A PROCEDURE BASED ON STATISTICAL CRITERIA FOR DISCRIMINATION BETWEEN STEADY STATE KINETIC MODELS

T. BÁRTFAI and B. MANNERVIK
Department of Biochemistry, University of Stockholm, Stockholm, Sweden

Received 19 June 1972
Revised version received 3 August 1972

1. Introduction

A major objective of enzyme kinetic studies is to establish the mathematical model that can describe the mechanism of action of an enzyme under investigation. It is also desirable to determine the numerical values of the parameters included in the model chosen. These problems are usually solved by means of computers. The main approaches in kinetics are simulation (digital and analog) and optimization (cf. Garfinkel et al. [1]). In simulation studies a model is built to describe the system investigated, whereas in the optimization approach a set of numerical values are determined to give the best fit of a predetermined model to experimental data. The present paper is an attempt to develop and summarize criteria for selection of one of several rival models using the optimization techniques widely applied in steady state kinetics (Johansen and Lumry [2]; Berman et al. [3]; Bliss and James [4]; Dammkoehler [5]; Hanson et al. [6]; Kowalik and Morrison [7]; Cleland [8, 9]; Reich et al. [10]). The selection of the best model is generally performed by rejection of all models, which are considered inappropriate by statistical or physico-chemical criteria. A design criterion is introduced to help finding the region of experimentation most suitable for discrimination purposes.

2. Regression procedure

The models of steady state kinetics are nonlinear. To estimate the best set of parameter values of a model, nonlinear optimization techniques are generally used. These procedures can be classified (cf. Swann [11] and Kowalik and Osborne [12]) as

i) direct search methods (methods of alternating variable [11], Hooke and Jeeves [13], Rosenbrock [4], Davies et al. (D.S.C.) [11], Powell [15], and the simplex method [16].
ii) gradient methods (the steepest descent method and Newton's method [11], Davidon's method [17] and the version of the former improved by Fletcher and Powell [18], and the compromise of the steepest descent method and Newton's method given by Marquardt [19]).

All of these procedures are iterative and require a set of primary estimates of the parameter values to start the iteration. In the discussion below we shall consider nonlinear models, which after suitable transformations can be treated by means of linear regression to compute the primary estimates required in the subsequent nonlinear regression. Many steady state kinetic equations are of this type, as they contain only one term in the numerator and can be linearized by inversion. However, more complex equations which are obtained with e.g. allosteric [20] and binding models [21] cannot be linearized.

In such cases primary estimates of the parameters can be chosen by trial and error [22] within constraints set by physical limits or other considerations. The procedure for discrimination and experiment design described below is applicable also for these models.

 North-Holland Publishing Company – Amsterdam

3. Discrimination based on a given set of data

3. 1. *Development of discrimination criteria*

Imagine a series of measured initial velocities, v_i, with corresponding sets of reactant (i.e. substrate, inhibitor, activator etc.) concentrations, $X_i = (X_{i1}, X_{i2}, \ldots, X_{ip})$

$$\mathbf{X} \overset{\frown}{} v \qquad (1)$$

where v is a vector $(v_1, v_2, \ldots, v_i, \ldots, v_n)$ and $\mathbf{X}$ is a matrix composed of n rows of reactant vectors, X_i. The aim of the investigation is to discriminate between possible models, η_j

$$v_i = \eta_j(X_i, K_j) \qquad (j = 1, 2, \ldots, k) \qquad (2)$$

where $K_j = (K_{0j}, K_{1j}, \ldots, K_{pj})$ is a vector of constants for the j-th model, and to compute estimates, $\theta_j = (\theta_{0j}, \theta_{1j}, \ldots, \theta_{pj})$, of the parameter values, K_j, of the best model.

We shall follow a computing scheme, which divides the nonlinear regression into two stages:

1) linear regression with transformed models to obtain primary estimates for
2) a nonlinear optimization.

Consider a hyperbolic equation (3)

$$v_i = \frac{K_0 X_{i1} X_{i2}}{X_{i1} X_{i2} + \sum_r K_r X_{ir}} \quad \begin{matrix} (i = 1, 2, \ldots, n) \\ (r = 1, 2) \end{matrix} \qquad (3)$$

and the linear equation system (4) obtained by inversion of (3)

$$\begin{bmatrix} \frac{1}{v_1} \\ \frac{1}{v_2} \\ \vdots \\ \frac{1}{v_n} \end{bmatrix} = \begin{bmatrix} 1 & \frac{1}{X_{11}} & \frac{1}{X_{12}} \\ 1 & & \vdots \\ \vdots & & \\ 1 & & \frac{1}{X_{n2}} \end{bmatrix} \begin{bmatrix} K_0^* \\ K_1^* \\ K_2^* \end{bmatrix}, \text{ or } a = \mathbf{B}\,K^* \qquad (4)$$

where K^* designates the vector of transformed constants.

