d(\ln x)} \quad (3)$$

whereas (xviii)–(xxi) are Monod–Wyman–Changeux models of degree n (Monod *et al.*, 1965). For these we used the binding polynomial:

$$N = [\lambda_1(1 + K_1 x)^n + \lambda_2(1 + K_2 x)^n] / (\lambda_1 + \lambda_2)$$

where $\lambda_1 + \lambda_2 = 1$, and again:

$$y = \frac{1}{n} \cdot \frac{d(\ln N)}{d(\ln x)}$$

Another binding model studied involves groups of n_i independent identical sites, each with binding constant K_i . This gives the binding polynomial:

$$N = (1 + K_1 x)^{n_1} (1 + K_2 x)^{n_2} \dots (1 + K_q x)^{n_q}$$

and the saturation function:

$$y = \frac{1}{\sum_{i=1}^q n_i} \cdot \sum_{i=1}^q \frac{n_i K_i x}{1 + K_i x}$$

Note that, whereas the Monod–Wyman–Changeux model gives only positively co-operative binding, this last model gives negatively co-operative binding, the binding isotherms and double-reciprocal and Scatchard plots being without inflexions (Bardsley & Waight, 1978b). In mechanisms (xxii), (xxiii) and (xxiv) the summation was over two, three or four sites respectively with $n_i = 1$.

Choice of rate and binding constants

There is obviously some statistical distribution of rate constant values in the totality of enzymes, since

extreme values for k_i of magnitude 10^{-10} s^{-1} or 10^{10} s^{-1} are less likely to occur than more modest values around 10^2 s^{-1} , say. Equally, there is no conceivable way of estimating the density (i.e. probability density function) of this distribution of k values, say $f(k)$, and in any case every enzyme has an individual set of constants evolved to fit a certain physiological function. So a particular enzyme could have a set of statistically unlikely rate constants that would not be representative of $f(k)$. At best we can only hope to model potential situations by choosing a large number of examples with rate constants that, broadly speaking, are more likely to have chemically reasonable values. For reasons outlined elsewhere (Wardell *et al.*, 1982), we have fixed on a lognormal distribution with second-order constants on average about three to four orders of magnitude larger than first-order ones. In the examples simulated we discarded about 10% of the examples because of 'premature saturation' and another 10% because of 'slow saturation', since we were unwilling to compromise the generality of our approach by using too narrow a range of k_i values. Details concerning the choice of rate and binding constants are given in Table 1.

Generation of simulated data

We wrote a program in FORTRAN that accessed subroutines for each of the models. The program accepted input for the type of model and distribution of k_i or K_i and then printed out $v([S])$ or $y(x)$ points over a wide range of substrate or ligand concentration. If the particular choice of k_i or K_i led to a satisfactory curve in the physiological range, the program then called for a further input of number of points, types of error associated with the points and range of independent variable. K_i or k_i values were then printed out, along with coefficients α_i and β_i and all the Sylvester resultants. Also, a data file was saved with the dependent variable, independent variable and standard deviation associated with the estimate in a form ready for curve-fitting. The facility was also available for input of k_i values in order to generate data with different error structure but with the same coefficients. The rate equations were calculated by a PASCAL program, provided by K. Indge (Indge & Childs, 1976), that calculated coefficients α_i and β_i and then converted them into a form ready for simulation. The type of data files generated were chosen so as to mimic as closely as possible the type of results found by ourselves and others (Askelof *et al.*, 1976; Storer *et al.*, 1975; Cornish-Bowden & Endrenyi, 1981), and details are given in Table 1.

Details of curve-fitting

Simulated kinetic data were initially fitted to eqn. (1) to get an estimate of the relative magnitude of the

Fig. 1. Enzyme mechanisms simulated in the present paper

Only the nodal diagrams are shown, rate constants being omitted for clarity. Product-release steps are identified by a full arrow head, reversible steps have single-headed arrows. The mechanisms are implicit, i.e. ES represents $ES=EP$ except for (xiii). (i), (ii), (iii) and (iv) A family of sequential models in which each state of substrate occupancy, i.e. E, ES_1 , ES_2 etc., is identical whichever of the n sites is filled. The species ES_1 , ES_2 , ..., ES_n , however, can have different kinetic properties. (v) A one-site/two-state model with the two conformational forms R and T. (vi) The modifier mechanism in which addition of modifier, M, leads to altered kinetic properties. (vii) The Random Bi Bi model with substrates A and B and products P and Q. (viii) The homosubstrate modifier mechanism in which substrate binds to a non-productive modifier site, leading to altered kinetic properties. (ix) A model with two distinct active sites. (x) A model proposed for alkaline phosphatase. (xi) and (xii) A model proposed for peroxidase. (xiii) The explicit version of the homosubstrate modifier mechanism.

coefficients α_i and β_i . There is no general rule regarding the relative magnitude of these coefficients, and the ratio of highest to lowest can often be very large, e.g. 10^{20} or greater, depending on the mechanism and choice of distribution for k_i values. Normalizing to eqn. (2) gave new starting estimates for the coefficients a_i and b_i and refitting the data. This procedure did not result in any further decrease in the weighted sum of squares for 1:1 and 2:2 curves, and only occasionally did it result in slightly improved fit for 3:3 and 4:4 functions. Several sets of starting estimates were used. One set used the same starting point for all coefficients, and another

had $\alpha_{i+1} = 10^3 \alpha_i$ and $\beta_{i+1} = 10^3 \beta_i$ for starting estimates. To check that the lowest weighted sum of squares had been located, the data were also fitted, by using eqn. (2), with the known, exact, values of coefficients as starting estimates. This procedure is obviously impossible in real-life experiments, and it was only used as a check that optimum fit had been achieved. The program was vindicated, in that in no case was the fit substantially better than that obtained by using a blind start. Simulated binding data were fitted directly to the K_i values of eqn. (3).

