Biochimica et Biophysica Acta, 1080 (1991) 231-236
© 1991 Elsevier Science Publishers B.V. All rights reserved 0167-4838/91/$03.50

BBAPRO 34040

Experimental designs for estimating the parameters of the Michaelis-Menten equation from progress curves of enzyme-catalyzed reactions

Ronald G. Duggleby and Robyn B. Clarke

Department of Biochemistry, University of Queensland, Queensland (Australia)

(Received 30 April 1991)

Key words: Progress curve, Experimental design; Enzyme kinetics; Michaelis-Menten kinetics; Kinetic analysis

When the progress curve for an enzyme catalysed reaction follows the integrated Michaelis-Menten equation, the maximum velocity and Michaelis constant can be determined from a single such curve. In this paper, an experimental design for collecting the data is proposed which is close to optimum in the sense that it produces the smallest standard error in the estimate of the Michaelis constant. This design involves choosing an initial substrate concentration which is approx. two or three times the value of the Michaelis constant.

Introduction

It has been known for many years that the progress curves of enzyme-catalyzed reactions depend on the kinetic properties of the enzyme. It follows that from an analysis of the shape of these curves it is possible to extract information concerning the kinetic constants. Considerable advances have been made towards this aim (e.g., Duggleby and Wood [1] and references therein) and it is now possible to analyse quite complex reactions.

Attention in the study of enzyme progress curves has focussed almost entirely upon the problems of data analysis, with experimental design being considered only rarely [2,3]. However, the success of the analysis will always depend on whether the progress curves contain sufficient information to define the parameters being estimated.

The simplest of cases is a stable enzyme catalysing an irreversible reaction with a single substrate and for which none of the reaction products is inhibitory. For such a reaction, a single progress curve is sufficient to estimate the maximum velocity and the Michaelis constant, the two kinetic parameters that define this system. To exploit this capability, it is necessary to decide three questions. What substrate concentration should be used? How many data points should be collected along the curve? And how should the points be spaced? The second of these questions involves a complex interplay between the desired accuracy of the results required, the experimental accuracy of the measurement system, and the amount of effort one is prepared to expend. It is the first and last of these three questions which will be addressed here.

Correspondence: R.G. Duggleby, Department of Biochemistry, University of Queensland, QLD 4072, Australia.

Theory

Fundamentals

When a stable enzyme catalyses an irreversible reaction with a single substrate and which is not subject to product inhibition, the progress curve is defined by the integrated Michaelis-Menten equation (Eqn. 1).

$$V_m t = z - K_a \ln(1 - z/A_0) \quad (1)$$

In this equation, V_m (the maximum velocity) and K_a (the Michaelis constant) are the kinetic parameters to be determined from the data, while A_0 is the initial concentration of the substrate. The amount of product formed by reaction at any time (t) is represented by z and a progress curve will consist of a group of measurements of z at various times. It should be stressed that Eqn. 1 only applies to reactions which are irreversible, either intrinsically or by virtue of some auxiliary reaction which removes the product.

Three general conclusions about the form of a progress curve may be drawn from Eqn. 1: firstly, V_m

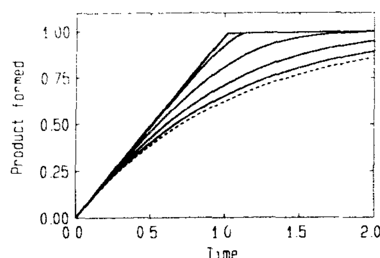

Fig. 1. Theoretical progress curves for a range of substrate concentrations. Progress curves were calculated from Eqn. 1, at a K_a value of 1.0 while varying A_0 to give A_0/K_a ratios (solid lines; top to bottom) of 1024, 32, 4, 1 and 0.25; in order to emphasise the differing shapes of the curves, V_m was varied to equalise the initial velocities while the amount of product was expressed as a fraction of A_0. Thus, each curve has a common tangent at $t = 0$ and a common asymptote at $t = \infty$. The broken line represents an exponential curve.

determines only the scale of the time axis; secondly, the shape of the curve is dependent upon the relative values of A_0 and K_a; and thirdly, at a fixed ratio of A_0/K_a, the scale of the concentration axis depends only on A_0. The second point is illustrated in Fig. 1.

