Analysis of Kinetic Data for Allosteric Enzyme Reactions as a Nonlinear Regression Problem

J. KOWALIK AND J. F. MORRISON
The Computer Centre and the Department of Biochemistry
Australian National University, Canberra, Australia

Communicated by Richard Bellman

ABSTRACT

The effect of substrate concentration on the initial velocity of the reaction catalyzed by the allosteric enzyme, nucleoside diphosphatase, has been determined. When the results were plotted in double reciprocal form, concave-up nonrectangular hyperbolas were obtained. Thus the data are qualitatively in accord with a mechanism that involves the interdependent reaction of two molecules of $MgIDP^-$ with the enzyme and the formation of product from both enzyme-MgIDP and enzyme-$(MgIDP)_2$ complexes. Further, it has been shown that the data can be fitted to the initial velocity equation for the proposed mechanism, which is the ratio of two quadratic polynomials in substrate concentration, and values have been obtained for the various kinetic constants. For the purpose of fitting the data, a variable metric method has been developed and used, since the Gaussian method did not always give convergency. The relative merits of these two procedures are discussed.

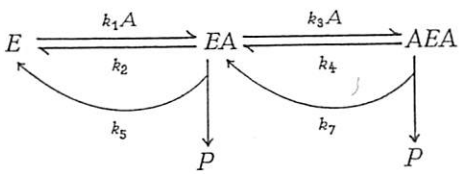
It is well established that a number of cellular reactions are catalyzed by allosteric enzymes, which differ from the classical enzymes in that they do not exhibit Michaelis-Menten kinetics and are subject to the influence of modifiers [1]. Recognition of this type of enzyme has come frequently from the finding that a plot of the initial reaction velocity against the substrate concentration yields a sigmoidal-shaped curve. The majority of kinetic investigations, however, have been primarily qualitative and hence not designed to yield information about the overall reaction mechanism or values for kinetic constants. A certain amount of emphasis has been placed on the use of Hill plots to determine the number of substrate molecules involved in the reaction [1], but the

Mathematical Biosciences 2, 57–66 (1968)

resulting conclusions are of doubtful validity [2]. This is especially true if the maximum velocity of the reaction has not been determined precisely.

Various general theories, involving a thermodynamic approach and the interaction of enzyme subunits, have been proposed to account for the kinetic results obtained for allosteric enzyme reactions [3–5]. But with some exceptions [6–9], it would seem that there has been a tendency to overlook or discount the possibility that such results might well be explained in terms of classical enzyme kinetics. The reason for this is not clear, especially when it is realized that the interdependent reaction of two molecules of substrate at the active site of an enzyme can give rise to sigmoidal curves for plots of initial velocity against substrate concentration. When data that give such sigmoidal curves are plotted in double reciprocal form, either parabolas or nonrectangular hyperbolas can be obtained, and these types of plots have been illustrated commonly in the literature [7, 10–12]. Further, it has been shown that the parabolic double reciprocal plot obtained from studies on $NADH_2$ oxidase is consistent with the compulsory addition of two molecules of substrate to the enzyme before product is released [7].

As part of a program to obtain more detailed information about the mechanism of allosteric enzyme reactions, kinetic studies have been undertaken in this laboratory of the reaction catalyzed by IDPase [13]. These studies have indicated that double reciprocal plots of initial velocity against the concentration of $MgIDP^-$ yield nonrectangular hyperbolas (cf. Fig. 1) and the previous report that linear plots are obtained in the presence of $MgATP^{2-}$ has been confirmed. Thus the reaction can be described, at least qualitatively, by the mechanism



where A and P represent $MgIDP^-$ and products, respectively. The initial velocity equation for the reaction can be expressed as

$$v = \frac{V[A^2 + (k_5/k_7)K_{a2}A]}{K_{a1}K_{a2} + K_{a2}A + A^2} \tag{1}$$

where $V = k_7 e_t$, $K_{a1} = (k_2 + k_5)/k_1$, and $K_{a2} = (k_4 + k_7)/k_3$; or in reciprocal form as

$$\frac{1}{v} = \frac{1 + K_{a2}(1/A) + K_{a1}K_{a2}(1/A)^2}{V[1 + (k_5/k_7)K_{a2}(1/A)]}, \tag{2}$$

which represents an equation of a nonrectangular hyperbola or a 2/1 function [14]. The initial velocity equation may also be written in general form as

$$v = \frac{V(A^2 + dA)}{A^2 + bA + c}. \tag{3}$$

In an endeavor to determine if the experimental data could be fitted quantitatively to Eq. (3), a number of sets were analyzed by means of a computer program written by Cleland [14]. However, significant fits were not obtained, either because of convergence to negative values of some of the kinetic parameters or because of divergency. In the latter case, the values for some parameters increased markedly and continued to increase as a function of the number of iterations. Thus it appeared that either the data were not sufficiently good or the reaction mechanism could not be described as outlined in the foregoing. To distinguish between these alternative explanations, efforts have been directed toward the development of a new computer program for the analysis of the data, one that utilizes a gradient minimization method rather than the Gaussian procedure used by Cleland. A general indication of its operation and a comparison of the two methods are elaborated in the following.