In each case 10^4 random searches of parameter space were made, followed by a Simplex and then

Table 1. *Details of data stimulation*

For (i) the choice of parameters, the general details are as follows. Rate constants, k_i , were all chosen to obey detailed balancing, all possible cycles of microscopic reversibility being exploited in a uniformly random manner. Parameters, p , followed a lognormal distribution, i.e. $\log p \sim N(\mu, \sigma^2)$, and values outside fixed ranges of $\pm 2\sigma$ were discarded. For (ii) the choice of data spacing and error, the general details are as follows. The exact v_i and y_i values were converted into simulated experimental values, $v_{\text{exp.}}$ and $y_{\text{exp.}}$, by adding normally distributed random error, and to each point was associated a random standard deviation obtained from a normal distribution. Values of the dependent variable or error falling outside the allowed range of $\pm 2\sigma$ were given the extreme value of the range. Experimental points were spaced logarithmically.

(i) Choice of parameters

Parameter	μ	σ	Range
2nd-order k_i for S addition (a)	5	2	$10^1 \leq k_i \leq 10^9 \text{ M}^{-1} \cdot \text{s}^{-1}$
(b)	6	2	$10^2 \leq k_i \leq 10^{10} \text{ M}^{-1} \cdot \text{s}^{-1}$
2nd-order k_i for M addition	2	2	$10^{-2} \leq k_i \leq 10^6 \text{ M}^{-1} \cdot \text{s}^{-1}$
1st-order k_i for chemical steps	2	2	$10^{-2} \leq k_i \leq 10^6 \text{ s}^{-1}$
1st-order k_i for conformational change	1	2	$10^{-3} \leq k_i \leq 10^5 \text{ s}^{-1}$
Concentration of second substrate B	-3	1	$[B] \leq 1 \text{ M}$
Concentration of modifier M	-3	1	$[M] \leq 1 \text{ M}$
Binding constant K_i	4	2	$1 \leq K_i \leq 10^8 \text{ M}^{-1}$
Conformational constant	0	2	$10^{-4} \leq \lambda_1/\lambda_2 \leq 10^4$

(ii) Choice of data spacing and error

Data type	No. of points	Spread of points	Error in dependent variable		Range of dependent variable	Standard deviation in dependent variable		Range of standard deviation in dependent variable
			μ	σ		μ	σ	
K(i)	15	$5 \times 10^{-6} \text{ M} \leq [S] \leq 5 \times 10^{-3} \text{ M}$	0	2.5%	$0.95 v_i \leq v_{\text{exp.}} \leq 1.05 v_i$	5%	0	$0.05 v_i = \sigma_i$
K(ii)	15	$5 \times 10^{-6} \text{ M} \leq [S] \leq 5 \times 10^{-3} \text{ M}$	0	2.5%	$0.95 v_i \leq v_{\text{exp.}} \leq 1.05 v_i$	5%	1.25%	$0.025 v_i \leq \sigma_i \leq 0.075 v_i$
K(iii)	15	$5 \times 10^{-6} \text{ M} \leq [S] \leq 5 \times 10^{-3} \text{ M}$	0	5%	$0.9 v_i \leq v_{\text{exp.}} \leq 1.1 v_i$	9%	3%	$0.03 v_i \leq \sigma_i \leq 0.15 v_i$
K(iv)	20	$5 \times 10^{-6} \text{ M} \leq [S] \leq 5 \times 10^{-2} \text{ M}$	0	5%	$0.9 v_i \leq v_{\text{exp.}} \leq 1.1 v_i$	9%	3%	$0.03 v_i \leq \sigma_i \leq 0.15 v_i$
K(v)	10	$10^{-5} \text{ M} \leq [S] \leq 10^{-3} \text{ M}$	0	2.5%	$0.95 v_i \leq v_{\text{exp.}} \leq 1.05 v_i$	5%	1.25%	$0.025 v_i \leq \sigma_i \leq 0.075 v_i$
B	15	$10^{-6} \text{ M} \leq x_1 \leq [x] \leq x_2 \leq 10^{-1} \text{ M}$ $x_2 = 10^3 x_1, y(x_1) \leq 0.1$ and $y(x_2) \geq 0.9$	0	2.5%	$0.95 y_i \leq y_{\text{exp.}} \leq 1.05 y_i$	5%	1.25%	$0.025 y_i \leq \sigma_i \leq 0.075 y_i$

a Migrad routine, and the print-out included best-fit parameters, estimated errors, number of calls to function and weighted sum of squares at the end of each step. Convergence criteria, the estimated variance-co-variance matrix, final weighted and unweighted sum of squares and calculated and experimental values were given, along with information as to whether parameter limits were approached. New starting estimates were input as required, and this was done interactively until no further improvement in fit occurred.