As is well known from the theory of linear regression (cf. [23]), the least squares estimate, θ^*, of the parameter vector, K^*, is given by (5), where $\mathbf{B}^T$ is the transpose of $\mathbf{B}$,

$$\theta^* = (\mathbf{B}^T\mathbf{B})^{-1}\,\mathbf{B}^T\,a \qquad (5)$$

The parameter values obtained from (5) are used as primary estimates in the nonlinear regression procedure. The latter method gives convergence rapidly if the primary estimates are sufficiently accurate. It is convenient to write the computer programs so that a print-out of the vector of residuals and the residual sum of squares (6) is obtained

$$Q_j^2 = \sum_{i=1}^{n} (v_i - \eta_j(X_i, \theta_j))^2/(n-p-1) \qquad (6)$$

in addition to the parameter values and their standard errors.

Assume that two rival models, j and k, are considered and that programs are available to analyze experimental data according to these models. Furthermore, suppose that the true model is either j or k, making our aim to identify this model and refine the corresponding parameter values.

Now let us examine the computing stages and the results of the calculations in order to formulate criteria, which make possible the rejection of the incorrect model.

3. 1. 1. Criteria based on the success of regression

A) *Linear regression.* Eqn. (4) can be solved if the matrix $(\mathbf{B}^T\mathbf{B})$ can be inverted. Failure of inversion can be due to singularity of the matrix, which implies that the model is overdetermined and not acceptable. However, it has to be confirmed that the failure of inversion is not caused by numerical difficulties (ill-conditioning) [3].

B) *Nonlinear regression.* The solution of Eqn. (4), used as primary estimates of the parameters, is sufficiently accurate to give rapid convergence with most of the programs. If only one of the models gives convergence, the other one is rejected. (Several sets of primary estimates should be used to start the iteration before a model is regarded to be defective.)

3. 1. 2. Criteria based on examination of the results of regression.

C) *Parameter values.* Models giving unreasonable parameter values are rejected. The significance of the values obtained can be tested by the statistical t-test, in which the ratio of a parameter value to its corresponding standard error is compared with the proper table value. An estimated parameter value, θ_r, satisfying inequality (7) is considered redundant (cf. [9 and 22 a]).

$$\frac{\theta_r}{\text{s.e.}\,(\theta_r)} < t \tag{7}$$

Regression with unconstrained parameter values may give negative values in the absence of product inhibition, which is inconsistent with the theory of steady state kinetics. Negative values can be avoided by introduction of constraints. Minimization of Q^2, in this case, will normally result in a parameter value equal to one of the limit values [11], and further analysis will show redundancy of parameters.

D) *Residuals.* The residuals of model j, q_{ij} $(= v_i - \eta_j(X_i, \theta_j))$, should be examined as functions of the reactant or velocity values by suitable plots or Tukey statistics [24]. Models which do not approximate the requirements (8) and (9)

$$E(\sigma^2) = Q^2 \tag{8}$$

$$q_{ij} \sim N(0,\sigma^2) \tag{9}$$

are rejected.

E) *Residual sum of squares.* The residual sum of squares, Q^2, should be compared for the two models. If none of the first four criteria (A–D) allows elimination of one model, the zero hypothesis H_0: $Q_j^2 = Q_k^2$ should be examined by the F-test. If H_0 is rejected, the model with the smaller Q^2 should be chosen.

The goodness of fit in the finally accepted model can be examined by the F-test, using the quotient of the sum of squares about the mean, Q^2_{mean}, and sum of squares due to the regression, Q^2_{reg} [23].

$$R^2 = \frac{Q^2_{\text{reg}}}{Q^2_{\text{mean}}} = \frac{\sum_i (\eta(X_i, \theta) - \bar{v})^2}{\sum_i (v_i - \bar{v})^2} \tag{10}$$

R^2 should not be far from unity in a good regression. Furthermore, a test for lack of fit should be carried out to confirm the adequacy of the model chosen [23].

4. Development of experiment design for discrimination

The discrimination criteria described in the previous section are very useful if measured values are available. However, it is possible that the experiment has not been carried out at reactant concentrations giving the optimal conditions to solve the discrimination problem. In this case we have to find the experimental conditions, securing the optimal use of criteria (A–E).

The only criterion of (A–E) which is not based on deficiencies of models is (E). By this criterion a comparison of the Q^2-values given by the models is made, allowing retention of the model giving the significantly smallest Q^2-value. Thus experimental conditions should be elaborated to maximize Q^2 in the incorrect model, as suggested by Hunter and Reiner [25]. It is assumed that the Q^2-value of the correct model is not affected by the choice of reactant concentrations, whatever their values, whereas the incorrect model will exhibit a greater Q^2-value by experimentation in a particular region of reactant concentrations.