At high A_0/K_a, the curve is approximated by a straight line which shows an abrupt transition as z nears A_0. Over most of its length, the line depends only on V_m and data collected in an experiment such as this would define V_m quite well, but would not contain much information from which to determine K_a. At low A_0/K_a, the curve is approximated by an exponential which has A_0 as an asymptote. The shape of the curve depends on the ratio V_m/K_a and data obtained in such an experiment would define this ratio quite well but would not contain much information from which to determine either V_m or K_a separately. Between these extremes, the curve is a combination of linear and exponential elements. If the objective of the experiment is to determine both V_m and K_a, A_0 should be neither very high nor very low relative to K_a.

Design criteria

A general design criterion has been described by Box and Lucas [4], based on selecting the independent variables so as to minimize of the volume of the joint confidence region in parameter space. This volume is proportional to the determinant of a matrix $(X^TX)^{-1}$, where X is a matrix in which each element is formed by evaluating one of the partial derivatives at one of the design points. These partial derivatives, calculated by differentiating the fitted equation ('the model') with respect to the parameters, are independent of the parameters for linear models and the design can be chosen without any knowledge of the likely values of the parameters. For nonlinear models, such as Eqn. 1, the partial derivatives are functions of the parameters and the design can only be selected if some estimate of the parameters values is at hand.

The application of the Box-Lucas criterion to initial velocity enzyme kinetic studies has been explored by a number of workers [5-11] while Markus and Plesser [2] reported a case study where this criterion was used in progress curve experiments.

Two other design criteria have been proposed for enzymological studies. The first is the use of a 'discrimination function' (Mannervik [12] and references therein) and its use is restricted to situations where one is trying to determine which of two rival models is most appropriate. The discrimination function is used to find the experimental conditions which lead to the greatest difference in the predictions of the two models. The other design criterion, and one of more relevance to the present study, is parameter sensitivity analysis [10]. The influence of an independent variable on the partial derivatives is assessed and the design is then based on values of the independent variable where these derivatives are largest. A report of this type of approach for the design of progress curve experiments has appeared [3] in which it was concluded that for data obeying Eqn. 1, the most useful data is that collected between 60 and 80% substrate utilization using 'the highest practicable initial substrate concentration'.

In the present work, the focus will be on experimental designs which minimize the standard error of the estimate of K_a, for two reasons. First, the K_a is usually of greater interest than V_m since it is independent of the enzyme concentration which itself may be difficult to determine with accuracy. Second, as will be shown, the coefficient of variation of K_a (i.e. standard error of K_a as a fraction of the value of K_a) is always larger than that of V_m. Thus a design which gives a good estimate of K_a will give an even better estimate of V_m. Similarly, the coefficient of variation of K_a is always larger than that of the ratio V_m/K_a and some enzymologists (e.g., Cleland [13]) prefer V_m/K_a to K_a as a 'fundamental constant' describing the properties of an enzyme.

Calculations

The partial derivatives of Eqn. 1 are given as Eqns. 2 and 3, while the matrix of second derivatives (X^TX) is given as Eqn. 4.

$$\delta z/\delta V_m = t/[1 + K_a/(A_0 - z)] \quad (2)$$

$$\delta z/\delta K_a = \ln(1 - z/A_0)/[1 + K_a/(A_0 - z)] \quad (3)$$

$$X^TX = \begin{vmatrix} \sum w(\delta z/\delta V_m)^2 & \sum w(\delta z/\delta V_m)(\delta z/\delta K_a) \\ \sum w(\delta z/\delta V_m)(\delta z/\delta K_a) & \sum w(\delta z/\delta K_a)^2 \end{vmatrix} \quad (4)$$