Use of the F test

We wrote a program that accepted inputs of Q_1 and Q_2 , the lowest weighted sums of squares for regression to models 1 and 2 with m_1 and m_2 parameters respectively. The expression used to calculate Q_j was:

$$Q_j = \sum_{i=1}^n q_{ij}^2 / \sigma_{ij}^2$$

where q_{ij} is the residual and σ_{ij}^2 is the variance. The number of experimental points was n , and the program calculated the likelihood-ratio-test statistic F (Lindgren, 1976) given by:

$$F(m_2 - m_1, n - m_2) = \frac{(Q_1 - Q_2) / (m_2 - m_1)}{Q_2 / (n - m_2)}$$

and compared the value with 95 and 99 percentiles of the appropriate F distribution. The results for this procedure are summarized in Table 2.

We also calculated an F statistic by using the formula:

$$\frac{(n - m_1)Q_1 - (n - m_2)Q_2}{(m_2 - m_1)Q_2} = F(m_2 - m_1, n - m_2) + \frac{Q_1}{Q_2}$$

given by Pettersson & Pettersson (1970), but the results were not sufficiently different to warrant inclusion.

4:4 curve-shape features

We naturally expected some difficulty in detecting fourth-degree terms in 4:4 data. In binding models we are trying to separate four from three parameters, but in kinetic models we have to detect eight rather than six parameters, and this will only be possible when the 4:4 data have some distinctive high-degree feature. So far, the 4:4 curves have not been classified, but it is known that three turning points are possible (Bardsley *et al.*, 1980b) and can be given by kinetic schemes [e.g. Fig. 4(a) of Bardsley *et al.* (1980a), noting that k_{-6} was printed as 1 rather than 0.001]. Accordingly, we wrote a program that generated coefficients, located the roots of $dv/d[S] = 0$, and tested the magnitude of these roots. In this way the probability of obtaining one distinctive 4:4 feature, three positive turning points, was assessed, and the chance of this falling in the experimentally accessible range was computed. A number of examples of mechanism (xiii) with three turning points in physiological ranges of substrate concentrations were selected for curve-fitting, but the F -test results for this are not included in the data in Table 2.

Results

Table 1 summarizes the choice of k_i and K_i values used in the kinetic and binding models, and gives

Fig. 2. Success in detecting high-degree terms by use of the F test

The overall average proportions of successes from the results of Table 2 are plotted for degree 1:1, 2:2, 3:3 and 4:4. The probability of detecting the correct degree (at 95% confidence) falls off rapidly as the degree increases, but the probability of detecting terms of at least second degree, i.e. deviations from Michaelis–Menten kinetics, increases slightly with increase in degree. O, Fixing the correct degree for kinetic models; Δ, fixing the correct degree for binding models; □, detecting second-order terms for kinetic models; ●, detecting second-order terms for binding models.

details of the types of random error introduced into the calculated data to simulate reality.

Table 2 gives the values obtained by curve-fitting and application of the F test. The first column is the number of the model named in the second column. Then come the fittings for 5:5/4:4, 4:4/3:3 etc. in which the proportion at 95% confidence level is given (99% in parentheses). The last column gives the proportion of successes after allowing for false positives.

Table 3 gives the proportion of simulated curves giving no, one, two or three turning points for 4:4 kinetic mechanisms. Numbers in parentheses are for curves with these features in the experimentally accessible range.

Fig. 1 shows the nodal diagrams to allow easy identification of the models.

Fig. 2 illustrates the success in using the F test for determining correct degree and for detecting deviation from Michaelis–Menten kinetics.

Discussion

The simulation study that we have undertaken is likely to prompt at least six searching questions to the critical reader, and we must state and discuss these.

Question 1. How close to reality was the simulation study?

Question 2. Were there any substantial differences between alternative types of data structure?

Question 3. Did different models of the same degree behave in a similar fashion?

Question 4. What proportion of false positives are to be expected when using the F test?

Question 5. How well did the minimization procedure estimate parameters?

Question 6. What can we conclude about the use of the F test for detecting the true degree?

Dealing first of all with question 1, we emphasize that 13 kinetic and 11 binding models were simulated, a total of 530 sets of data. It is unlikely that further simulations would lead to any substantial change in our conclusions. All the curves fitted lay in the experimentally accessible range, and k_i and K_i values were allowed to vary over many orders of magnitude. We also experimented with several wide ranges of parameters, but, since this led to an increase in the number of rejections due to simulated data lying outside the physiological range, our opinion is that the curves used were representative of those generated by actual enzymes. To answer the content of question 2, we mention that 10, 15 or 20 data points were generated over two, three or four orders of magnitude by using errors on the $v([S])$ or $y(x)$ points of up to 10% and allowing the standard deviation of each point to be either a fixed percentage of the dependent variable or to be

Table 2. *F*-test results

Details of the models are in Fig. 1, and the types of data, i.e. $K(i)$, $K(ii)$, $K(iii)$, $K(iv)$, $K(v)$, (a) , (b) or B , are summarized in Table 1. The curve-fitting procedure and use of the *F* test are described in the text. Sometimes the same set of coefficients was used to generate data with alternative error structures. When this is done, the entry is labelled 1st set. Figures entered under the title 2nd set are for a completely new set of rate constants. Abbreviation: MWC, Monod-Wyman-Changaux model.