Consider the residuals of models j and k in the i-th experimental point

$$\begin{aligned} q_{ij} &= v_i - \eta_j(X_i, \theta_j) = \epsilon_i + \epsilon_{ij} + \epsilon_{ij_h} \\ q_{ik} &= v_i - \eta_k(X_i, \theta_k) = \epsilon_i + \epsilon_{ik} + \epsilon_{ik_h} \end{aligned} \tag{11}$$

where ϵ_i is the experimental error (which can be estimated from replicate measurements and which is assumed to follow $\epsilon_i \sim N(0,\sigma^2)$); ϵ_{ij} is the error caused by inadequacy of model j; and ϵ_{ij_h} is the error of the h-th nonlinear regression program. For the present purpose we regard all the regression programs as giving the same error ($\epsilon_{ij_h} \approx \epsilon_{ik_h}$).

Define a discrimination function g:

$$g(X_i, \theta_j, \theta_k) = |q_{ij} - q_{ik}| = |\epsilon_{ij} - \epsilon_{ik}| \tag{12}$$

Optimal conditions for discrimination between models j and k are obtained, if experiments are carried out at the point X_i in the reactant space, which maximizes the discrimination function g. This point can be determined by means of numerical or analytical differentiation of g, if some primary estimates of θ_j and θ_k are available. (Differentiation of g is feasible because the experimental points can be chosen sufficiently close to define g as a continuous function of X_i).

Now the necessary and sufficient conditions for a local maximum of g are (13) and (14):

$$\left(\frac{\partial g(X_i, \theta_j, \theta_k)}{\partial X_{ir}}\right)_{X'_i} = 0 \quad (r = 1, 2, \ldots, p) \quad (13)$$

where X'_i is the vector X_i lacking the r-th element, and

$$D_i = 0; \Delta_0, \Delta_1, \ldots, \Delta_p \text{ alternately positive and negative} \quad (14)$$

where D_i (the Hessian of g) is the determinant of second-order partial derivatives with respect to X_i and Δ_{p-r} ($r = 0, 1, \ldots, p$) are the determinants obtained by deleting the last r rows and columns of D_i ($\Delta_0 = 1$) [26]. The maximum may lie outside the available space of reactant concentrations, but the knowledge of the concentrations maximizing g nevertheless can be utilized in finding the most suitable region for experimentation.

In the case that constraints are imposed on the reactant concentrations, criterion (13) is enlarged by introduction of Lagrange multipliers for the side conditions [26]. In a similar manner constraints on the estimates of the parameter vector, K_j, affect the search of the minimum of Q^2 in the nonlinear optimization process.

5. An example of design for discrimination

A typical discrimination problem arises when a general equation is examined for redundant terms to select the best model, and we do not desire a crude model, describing the major features only, but a complete model reflecting even the minor effects. The g function, predicting the suitable region of experimentation, can help to reach this goal.

Consider an enzyme noncompetitively inhibited by I, where the steady state kinetics are described by model j

$$v_j = \frac{VA}{K_m(1 + I/K_1) + A(1 + I/K_2)} \quad (15)$$

where A and I are the concentrations of substrate and inhibitor, respectively; K_m is the Michaelis constant; K_1 and K_2 are inhibition constants, and V is the maximal velocity.

Suppose now that $K_1 \gg K_2$ which will cause considerable difficulties in detecting the slope effect [27] in a double reciprocal plot. A consequence of this condition ($K_1 \gg K_2$) is that also the uncompetitive inhibition pattern (16) will fit the data

$$v_k = \frac{V'A}{K'_m + A(1 + I/K'_2)} \quad (16)$$

where the constants are used in the same meaning as in eqn. (15).

To resolve this ambiguity by further experiments we introduce the function

$$g = g(A, I, V, V', K_m, K'_m, K_1, K_2, K'_2) = |v_j - v_k| \quad (17)$$

Because of the very small effect of K_1 on the other parameter estimates, the corresponding primed and unprimed constants differ only slightly. Therefore, as a first approximation, the corresponding pairs can be regarded to be equal making g:

$$g = g(A, I, V, K_m, K_1, K_2) \quad (18)$$

Examination of g shows that zero or infinite values of A and I all give $g = 0$. This indicates a finite maximum, which is obtained from:

$$\left(\frac{\partial g}{\partial A}\right)_I = 0 \text{ and } \left(\frac{\partial g}{\partial I}\right)_A = 0 \quad (19)$$