The general problem involves the fitting of K sets of values

$$y_k, \mathbf{X}_k = (X_{1k}, X_{2k}, \ldots, X_{nk}); \qquad k = 1, 2, \ldots, K \tag{4}$$

to the nonlinear function $f(\mathbf{X}_k, \mathbf{p})$ in the parameter $\mathbf{p} = (p_1, p_2, \ldots, p_R)$ so as to obtain values of $\mathbf{p}$ for which the sum of squares

$$F(\mathbf{p}) = \sum_{k=1}^{K} (f(\mathbf{X}_k, \mathbf{p}) - y_k)^2 = \sum_{k=1}^{K} d^2(\mathbf{X}_k, \mathbf{p}) \tag{5}$$

is minimized. If additional restrictions, which require that all or selected p_r values be positive, must be imposed on the solution, then this requirement can be met by introducing a new variable $\bar{p}_r$ in place of p_r, where $\bar{p}_r^2 = p_r$. In general, such a nonlinear problem cannot be solved in one computational step and an iterative procedure has to be used.

If $\mathbf{p}^0$ represents an approximate solution of the foregoing problem, then an effective iterative procedure should be able to determine a new vector $\mathbf{p}^1$ that gives a better value of $F(\mathbf{p})$ such that $F(\mathbf{p}^1) < F(\mathbf{p}^0)$. One

possible method for obtaining successive $\mathbf{p}^{i+1}$ values from the previous $\mathbf{p}^i$ values is to solve the so-called Gaussian normal set of equations (6).

$$(A^TA)\delta p^i = -A^T\,\mathbf{d}(\mathbf{X}, \mathbf{p}^i) \tag{6}$$

where A is a $(K \times R)$ matrix with the elements

$$a_{k,r} = \left.\frac{\partial f(\mathbf{X}_k, \mathbf{p})}{\partial p_r}\right|_i, \tag{7}$$

$\mathbf{d}(\mathbf{X}, \mathbf{p}^i)$ is the following column vector

$$\mathbf{d}(\mathbf{X}, \mathbf{p}^i) = (d(\mathbf{X}_1, \mathbf{p}^i), d(\mathbf{X}_2, \mathbf{p}^i), \ldots, d(\mathbf{X}_K, \mathbf{p}^i))^T, \tag{8}$$

and T denotes a transposition. Equation (6) is obtained from the quadratic function of the Taylor expansion of $F(\mathbf{p})$ in the neighborhood of $\mathbf{p}^i$. The new approximation is now given by

$$\mathbf{p}^{i+1} = \mathbf{p}^i + \delta p^i = \mathbf{p}^i - (A^TA)^{-1}A^T\,\mathbf{d}(\mathbf{X}, \mathbf{p}^i). \tag{9}$$

When relationship (10) is introduced

$$\begin{aligned}\operatorname{grad} F(\mathbf{p}^i) &= \left[\frac{\partial F(\mathbf{p}^i)}{\partial p_1}, \quad \frac{\partial F(\mathbf{p}^i)}{\partial p_2}, \ldots, \frac{\partial F(\mathbf{p}^i)}{\partial p_R}\right]^T \\ &= 2A^T\,\mathbf{d}(\mathbf{X}, \mathbf{p}^i),\end{aligned} \tag{10}$$

it follows from (9) that

$$\mathbf{p}^{i+1} = \mathbf{p}^i - \tfrac{1}{2}(A^TA)^{-1}\operatorname{grad} F(\mathbf{p}^i). \tag{11}$$