Model (see Fig. 1)	Degree	Data type (see Table 1)	Proportion of <i>F</i> -test successes for $n:n$ as opposed to $(n-1):(n-1)$ rational functions at 95% (99%) confidence level in each sample, i.e. set of ten examples				Proportion of correct assignment of degree in each sample, i.e. set of ten examples at 95% (99%) confidence level after allowing for false positives
			5:5/4:4	4:4/3:3	3:3/2:2	2:2/1:1	
(i) Sequential 1:1	1:1	$K(i)(b)$	—	—	0 (0)	0 (0)	1 (1)
		$K(ii)(b)$	—	—	0 (0)	0.1 (0)	0.9 (1)
		$K(iii)(b)$	—	—	0 (0)	0.1 (0)	0.9 (1)
		$K(iv)(b)$	—	—	0.1 (0)	0 (0)	1 (1)
		$K(v)(b)$	—	—	—	0 (0)	1 (1)
		$K(ii)(b)$ 2nd set	—	—	0 (0)	0.1 (0)	0.9 (1)
(ii) Sequential 2:2	2:2	$K(i)(b)$	—	0 (0)	0 (0)	1 (1)	1 (1)
		$K(ii)(b)$	—	0 (0)	0 (0)	1 (1)	1 (1)
		$K(iii)(b)$	—	0 (0)	0 (0)	1 (1)	1 (1)
		$K(iv)(b)$	—	0 (0)	0 (0)	1 (1)	1 (1)
		$K(v)(b)$	—	—	0 (0)	0.9 (0.9)	0.9 (0.9)
		$K(ii)(b)$ 2nd set	—	0 (0)	0.1 (0)	0.9 (0.8)	0.9 (0.8)
(iii) Sequential 3:3	3:3	$K(i)(b)$	—	0 (0)	0.5 (0.5)	0.9 (0.9)	0.5 (0.5)
		$K(ii)(b)$	—	0 (0)	0.5 (0.5)	0.9 (0.9)	0.5 (0.5)
		$K(iii)(b)$	—	0 (0)	0.4 (0.4)	0.9 (0.9)	0.4 (0.4)
		$K(iv)(b)$	—	0 (0)	0.8 (0.8)	1 (0.9)	0.8 (0.8)
		$K(v)(b)$	—	0 (0)	0.3 (0.3)	0.8 (0.6)	0.3 (0.3)
		$K(ii)(b)$ 2nd set	—	0 (0)	0.6 (0.6)	1 (0.9)	0.6 (0.6)
(iv) Sequential 4:4	4:4	$K(i)(b)$	0.2 (0)	0.3 (0.3)	0.8 (0.8)	0.9 (0.9)	0.3 (0.3)
		$K(ii)(b)$	0 (0)	0.4 (0.3)	0.9 (0.9)	0.8 (0.8)	0.4 (0.3)
		$K(iii)(b)$	0 (0)	0.2 (0.2)	0.9 (0.8)	0.9 (0.9)	0.2 (0.2)
		$K(iv)(b)$	0.3 (0.2)	0.4 (0.3)	0.9 (0.9)	0.9 (0.7)	0.4 (0.3)
		$K(v)(b)$	—	0 (0)	0.7 (0.6)	0.9 (0.7)	0 (0)
		$K(iii)(b)$ 2nd set	0 (0)	0.4 (0.3)	0.5 (0.4)	0.9 (0.9)	0.4 (0.3)
(v) One-site/two state (MWC implicit)	2:2	$K(i)(a)$ 1st set	—	0 (0)	0 (0)	0.3 (0.3)	0.3 (0.3)
		$K(iii)(a)$ 2nd set	—	—	0.1 (0.1)	0.3 (0.2)	0.3 (0.2)
(vi) Modifier (implicit)	2:2	$K(i)(a)$ 1st set	—	0 (0)	0 (0)	0.1 (0.1)	0.1 (0.1)
		$K(iii)(a)$ 2nd set	—	—	0.1 (0)	0.2 (0.1)	0.1 (0.1)
(vii) Random Bi Bi	2:2	$K(i)(a)$ 1st set	—	0 (0)	0 (0)	0.2 (0.2)	0.2 (0.2)
		$K(iii)(a)$ 2nd set	—	0 (0)	0 (0)	0.3 (0.2)	0.3 (0.2)
(viii) Homosubstrate modifier (implicit)	3:3	$K(i)(a)$ 1st set	—	0 (0)	0.2 (0.1)	0.8 (0.8)	0.2 (0.1)
		$K(iii)(b)$ 2nd set	—	0 (0)	0 (0)	0.7 (0.7)	0 (0)
(ix) Two-site (implicit)	3:3	$K(i)(a)$ 1st set	—	0 (0)	0 (0)	0.8 (0.7)	0 (0)
		$K(iii)(b)$ 2nd set	—	0 (0)	0.1 (0.1)	0.6 (0.6)	0.1 (0.1)
(x) Alkaline phosphatase	3:3	$K(i)(a)$ 1st set	—	0 (0)	0.2 (0.2)	0.9 (0.8)	0.2 (0.2)
		$K(iii)(a)$ 2nd set	—	0.1 (0)	0.1 (0)	0.7 (0.7)	0.1 (0)
(xi) Peroxidase ([B] varied)	3:3	$K(i)(b)$ 1st set	—	0 (0)	0 (0)	0.4 (0.3)	0 (0)
		$K(iii)(b)$ 2nd set	—	0 (0)	0 (0)	0.2 (0.2)	0 (0)
(xii) Peroxidase ([A] varied)	4:4	$K(i)(b)$ 1st set	0 (0)	0 (0)	0 (0)	0.7 (0.7)	0 (0)
		$K(iii)(a)$ 2nd set	0.1 (0)	0 (0)	0.2 (0.1)	0.6 (0.5)	0 (0)
(xiii) Homosubstrate (explicit)	4:4	$K(i)(b)$ 1st set	0 (0)	0 (0)	0.4 (0.3)	0.9 (0.9)	0 (0)
		$K(iii)(b)$ 2nd set	0 (0)	0 (0)	0.1 (0.1)	0.8 (0.8)	0 (0)
(xiv) Generalized binding	1:1	B	—	—	—	0 (0)	1 (1)
(xv) Generalized binding	2:2	B	—	—	0.1 (0)	1 (0.9)	0.9 (0.9)
(xvi) Generalized binding	3:3	B	—	0 (0)	1 (0.9)	0.9 (0.8)	1 (0.9)
(xvii) Generalized binding	4:4	B	0 (0)	0.9 (0.9)	0.9 (0.9)	0.7 (0.7)	0.9 (0.9)
(xviii) MWC binding	1:1	B	—	—	—	0 (0)	1 (1)