In this matrix, w is a weighting factor which is inversely proportional to the variance of the corresponding observation. In the present paper, all observations are

assumed to be equally accurate and the weights were set to unity. The determinant (D) of X^TX is given as Eqn. 5 while Eqns. 6–8 are the standard errors of V_m, K_a and V_m/K_a.

$$D = \sum w(\delta z/\delta V_m)^2 \sum w(\delta z/\delta K_a)^2 - \left[\sum w(\delta z/\delta V_m)(\delta z/\delta K_a)\right]^2 \quad (5)$$

$$\sigma(V_m) = \sigma_{pe}\sqrt{\left[\sum w(\delta z/\delta K_a)^2/D\right]} \quad (6)$$

$$\sigma(K_a) = \sigma_{pe}\sqrt{\left[\sum w(\delta z/\delta V_m)^2/D\right]} \quad (7)$$

$$\sigma(V_m/K_a) = (V_m/K_a)\sqrt{\left[\{\sigma(V_m)/V_m\}^2 + \{\sigma(K_a)/K_a\}^2 - c\right]} \quad (8)$$

where

$$c = \sigma_{pe}^2\left[2\sum w(\delta z/\delta V_m)(\delta z/\delta K_a)/(DV_mK_a)\right] \quad (9)$$

It should be noted that the expressions for the standard errors are proportional to σ_{pe} which is the standard error of the experimental data. Thus the accuracy of the data determines the size of the errors in the parameters. However, for the purposes of comparing different designs, it will be assumed that the relative accuracy of the data is not influenced by the experimental design. In other words, observations are equally accurate anywhere along a progress curve and that if the substrate concentration is changed by a certain factor, the size of experimental errors changes by this same factor.

The formulae used to calculate these standard errors, while based on those which are used widely in nonlinear regression analysis, may be inexact as they are derived from a truncated Taylor series. There have been no studies on whether these formulae give realistic values when applied to progress curves although here it is assumed that they do and some limited evidence on this point will be presented.

Experimental procedures

Data collection

Progress curves for pyruvate kinase were determined at 30° in 2.7 ml reaction mixtures containing 5.6 mM ADP, 11 mM $MgCl_2$, 22 mM KCl and 10 IU lactate dehydrogenase in 0.1 M Tris-HCl buffer (pH 8.0). Phospho*enol*pyruvate was varied over the range from 20 to 200 μM and sufficient NADH added to ensure that there was an excess over phospho*enol*pyruvate of between 60 and 120 μM. The reaction was started by adding between 0.2 and 0.7 IU of pyruvate kinase, the amount being chosen so that approx. 10% of the initial phospho*enol*pyruvate remained after 20–30 min. After a set period of incubation, the reaction was stopped by adding 0.3 ml of 5% sodium dodecyl sulphate and the absorbance at 340 nm was determined. Between 7 and 10 identical reaction mixtures were prepared for each progress curve and the reaction stopped after different times.

The reaction catalysed by pyruvate kinase has two substrates and two products. However, under the conditions employed (coupling with lactate dehydrogenase to remove pyruvate and a high MgADP concentration which remains effectively constant and also prevents product inhibition by accumulating MgATP), it behaves in progress curve experiments as if it were a single substrate, irreversible reaction [14].

Data analysis

A general progress curve equation [15] was fitted to the primary absorbance data by nonlinear regression using the DNRP53 computer program [16]. From this analysis, the initial velocity, half-time for reaction, total absorbance change and starting absorbance were determined. Each datum was subtracted from this starting absorbance and the amount of pyruvate formed by the pyruvate kinase reaction was then calculated. In addition, the time axis for all progress curves was normalised to a common scale by multiplying by the pyruvate kinase concentration. Normalised progress curves obtained at the same concentration of phospho*enol*pyruvate, but with different amounts of pyruvate kinase, were found to be superimposible, showing that the system is stable under the experimental conditions employed [17].