It is clear from (11) that the Gaussian method is essentially a steepest-descent procedure with the metric matrix A^TA. This matrix is always positive definite and consequently the direction vector

$$\mathbf{D}^i = -(A^TA)^{-1}\operatorname{grad} F(\mathbf{p}^i) \tag{12}$$

must lead to a smaller value of the function in every iterative step. If it is assumed that the vector $A^T\mathbf{d}$ does not vanish and that the matrix A^TA is nonsingular, the Gaussian method should produce a sequence of values for $\mathbf{p}^i$, $i = 1, 2, \ldots$, such that $F(\mathbf{p}^1) > F(\mathbf{p}^2) > \cdots$, which would indicate that the method is stable. Unfortunately, unless certain conditions are satisfied, the Gaussian method does not converge to a solution. The practical difficulty in finding δp^i from system (6) is that these equations are ill-conditioned. This is because the vectors of the A matrix are often

nearly linearly dependent. Ill-conditioning is, therefore, an inherent feature of the normal set of equations and the main reason for the failure of the Gaussian method. If the matrix A^TA is ill-conditioned, the solution of system (6) gives very large steps δp^i for which $F(\mathbf{p}^{i+1})$ may be considerably greater than $F(\delta p^i)$. This arises because the relationship given in (6) is approximate and valid only in the neighborhood of $\mathbf{p}^i$. Another disadvantage of the Gaussian method is that the vector $\mathbf{D}^i$ is frequently almost perpendicular to the grad $F(\mathbf{p}^i)$ [18]. Therefore, even a modified Gaussian method [16], which uses a search for a minimum along δp^i

$$\min_{\lambda} F(\mathbf{p}^i + \lambda\, \delta p^i) \tag{13}$$

and then gives

$$\mathbf{p}^{i+1} = \mathbf{p}^i + \lambda^{\min} \delta p^i, \tag{14}$$

may break down, because it sometimes produces unrealistically small values of λ. In general, two requirements are necessary to assure the stability of the Gaussian method when this method is used to solve nonlinear regression problems. They are: (1) the matrix A^TA must be well conditioned; and (2) the starting value $\mathbf{p}^0$ must be a good approximation of the solution and the observed data y_k must be sufficiently accurate to ensure that (A^TA) is a good approximation to the matrix of second-order terms in the Taylor expansion of $F(\mathbf{p})$.

The least squares method can be improved by a number of modifications [15–19]. The major effort has been devoted toward: (a) an improvement of the conditioning of the linear system (6); and (b) reduction of the angle between δp^i and $-$ grad $F(\mathbf{p}^i)$ so that it becomes more acute. However, another possible approach to the solution of a nonlinear least squares problem is the direct use of the gradient minimization method. In this procedure, the new vector $\mathbf{p}^{i+1}$ is defined from the current one according to the equation

$$\mathbf{p}^{i+1} = \mathbf{p}^i - \lambda^i H^i \operatorname{grad} F(\mathbf{p}^i) \tag{15}$$

where H^i is a positive definite matrix and λ^i minimizes $F(\mathbf{p}^{i+1}) = F(\mathbf{p}^i, \lambda^i)$. Such an iterative procedure avoids the necessity of solving Eq. (6). Depending on the choice of H^i matrices, a number of descent methods can be elaborated. That due to Davidon and improved by Fletcher and Powell [20] has proved to be the most powerful of all the optimization techniques currently available and has been used in the present work.

The basic principle of the procedure is to start with $H^0 = I$ and to modify H^i after each step by using the information gained in the preceding step. The method provides the minimization process with directions

$$\mathbf{D}^i = -H^i \operatorname{grad} F(\mathbf{p}^i), \tag{16}$$

which are the gradients relative to the metric $(H^i)^{-1}$. Thus it is always possible to find $\lambda^i > 0$ so that $F(\mathbf{p}^{i+1}) < F(\mathbf{p}^i)$. Positive definite matrices H^i are updated from step to step by easy matrix and vector operations. In addition, the H^i matrix converges to the inverse of Hessian, which is defined as

$$h_{m,n} = \frac{\partial^2 F(\mathbf{p})}{\partial p_m, \partial p_n} \tag{17}$$

when $\mathbf{p}^i$ approaches the optimum value. Thus the H^i matrix provides the information about the nature of the minimum. The method of obtain-