Table 2 (continued)

Model (see Fig. 1)	Degree	Data type (see Table 1)	Proportion of <i>F</i> -test successes for <i>n</i> : <i>n</i> as opposed to (<i>n</i> −1):(<i>n</i> −1) rational functions at 95% (99%) confidence level in each sample, i.e. set of ten examples				Proportion of correct assignment of degree in each sample, i.e. set of ten examples at 95% (99%) confidence level after allowing for false positives
			5:5/4:4	4:4/3:3	3:3/2:2	2:2/1:1	
(xix) MWC binding	2:2	B	—	—	0.1 (0.1)	0.7 (0.6)	0.7 (0.6)
(xx) MWC binding	3:3	B	—	0 (0)	0.4 (0.3)	0.8 (0.8)	0.4 (0.3)
(xxi) MWC binding	4:4	B	0.1 (0.1)	0.3 (0.1)	0.7 (0.7)	0.9 (0.9)	0.2 (0.1)
(xxii) Two independent sites	2:2	B	—	—	0.1 (0.1)	1 (1)	0.9 (0.9)
(xxiii) Three independent sites	3:3	B	—	0 (0)	0.4 (0.3)	0.9 (0.9)	0.4 (0.3)
(xxiv) Four independent sites	4:4	B	0 (0)	0.4 (0.3)	0.7 (0.7)	1 (1)	0.4 (0.3)

Summary of curve-fitting results

Type of model	Degree of model	No. of models	No. of curves fitted	Proportion of successes at 95% level (99%)	Proportion of successes in detecting 2:2 > 1:1 at 95% level (99%)	Proportion of false positives at 95% level (99%)
Kinetic	1:1	1	60	0.95 ± 0.05 (1 ± 0)	—	0.07 ± 0.05 (0 ± 0)
Kinetic	2:2	4	120	0.59 ± 0.40 (0.57 ± 0.41)	0.60 ± 0.39 (0.57 ± 0.41)	0.03 ± 0.05 (0.008 ± 0.03)
Kinetic	3:3	5	140	0.26 ± 0.26 (0.25 ± 0.27)	0.76 ± 0.23 (0.71 ± 0.22)	0.007 ± 0.03 (0 ± 0)
Kinetic	4:4	3	100	0.17 ± 0.19 (0.14 ± 0.15)	0.83 ± 0.11 (0.78 ± 0.13)	0.07 ± 0.11 (0.02 ± 0.07)
Binding	1:1	2	20	1.0 ± 0 (1.0 ± 0)	—	0 ± 0 (0 ± 0)
Binding	2:2	3	30	0.83 ± 0.12 (0.80 ± 0.17)	0.83 ± 0.12 (0.80 ± 0.17)	0.1 ± 0.1 (0.07 ± 0.06)
Binding	3:3	3	30	0.60 ± 0.35 (0.50 ± 0.35)	0.87 ± 0.06 (0.83 ± 0.06)	0 (0)
Binding	4:4	3	30	0.50 ± 0.36 (0.43 ± 0.42)	0.87 ± 0.15 (0.87 ± 0.15)	0.02 ± 0.04 (0.02 ± 0.04)