The integrated Michaelis-Menten equation (Eqn. 1) was fitted to the normalised progress curves using the DNRP53 computer program to obtain values and standard errors for V_m and K_a. A typical result is shown in Fig. 2.

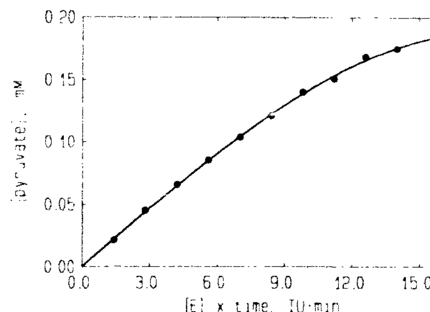

Fig. 2. A progress curve for the pyruvate kinase reaction. Conditions were as described in Experimental procedures, using 0.195 mM phospho*enol*pyruvate and 0.60 U of pyruvate kinase. The *line* represents the best fit of Eqn. 1 to the data and corresponds to values for V_m and K_a of $1.965 \pm 0.071 \cdot 10^{-2}$ mM IU^{-1} min^{-1} and $4.300 \pm 0.552 \cdot 10^{-2}$ mM, respectively.

TABLE I

Effect of data spacing on parameter errors

The standard errors of V_m, K_a and V_m/K_a were calculated using Eqns. 6–8 for two types of experimental design with data equally spaced on either the time axis or the product concentration axis. In each case the final point was the same, and corresponded to the % reaction shown in the first column. In all cases, A_0/K_a was set at 4.0 with nine points per curve. The actual predicted errors in the parameters would be calculated from the values shown by multiplying by the fractional error in the data and the parameter value. For example, if the data error was 0.01 (i.e. 1%) and V_m was 3.0, the first value shown in the second column (5.92) corresponds to a standard error of 0.01·3.0·5.92 = 0.1776

% Reaction	V_m		K_a		V_m/K_a	
	product	time	product	time	product	time
99	5.92	5.13	19.78	15.78	13.97	10.79
90	5.75	5.64	19.20	18.26	13.56	12.73
81	7.32	7.29	25.59	25.19	18.35	17.97
72	9.80	9.80	36.12	35.90	26.38	26.16
63	13.58	13.59	52.58	52.45	39.06	38.92
54	19.55	19.56	79.25	79.16	59.75	59.65
45	29.68	29.69	125.45	125.39	95.81	95.74

Results

Data spacing

In principle, data can be collected at any combination of points along a progress curve but in practice only two ways are likely. The first, and most obvious, is to collect data at preset, regular time intervals up to some limit. However, this may not be the best way as the points will be widely spaced on the concentration axis at early times when product is accumulating rapidly and closely spaced in concentration as the reaction slows towards the end. This possible deficiency may be overcome by collecting the data so that it is regularly spaced on the concentration axis. While this may not be feasible in all experimental systems, it is relatively easy to achieve if the reaction has a signal which can be monitored continuously; all that needs to be done is to note the time that the signal reaches the desired value. These two designs are compared in Table I from which a number of conclusions can be drawn.

Firstly, there is not a great deal to choose between the two data spacings although in the majority of cases, spacing by time is somewhat better than spacing by concentration. In the few instances where the latter design was better, the improvement was extremely small. This general pattern was also found at A_0/K_a ratios ranging from 0.1 to 10.0 (data not shown). Thus the simpler design to implement experimentally, collecting data at equal intervals of time, is also the better design from a statistical standpoint. Unless otherwise noted, all calculations shown below were made with this design.

Secondly, the data span (shown in the first column of Table I) has a very large influence and restricting the data to early parts of the progress curve cause the standard errors to become considerably larger. This point will be expanded upon later.