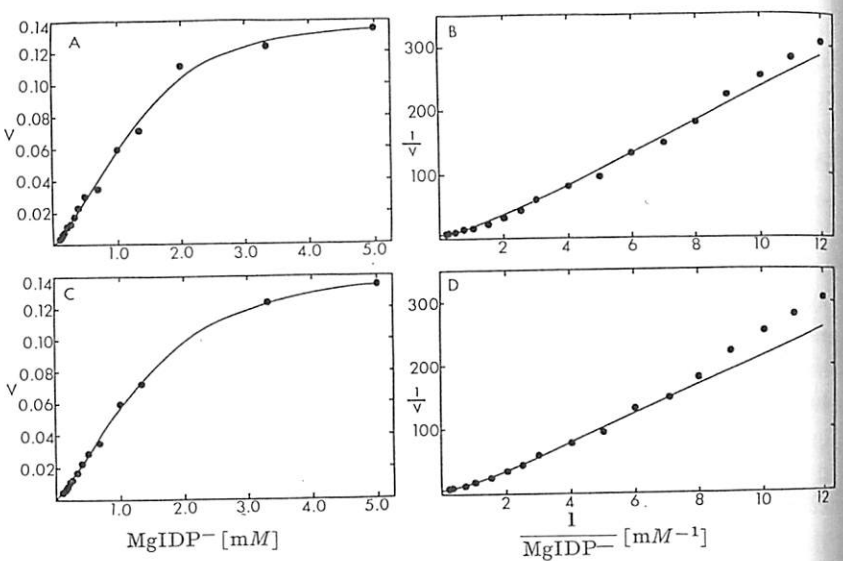


FIG. 1. Effect of the concentration of $MgIDP^-$ on the initial velocity of the IDPase reaction at pH 8.5. The solid lines represent the theoretical curves obtained by substitution of the computed values for V, b, c, and d into the initial velocity equation. Plots A and B illustrate the fit obtained when all 18 experimental values were analyzed by means of the present computer program. Plots C and D show the fit obtained when the velocity value determined at a substrate concentration of 2.0 mM was omitted and the remaining data were analyzed by either the present program or that of Cleland [14].

ing λ^i that locates the minimum of $F(\mathbf{p}^{i+1}) = F(\mathbf{p}^i, \lambda^i)$ along the relative gradient $\mathbf{D}^i$ is not difficult. The cubic interpolation in the Davidon method [20] has been used with satisfactory results.

The computer program, written in PL/1 and utilizing the principles outlined earlier, has been used to analyze initial velocity data obtained from a kinetic investigation of the IDPase reaction. The results (Fig. 1A, B) appear to indicate that the data do give a significant fit to the proposed velocity equation (3) when all experimental points are included in the analysis. Values, together with their standard errors, have been determined for the various parameters and are listed in Table I. It will be noted that the value for b, which must be positive on the basis of the postulated mechanism, is not significant. Analysis of the same data by means of Cleland's computer program [14] actually yielded a negative value for b (Table I). On the other hand, when the velocity value obtained at the relatively high $MgIDP^-$ concentration of 2.0 mM was omitted, analysis of the remaining data by either Cleland's or the present computer program gave the fits illustrated in Fig. 1C, D. Although similar positive values were obtained for the parameters (Table I), there was some variation in the values for their standard errors. The reason for omitting the value just mentioned was that initial analysis, using the present program, showed that its deviation from the curve (Fig. 1A) was virtually equivalent to three times the standard error of the overall fit and was, therefore, probably in error. An advantage of the present program is that it allows of the detection of any values that are probably in error because of factors over and above random variation. Neither of the foregoing data sets gave a significant fit to the equation [21] $v = VA^2/(A^2 + bA + c)$.

From a comparison of the results as illustrated in the form of a plot of velocity against $MgIDP^-$ concentration (Fig. 1C) and in the form of a double reciprocal plot (Fig. 1D), it would appear that in the former case, the experimental points give a better fit to the theoretical curve. This effect is apparent because the double reciprocal plot emphasizes any deviation of the low velocity points, which make only a small contribution to the residual sum of squares. But at the same time, it is this effect that precludes the use of graphical procedures as a means of obtaining precise values for the parameters. The differences in the magnitude of the values determined by graphical and computer techniques are shown in Table I.

In connection with the fitting of experimental data to the given velocity equation (3), it should be pointed out that in order to obtain a good fit, a large number of accurate experimental points are required.

TABLE I

COMPARISON OF VALUES OBTAINED FOR THE VARIOUS PARAMETERS OF THE INITIAL VELOCITY EQUATION[a]

Method of analysis	No. of experimental points	Parameter			
		V	b	c	d
Graphical	17 or 18	0.154	2.05	2.11	0.35
Computer					
Present program	18	0.129 ± 0.029	0.08 ± 0.80	3.28 ± 0.41	1.00 ± 0.39
Cleland's program	18	0.106 ± 0.023	-0.44 ± 0.47	4.70 ± 1.84	1.89 ± 1.35
Present program	17	0.136 ± 0.021	0.56 ± 0.23	3.86 ± 0.31	1.26 ± 0.28
Cleland's program	17	0.137 ± 0.015	0.58 ± 0.34	3.80 ± 1.21	1.22 ± 0.61