Total = 24 Total = 530

Average for 1:1 = 0.98 (1.0) Overall average = 0.79 (0.76)
 Average for 2:2 = 0.71 (0.69) Average for 2:2 = 0.72 (0.69)
 Average for 3:3 = 0.43 (0.38) Average for 3:3 = 0.82 (0.77)
 Average for 4:4 = 0.34 (0.29) Average for 4:4 = 0.85 (0.83)
 Overall average = 0.04 (0.01)

Table 3. Occurrence of complex *v*([S]) curves with 4:4 mechanisms

v([S]) plots for 4:4 functions can have up to three positive turning points, i.e. [S] values where $dv/d[S] = 0$ for [S] > 0. The probability for observing no, one, two or three turning points was computed by using a 4:4 rational function with arbitrary coefficients uniformly distributed on (0,1) and also mechanisms (iv), (xii), (xiii) with k_1 values as used elsewhere in the present paper. Values in parentheses are where all features occur in the range $5 \times 10^{-6} \text{ M} < [S] < 5 \times 10^{-3} \text{ M}$, i.e. the physiological range.

Mechanism	No. of cycles	Probability for 0, 1, 2 or 3 turning points (physiological range)			
		0	1	2	3
Rational functions	30 000	0.39	0.50 (0)	0.11 (0)	0.0029 (0)
(iv) Sequential	30 000	0.28	0.45 (0.23)	0.24 (0.067)	0.033 (0.0017)
(xii) Peroxidase ([A] varied)	70 000	0.38	0.60 (0.22)	0.019 (0.0018)	0.0034 (0.000029)
(xiii) Explicit homosubstrate mechanism	70 000	0.50	0.49 (0.28)	0.013 (0.0032)	0.0027 (0.00013)

random. We tried to cover all the contingencies that an experimentalist could encounter. Bearing in mind the findings of Askelof *et al.* (1976) and Storer *et al.* (1975) concerning error, and also relying on our own results, we consider that schemes *K*(i), *K*(ii) and *K*(iii) are probably the most representative of what can be realized in practice, and *F*-test results for schemes *K*(i), *K*(ii), *K*(iii) and *K*(iv) were very similar. In other words, the results were not strongly affected by changes of error structure in the ranges simulated by *K*(i)–*K*(iv) or *B*. The values given in

Table 1 show what is to be expected, namely that, as the number of points and range decreases, as in *K*(v), the proportion of *F*-test successes begins to fall off.

We are not able to provide a definitive answer to question 3, since only 20 curves fitted for each mechanism studied was not sufficient.

However, it does seem that, in very complex mechanisms, such as (xii) or (xiii), the coefficients in the rate equation, being such complicated functions of the rate constants, are somewhat buffered from

the effect of variation in the k_i values when compared with mechanisms such as (i)–(iv). Consequently, the coefficients for mechanism (iv) vary more than those of mechanisms (xii) or (xiii) as the k_i values vary, and so high-degree character is more easily detected. An extreme example of this is the binding models, where the probability of F -test success was generally higher than with kinetic models. However, a word of qualification is needed concerning this matter. In converting binding data into $y(x)$ form, it is necessary to know the number of sites per molecule, i.e. n . Hence what we have really been testing with the models (xiv)–(xxiv) is the extent to which data generated by a $n:n$ Adair equation can be equally well fitted by an Adair equation of lower degree. There are circumstances when binding data can be generated by an n th-degree Adair equation with some dependent parameters, and then it may be that an $m:m$ equation, $m < n$, gives an adequate fit, but it is not an m th-degree Adair equation, i.e. the numerator is not the logarithmic derivative of the denominator. We can, however, answer question 4, since the probability of false positives was found to be very small. This proportion did increase slightly when the statistic $F(m_2 - m_1, n - m_2) + Q_1/Q_2$ was used. There was little difference between results at 95% or 99% confidence with either statistic.

The K_i values of eqn. (3) were estimated very closely by our routine, as were the a_i/β_0 values of eqn. (1), i.e. a_i of eqn. (2). It was noted that in the high-degree curves sum-of-squares minima were sometimes found with coefficients very far removed from the actual ones, i.e. the error surface has local minima, but this was not found with 1:1 and 2:2 curves. We can answer question 6 by noting that the probability of detecting the true degree falls off steadily as the degree increases, as illustrated in Fig. 2. This is less marked with binding models, but perhaps attention should also be drawn to the fact that the probability of detecting terms of at least second degree was found to be quite high. Evidently it is much easier to prove that a rate equation has second-degree terms than that it has third-order or fourth-order coefficients. We might ask whether there could ever be objective evidence for degree 4:4 as opposed to 3:3, say.

We can now answer this question by reference to the results of Table 3. This Table shows what is fairly obvious but nevertheless very important. If we choose a set of k_i values for a 4:4 mechanism, then it is very likely that the curve will have only one or no turning points. The chances of obtaining a set of k_i values leading to a $v([S])$ curve with three turning points in the physiological range is very small. Clearly, before curve-fitting can differentiate degree 4:4 from 3:3, there must be some geometric feature of the $v([S])$ curve that has specific 4:4 character.