Thirdly, the standard error of K_a is larger than those of both V_m and V_m/K_a. As is documented later, this was also found to be generally true. However, although data in Table I indicate that the error in V_m is smaller than that in V_m/K_a, many instances were found where the relativities were reversed and this point will also be discussed below.

To verify that the standard errors calculated from Eqns. 6–8 which are presented in Table I and elsewhere in this paper are of about the right magnitude, a limited Monte Carlo study was performed. 25 progress curves were simulated with 1% error for the conditions shown in Table I, row 2, with data equally spaced on the product concentration axis. When these were analysed, the standard errors of V_m ranged from 0.0157 to 0.0758 with a mean of 0.0487, in reasonable agreement with the predicted value of 0.0575. Similarly, the standard errors of K_a (range: 0.0514 to 0.2558; mean: 0.1619) were comparable to the predicted value of 0.1920.

Initial substrate concentration

The data in Table I were all calculated at a fixed A_0/K_a ratio of 4.0; the effect of varying this ratio is illustrated in Table II. As expected V_m is poorly determined at low substrate concentrations and the error

TABLE II

Effect of substrate concentration on parameter errors

The standard errors of V_m, K_a and V_m/K_a were calculated using Eqns. 6–8 with nine data point equally spaced on the time axis. The final point corresponded to 75% utilization of the initial substrate concentration which is shown in the first column. In all cases, K_a was set at 1.0 so the values in the first column also represent A_0/K_a ratios

A_0	V_m	K_a	V_m/K_a
0.1	126.60	135.06	8.54
0.2	66.34	75.19	8.93
0.4	36.22	45.86	9.72
0.7	23.28	34.07	10.87
1.0	18.08	30.01	12.01
1.5	14.01	27.82	13.89
2.0	11.96	27.63	15.74
2.5	10.73	28.23	17.58
3.0	9.90	29.23	19.40
4.0	8.85	31.82	23.03
5.0	8.22	34.79	26.64
6.0	7.79	37.95	30.23
7.0	7.49	41.23	33.81
8.0	7.26	44.57	37.38
9.0	7.08	47.96	40.94
10.0	6.93	51.38	44.51

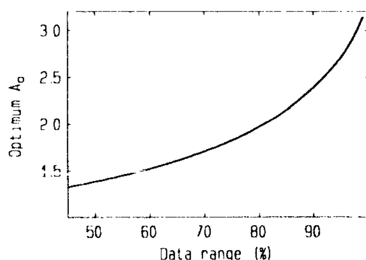

Fig. 3. Optimum initial substrate concentrations for progress curves. The K_a was set at 1.0 and the standard error was calculated using Eqn. 7, with nine data points equally spaced on the time axis; the final point corresponded to percent utilization (shown on the abscissa) of the initial substrate concentration. The ratio of A_0/K_a was varied until a value was found which gave the minimum standard error and it is this value which is plotted on the ordinate.

becomes progressively smaller as A_0 is increased. The converse pattern is seen for V_m/K_a; the error becomes progressively larger as A_0 is increased. The most interesting result is for K_a; it has a large error at both large and small A_0, passing through a minimum near an A_0/K_a ratio near 2.0. Thus, an intermediate substrate concentration is best for determining K_a and this has the added advantage of producing comparable and not excessive errors in both V_m and V_m/K_a.

Similar calculations to those shown in Table II were performed at many A_0/K_a ratios to determine the exact value which gave the lowest possible standard error for K_a. This optimum concentration was then determined for data ranges from 45 to 99% substrate conversion; the results are shown in Fig. 3. While there is clearly a dependence upon the data range, the effect is not massive and A_0/K_a varies from 1.33 to 3.14 between 45 and 99% substrate conversion.

Design efficiency

In selecting conditions to perform a progress curve experiment, and in particular choosing the best initial substrate concentration, it is necessary to know K_a. This is something of a paradox when the purpose of the experiment is to determine K_a. Thus, it is necessary to have an estimate of K_a, either from preliminary experiments or from educated guess-work.