[a] The graphical determination of values for b, c, and d involved the drawing of the curve of best fit to the experimental data when plotted in double reciprocal form. From this plot estimates for the initial slope of the curve $S1(1)$, slope of the asymptote $S1(2)$, intercept of the curve Int(1), and extrapolated intercept of the asymptote Int(2) were obtained and were substituted into the relationships

$$d = \frac{S1(1) - S1(2)}{\text{Int}(2) - \text{Int}(1)}; \qquad c = \frac{S1(2) \times d}{\text{Int}(1)}; \qquad b = d + \frac{S1(1)}{\text{Int}(1)}.$$

These are necessary to define the overall shape of the double reciprocal plot over a wide range of substrate concentrations so that good estimates of the slope and extrapolated intercept of the asymptote, as well as the initial slope and intercept of the curved portion of the plot, can be obtained (cf. Fig. 1 and Table I). It is these estimates that provide the starting point for the computer analysis. The need to have accurate values is related to the fact that functions of the type discussed can have local minima whose values are greater than that of the global minimum or places where the function is almost flat and not a minimum. Indeed, one of the inherent difficulties of the present method of analysis is its sensitivity to the value of the starting point. It should also be mentioned that the error in the higher reaction velocities should be small, as otherwise these deviations make a large contribution to the sum of squares and a less significant fit is obtained, especially in the region of low reaction velocities. Elimination of this difficulty might well require the use of analytical methods that involve the continuous monitoring of product formation.

While the present results seem to be in accord with the proposed mechanism for the IDPase reaction, it is not suggested that this is the only mechanism that can account for the results. But the availability of a versatile computer program for the analysis of data that give double reciprocal plots that conform to a nonrectangular hyperbola does make it possible to determine in a quantitative manner if the hypothesis presented earlier might be tenable for other allosteric enzymes. In this connection, it should be borne in mind that the elimination of mechanisms that are not in agreement with the experimental data is an important aspect of kinetic analysis. If the experimental data can be fitted to the type of equation presented earlier, then the kinetic analysis can be extended by deriving velocity equations that allow for the possible effects of modifiers and product inhibitors and by making a comparison of the experimental results with those expected on the basis of theory. A listing of the computer program is available from the authors.

REFERENCES

1 D. E. Atkinson, *Ann. Rev. Biochem.* **35**(1966), 85.
2 A. B. Roy, *J. Mol. Biol.* **10**(1964), 176.
3 J. Monod, J. Wyman, and J-P. Changeux, *J. Mol. Biol.* **12**(1965), 88.
4 D. E. Koshland, G. Nemethy, and D. Filmer, *Biochemistry* **5**(1966), 365.
5 M. M. Rubin and J-P. Changeux, *J. Mol. Biol.* **21**(1966), 265.

6 C. Frieden, *J. Biol. Chem.* **239**(1964), 3522.
7 A. Worcel, D. S. Goldman, and W. W. Cleland, *J. Biol. Chem.* **240**(1965), 3399.
8 B. D. Sanwal, C. S. Stachow, and R. A. Cook, *Biochemistry* **4**(1965), 410.
9 B. D. Sanwal and R. A. Cook, *Biochemistry* **5**(1966), 886.
10 R. Okazaki and A. Kornberg, *J. Biol. Chem.* **239**(1964), 275.
11 P. Datta and L. Prakash, *J. Biol. Chem.* **241**(1966), 5827.
12 P. Datta, *J. Biol. Chem.* **241**(1966), 5836.
13 M. Yamazaki and O. Hayaishi, *J. Biol. Chem.* **240**(1965), PC2761.
14 W. W. Cleland, *Biochim. Biophys. Acta* **67**(1963), 173.
15 D. D. Morrison, in *Proc. Jet Propulsion Lab. Seminar, Los Angeles (1960)*.
16 H. O. Hartley, *Technometrics* **3**(1961), 269.
17 D. P. Feder, in *Recent advances in optimization techniques* (A. Lavi and T. P. Vogl, eds.), pp. 5–21. Wiley, New York, 1966.
18 D. W. Marquardt, *J. Soc. Ind. Math.* **11**(1963), 431.
19 R. J. Pegis, D. S. Grey, T. P. Vogl, and A. K. Rigler, in *Recent advances in optimization techniques* (A. Lavi and T. P. Vogl, eds.), pp. 47–68. Wiley, New York, 1966.
20 R. Fletcher and M. J. D. Powell, *Computer J.* **6**(1963), 163.
21 W. W. Cleland, *Nature* **198**(1963), 463.