One such specific feature could be the three turning points just referred to. If the $v([S])$ curve were of this special variety and the geometric features were sufficiently pronounced to be detectable through the random noise of experimental error, then the F test would differentiate degree 4:4 from 3:3. An example to illustrate this is the explicit homomultisubstrate mechanism (xiii). The F test detected the true degree in an example with three turning points and established an important theoretical point. It has been shown (Solano-Muñoz *et al.*, 1981b) that, when the degree of a rate equation increases due to the inclusion of $ES \rightleftharpoons EP$ steps, the probability of getting a new curve-shape feature, though finite, is very small. We have now established that in such circumstances, an example being mechanisms (viii) and (xiii), the F test could detect an increase in degree and the explicit mechanism could be differentiated from the implicit one.

To avoid being misunderstood, we emphasize that it is not necessary for a curve to have multiple turning points or inflexions in order for the F test to detect high-degree terms. Other features such as steepness and curvature can also suffice.

We have claimed that the probability with which different 2:2 mechanisms give rise to the alternative types of $v([S])$ and double-reciprocal plots is quite characteristic of the individual mechanisms (Solano-Muñoz *et al.*, 1981a,b). Also, we showed that the 3:3 mechanisms are not so distinctive in the occurrence of this phenomenon (Wardell *et al.*, 1982). The 4:4 simulation results, although not so comprehensive, suggest that 4:4 mechanisms are all similar in the extent to which the alternative curve shapes can arise. Perhaps this is only to be expected, as the coefficients of the rational functions become increasingly complicated functions of the k_i , and variations in the contributions of the directed trees in the distribution expressions become buffered, so to speak. However, the finding that the homomultisubstrate modifier mechanism in explicit form (mechanism xiii) is just a typical 4:4 mechanism is a very powerful result. Most mechanisms with conformational states, several active or binding sites etc. lead to an increase in degree when specific steps are allowed for $ES \rightleftharpoons EP$. If the example just discussed is typical, it suggests that incredibly complicated $v([S])$ curves could be fitted by very simple mechanisms in explicit forms and so the possibility of model discrimination with rate equations of high degree seems very remote.

Finally, we confess that the probabilities computed in the present paper have to be viewed with caution, since they are from a limited sample and are subject to the choice of distribution for k_i , cut-off points for parameter magnitudes, error structure imposed and curve-fitting program. Nevertheless, such probabilities must exist since, if an exhaustive

analysis could be done, convergence to final values would occur asymptotically. For instance, we have claimed that the probability of detecting second-order terms in a rate equation of degree $\geq 2:2$ is of the order of 0.8 when simulations are made as closely as possible to reality. We would be surprised if the 'actual value', obtained in only real-life experiments, differed from this by much more than 10%. This argument is not dependent on the accuracy of the values in Tables 1 and 2, which can be predicted independently. For, if L 'successes' in N 'trials' are used to estimate a probability, $p \approx L/N$, by the Monte Carlo method, then L is a random binomial variable with mean Np and variance $Np(1-p)$.

We perhaps should mention that none of the resultants generated in the simulation were small, and there seemed to be no correlation between failure to detect high-degree terms and relatively small resultants. In other words, none of the high-degree curves could be thought of as modestly perturbed Michaelis–Menten equations. Also, none of the coefficients generated or calculated as best-fit coefficients could be discounted, i.e. there were no effectively dead-end substrate-inhibition examples.

Conclusion

Well-established methods of mathematical statistics can be used to justify the F test for model discrimination under the assumptions that there are a large number of determinations with no serious outliers, that the variance is constant and that the model is linear in the parameters. These assumptions do not hold in enzyme kinetics, and so extension of statistical techniques into an area where they are not strictly justified is based on the hope that perturbations such as non-linearities do not seriously affect the theoretical basis. The philosophy summed up by the saying 'half a loaf is better than no bread' has great appeal to a needy man.

It behoves us to proffer advice as to the choice of experimental conditions most conducive to success and to assess the probable chance of favourable outcome when analysing complex kinetics. In the present paper we have eschewed theoretical calculations and instead have sought to determine by an empirical approach the value of the F test. We envisage a situation where an experimenter has a stable enzyme/substrate or macromolecule/ligand system and the independent variable can be varied over two to four orders of magnitude, with some 10–20 points, each determined at least five times so that individual variances can be estimated. Putative outliers will have been eliminated or vindicated by redeterminations, and we shall suppose that the experimenter has grounds for suspecting substrate activation, substrate inhibition, several active sites,

homosubstrate- or modifier-binding sites, second substrates involved in random pathways, conformational states or other causes of co-operative effects. Such circumstances lead to $v([S])$ or $y(x)$ curves of high degree, and the experimenter would want to ask: what is the chance of correctly determining the degree of the $v([S])$ or $y(x)$ data by curve-fitting and use of the F test as recommended by Pettersson & Pettersson (1970)? The answer we can give is that the overall chance is about 0.6, but this decreased with increasing degree of the mechanism in the study we performed as follows: 1:1 (0.98), 2:2 (0.71), 3:3 (0.43) and 4:4 (0.34).