The question then arises of how much influence a poor estimate will have and this point is addressed in Table III. The efficiencies shown indicate how much smaller the standard error could have been if K_a had been known precisely while designing the experiment. Clearly, a poor estimate of K_a does not lead to a particularly bad experiment. For example, a 5-fold error (on either side) in the estimate of K_a gives an efficiency of approx. 60%; even if K_a had been known precisely in advance, only a 40% reduction in the standard error of K_a could have been expected.

TABLE III

Effect of an inaccurate estimate of K_a on the selection of an initial substrate concentration

The data consisted of 9 points equally spaced on the time axis. The final point corresponded to 90% utilization of the initial substrate concentration of 2.3874, which is optimal for a K_a value of 1.0. The actual K_a was varied from 0.1 to 10.0 (first column) and the standard error of K_a was calculated using Eqn. 7 at A_0 values of 2.3874 and at the optimum concentration for the actual K_a. The 'efficiency' of the design is the ratio of the standard errors at these two concentrations

K_a	Efficiency
0.1	0.3772
0.2	0.6000
0.4	0.8388
0.7	0.9728
1.0	1.0000
1.3	0.9849
1.6	0.9534
2.0	0.8997
2.5	0.8324
3.0	0.7698
5.0	0.5809
7.0	0.4633
10.0	0.3451

Experimental verification

To further test the proposed experimental design, the reaction catalysed by pyruvate kinase was selected for study. Under the conditions employed it behaves kinetically as if it is an irreversible reaction with a single substrate, phospho*enol*pyruvate. Preliminary investigations using initial velocity measurements (data not shown) indicated a K_a in the vicinity of 40 μM in agreement with earlier studies [14] under similar experimental conditions.

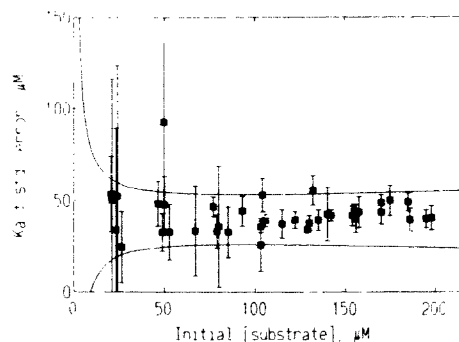

Fig. 4. Estimates of K_a obtained from a series of progress curves for pyruvate kinase. The initial phospho*enol*pyruvate concentration was varied from 20 to 200 μM and progress curves collected and analysed as dscribed in Experimental procedures. The *points* indicate the estimated values of K_a while the *bars* show the standard errors. The solid lines show the expected shape of the error envelope, calculated using Eqns. 2, 3, 5 and 7.

Experimental designs were as described above (data range = 90% substrate utilization), except that the substrate concentration was varied from 20 to 200 μM, corresponding to the range 0.5–5.0 times the estimated K_a. The results from 39 progress curve experiments are summarised in Fig. 4. All values of K_a are clustered around 40 μM (weighted mean = 39.6 ± 0.7 μM) and the values appear to be independent of the substrate concentration, as expected. However, the standard errors show a clear trend, being quite large at low substrate concentrations and decreasing significantly as the substrate concentration is increased. In most experiments, the standard error obtained at about 100 μM (2.5 times K_a) were reasonably small and were not reduced further at higher substrate concentrations. These results are in broad agreement with the theoretical predictions.

Discussion

The results presented here allow some recommendations concerning the design of progress curve experiments where the data follow the integrated Michaelis-Menten equation. The first factor that should be considered is the substrate concentration to be used, which may be limited by economic or technical considerations. As a rule of thumb it should be 2.5-times K_a (Fig. 3), or as high as possible if this concentration exceeds practical limits. Obviously there is paradox here as the purpose of the experiment is determine K_a, while K_a must be known to design the experiment. It turns out (Table III) that this is no real problem as a rather poor guess at what K_a might be still results in reasonably efficient designs.