The analytical solutions are:

$$A = \frac{K_m}{1 + I/K_2} \sqrt{1 + I/K_1} \qquad (20)$$

$$I = \frac{K_m + A}{\sqrt{A/K_2 (K_m/K_1 + A/K_2)}} \qquad (21)$$

Eqns. (20) and (21) give two choices for the best value of A for discrimination with a given value of I, and vice versa. The point of intersection of (20) and (21) defines the absolute maximum of g, which can be obtained graphically by plotting A versus I according to (20) and (21). Replicate experiments should be carried out in the point of intersection to estimate the pure error, ϵ_p. If the residual sum of squares in one model significantly exceeds ϵ_p, this model is rejected. The problem now is to optimize the conditions for parameter estimation in the better model. As described by Box and Lucas [28], this is equivalent to finding the points which maximizes the determinant of the information matrix, $|\mathbf{B}^T \mathbf{B}|$ of the linearized model or the determinant $|\mathbf{F}_j^T \mathbf{F}_j|$ in the nonlinear model (where $\mathbf{F}_j$ is the matrix of first derivatives of $\eta_j(X_i, K_j)$ with respect to the elements of the parameter vector, K_j; $\mathbf{F}_j$ is evaluated by means of the current estimate, θ_j, of K_j).

6. Suggested procedure of experimentation

Let us briefly summarize the procedure proposed in the present paper.

a) Experiments should be made in a large region of reactant concentrations (to cause failure of the incorrect model, fitting only in a limited region) and the parameter values computed for rival models. b) Using these parameter values, the maximum of the discrimination function (g) for pairs of models (not failing in the earlier examination) can be evaluated. Experiments should then be carried out to maximize g and to obtain Q^2. The Q^2-values should be examined according to criterion (E) (sect. 3.1.2.), and as soon as the difference between these values becomes significant the best model can be chosen. c) The remaining problem is to obtain the best estimates of the parameter values, which requires maximizing the determinant $|\mathbf{F}^T \mathbf{F}|$. Evidently, it may be possible to start at point b) if primary estimates are available from literature or other sources.

References

[1] D. Garfinkel, L. Garfinkel, M. Pring, S.B. Green and B. Chance, Ann. Rev. Biochem. 39 (1970) 473.
[2] G. Johansen and R. Lumry, Compt. Rev. Trav. Lab. Carlsberg 32 (1961) 185.
[3] M. Berman, E. Shahn and M.F. Weiss, Biophys. J. 2 (1962) 275.
[4] C.I. Bliss and A.T. James, Biometrics 22 (1966) 573.
[5] R.A. Dammkoehler, J. Biol. Chem. 241 (1966) 1955.
[6] K.R. Hanson, R. Ling and E. Havir, Biochem. Biophys. Res. Commun. 29 (1967) 194.
[7] J. Kowalik and J.F. Morrison, Math. Biosci. 2 (1968) 57.
[8] W.W. Cleland, Nature 198 (1963) 463.
[9] W.W. Cleland, Advan. Enzymol. 29 (1967) 1.
[10] J.G. Reich, G. Wangermann, M. Falck and K. Rodhe, European J. Biochem. 26 (1972) 368.
[11] W.H. Swann, FEBS Letters 2 Suppl. (1969) 39.
[12] J. Kowalik and M.R. Osborne, Methods for Unconstrained Optimization Problems (American Elsevier Publishing Co., New York, 1968).
[13] R. Hooke and T.A. Jeeves, Direct search solution of numerical and statistical problems, J.A.C.M. 8 (1961) 212.
[14] H.H. Rosenbrock, The Computer J. 3 (1960) 175.
[15] M.J.D. Powell, The Computer J. 7 (1964) 155.
[16] J.A. Nelder and R. Mead, The Computer J. 7 (1965) 308.
[17] W.C. Davidon, A.E.C. Reasearch and Development Report ANL-5990 (Rev.) (1959).
[18] R. Fletcher and M.J.D. Powell, The Computer J. 6 (1963) 163.
[19] D.W. Marquardt, J. Soc. Ind. Appl. Math. 11 (1963) 431.
[20] J. Monod, J. Wyman and J.-P. Changeux, J. Mol. Biol. 12 (1965) 88.
[21] E.E. Baulieu and J.-P. Raynaud, European J. Biochem. 13 (1970) 293.
[22] A. Cornish-Bowden and D.E. Koshland, Jr., Biochemistry 9 (1970) 3325.
[22a] G. Pettersson and I. Pettersson, Acta Chem. Scand. 24 (1970) 1275.
[23] N.R. Draper and H. Smith, Applied Regression Analysis (John Wiley & Sons Inc., NewYork, N.Y., 1966).
[24] F.J. Anscombe and J.W. Tukey, Technometrics 5 (1963) 141.
[25] W.G. Hunter and A.M. Reiner, Technometrics 7 (1965) 307.
[26] T.M. Apostol, Mathematical Analysis (Addison-Wesley Publ. Co., Reading, Mass., 1957).
[27] W.W. Cleland, Biochim. Biophys. Acta 67 (1963) 173.
[28] G.E.P. Box and H.L. Lucas, Biometrika 46 (1959) 77.