Only slight differences are expected between 95% and 99% confidence levels, and use of $F(m_2 - m_1, n - m_2) + Q_1/Q_2$ as F statistic leads to a slightly higher acceptance value but also gives an increase in the number of false positives. An important theoretical point is that, when an explicit version of an enzyme mechanism allowing for $ES \rightleftharpoons EP$ steps leads to increase in degree over the implicit version, there is an extremely small, but nevertheless finite, chance of detecting the higher-degree terms by the F test. It is likely that the overall chance of detecting deviations from Michaelis–Menten kinetics, i.e. terms of at least second degree, is about 0.8, and the probability of estimating binding constants or outer coefficients of rate equations within an order of magnitude is about 0.85, but inner coefficients of 3:3 and 4:4 functions can only be reliably estimated when the curve is of a complicated type. If an experimenter performs a sequence of, say, ten or so different $v([S])$ plots at different concentrations of second substrate or modifier, as we have done, and finds that the F test distinguishes degree 4:4 from 3:3 with only one or two of the curves but degree 2:2 from 1:1 with all of them, this is to be interpreted as normal behaviour and indicates a mechanism of minimum degree 4:4.

F. J. B. thanks the British Council for a scholarship, and A. J. W. thanks the North Western Regional Health Authority for financial support during this project. Again we thank Dr. K. J. Indge for help with computation and his never-failing patience in helping us to check the rate equations.

References

- Askelof, P., Korsfeldt, M. & Mannervik, B. (1976) *Eur. J. Biochem.* **69**, 61–67
- Bardsley, W. G. (1976) *Biochem. J.* **153**, 101–117
- Bardsley, W. G. (1977a) *J. Theor. Biol.* **67**, 121–139
- Bardsley, W. G. (1977b) *Biochem. Soc. Trans.* **5**, 756–758
- Bardsley, W. G. & Childs, R. E. (1975) *Biochem. J.* **149**, 313–328
- Bardsley, W. G. & Waight, R. D. (1978a) *J. Theor. Biol.* **70**, 135–156

- Bardsley, W. G. & Waight, R. D. (1978b) *J. Theor. Biol.* **72**, 321–372
- Bardsley, W. G., Leff, P., Kavanagh, J. P. & Waight, R. D. (1980a) *Biochem. J.* **187**, 739–765
- Bardsley, W. G., Woolfson, R. & Wood, R. M. (1980b) *J. Theor. Biol.* **85**, 45–51
- Childs, R. E. & Bardsley, W. G. (1975) *Biochem. J.* **145**, 93–103
- Cleland, W. W. (1967) *Adv. Enzymol. Relat. Areas Mol. Biol.* **29**, 1–32
- Cornish-Bowden, A. & Endrenyi, L. (1981) *Biochem. J.* **193**, 1005–1008
- Endrenyi, L. (ed.) (1981) *Kinetic Data Analysis*, Plenum Press, New York
- Feraudi, M., Kohlmeier, M. & Schmolz, G. (1977a) *J. Mol. Catal.* **2**, 171–177
- Feraudi, M., Kohlmeier, M., Glaser, W. & Keleti, T. (1977b) *Acta Biochim. Biophys. Acad. Sci. Hung.* **12**, 317–327
- Garfinkel, L., Kohn, C. & Garfinkel, D. (1977) *CRC Crit. Rev. Bioeng.* **2**, 329–361
- Haarhof, K. N. (1969) *J. Theor. Biol.* **22**, 117–150
- Hill, C. M., Waight, R. D. & Bardsley, W. G. (1977) *Mol. Cell. Biochem.* **15**, 173–178
- Indge, K. J. & Childs, R. E. (1976) *Biochem. J.* **155**, 567–570
- Lam, C. F., Schatz, N. & Priest, D. (1977) *Comput. Biol. Med.* **7**, 209–221
- Lindgren, B. W. (1976) *Statistical Theory*, pp. 542–545, Macmillan, New York
- Mannervik, B. (1975) *BioSystems* **7**, 101–119
- Matthews, R. H. & Alben, J. (1978) *J. Theor. Biol.* **72**, 219–230
- Monod, J., Wyman, J. & Changeux, J.-P. (1965) *J. Mol. Biol.* **12**, 88–118
- Ottaway, J. H. & Apps, D. K. (1972) *Biochem. J.* **130**, 861–870
- Pettersson, G. & Pettersson, I. (1970) *Acta Chem. Scand.* **24**, 1275–1286
- Reich, J. G. (1970) *FEBS. Lett.* **9**, 245–251
- Solano-Muñoz, F., McGinlay, P. B., Woolfson, R. & Bardsley, W. G. (1981a) *Biochem. J.* **193**, 339–352
- Solano-Muñoz, F., Bardsley, W. G. & Indge, K. J. (1981b) *Biochem. J.* **195**, 589–601
- Storer, A. C., Darlison, M. G. & Cornish-Bowden, A. (1975) *Biochem. J.* **151**, 361–367
- Waight, R. D. & Bardsley, W. G. (1977) *Biochem. Soc. Trans.* **5**, 785–789
- Waight, R. D., Leff, P. & Bardsley, W. G. (1977) *Biochem. J.* **167**, 787–798
- Wardell, J. M., Kavanagh, J. P., Bardsley, W. G. & Wood, R. M. (1982) *J. Theor. Biol.* **95**, 465–487
- Whitehead, E. P. (1979) *J. Theor. Biol.* **80**, 355–381
- Wilkinson, G. N. (1961) *Biochem. J.* **80**, 324–332
- Wyman, J. (1972) *Curr. Top. Cell. Regul.* **6**, 209–226