Once the substrate concentration to use has been decided, it is now necessary to determine where and how the data are to be spaced along the progress curve. The results (Table I) indicate that data should be collected until the reaction is 90% completed but that there is no advantage in extending beyond this point. A secondary advantage of limiting the data span is that as catalysis slows due to substrate exhaustion, it would take a long time to collect data from the last few percent of the reaction.

The conclusions above that A_0 should be 2–3-times K_a and that data should be collected until the reaction is about 90% complete are somewhat at variance with those of Vandenberg et al. [3] who suggest that A_0 should be as high as possible while the reaction should be followed to 60–80% completion. This differing conclusion is the result of the different criterion used to select the best design. Vandenberg et al. [3] base their design upon parameter sensitivity analysis which involves determining maxima for the partial derivatives (Eqns. 2 and 3). While it is clear that an optimum design must be one in which parameter values affect the shape of the progress curve, it does not necessarily follow that maximising the partial derivatives results in the best design.

The design advocated here is one which minimises the standard error of K_a, which is usually the parameter of most interest. If it is argued that V_m or V_m/K_a are of more interest in some circumstances, it is worth noting that both V_m and V_m/K_a always have smaller standard errors than K_a (Tables I and II). Moreover, conditions for minimising the error in V_m are those which maximise the error in V_m/K_a, and vice versa. Thus, selecting conditions to achieve the smallest error in K_a is a useful compromise.

This experimental design is useful only when the enzyme concerned catalyses an irreversible reaction with one substrate and where product inhibition is insignificant. Only in these circumstances is the progress curve described by the integrated Michaelis-Menten equation. When there is more than one substrate, or product inhibition, or reversibility, more complex kinetic equations are needed and the parameters can only be determined from the combined results of several progress curves. Further work is required to determine the best designs in such situations.

Acknowledgement

This work was supported by grant A08715141 from the Australian Research Council.

References

1 Duggleby, R.G. and Wood, C. (1989) Biochem. J. 258, 397–402.
2 Markus, M. and Plesser, T. (1976) Biochem. Soc. Trans. 4, 361–364.
3 Vandenberg, J.I., Kuchel, P.W. and King, G.F. (1986) Analyt. Biochem. 155, 38–44.
4 Box, G.E.P. and Lucas, H.L. (1959) Biometrika 46, 77–90.
5 Bártfai, T. and Mannervik, B. (1972) FEBS Lett. 26, 252–256.
6 Bezeau, M. and Endrenyi, L. (1986) J. Theor. Biol. 123, 415–430.
7 Duggleby, R.G. (1979) J. Theor. Biol. 81, 671–684.
8 Duggleby, R.G. (1981) in Kinetic Data Analysis: Design of Experiments for Estimating Enzyme and Pharmacokinetic Parameters (Endrenyi, L., ed.), pp 169–179. Plenum Press, New York.
9 Endrenyi, L. (1981) in Kinetic Data Analysis: Design of Experiments for Estimating Enzyme and Pharmacokinetic Parameters (Endrenyi, L., ed.), pp. 137–167. Plenum Press, New York.
10 Kanyár, B. (1978) Acta Biochim. Biophys. Acad. Sci. Hung. 13, 153–160.
11 Reich, J.G. (1970) FEBS Lett. 9, 245–251.
12 Mannervik, B. (1982) Methods Enzymol. 87, 370–390.
13 Cleland, W.W. (1975) Accounts Chem. Res. 8, 145–151.
14 Duggleby, R.G. (1983) Biochim. Biophys. Acta 744, 249–259.
15 Duggleby, R.G. (1985) Biochem. J. 228, 55–60.
16 Duggleby, R.G. (1984) Comput. Biol. Med. 14, 447–455.
17 Selwyn, M.J. (1965) Biochim. Biophys. Acta 105, 193–195.