REVIEW LETTER

PARAMETER ESTIMATION AND ENZYME KINETIC MODELS

Jens G. REICH
Forschungszentrum für Molekularbiologie und Medizin,
Deutsche Akademie der Wissenschaften, Berlin-Buch, GDR

Received 6 July 1970

1. Introduction

Biochemists have always sought to use kinetic and other measurements to derive a description of enzyme catalysis at the molecular level. However accurate the data and sophisticated the experimental techniques, all efforts may be futile unless they are coupled with appropriate modeling and simulation strategies to construct a mathematical model of the object of study.

An adequate model might be a simple formula such as the Michaelis equation or alternatively it may be an extremely complex set of equations. This depends entirely on the available data. The survey presented here aims to show that the mathematical techniques which lead to the kinetic model are by no means as trivial as they appear to be from the study of simple textbook cases. It will be seen that in many recent papers the difficulties have been investigated, but so far satisfactory solutions for the more complicated cases have not been described.

2. Nomenclature

To avoid confusion with the non-mathematical use of the term, a *model* should be taken in this letter to be a mathematical formula relating certain input variables (substrate and modifier concentrations) to observable output variables (reaction rate, relative saturation etc.). A model has a mathematical structure (for example, it produces hyperbolic curves), and has *parameters* (kinetic constants) which are to be estimated from the experimental data. As a rule, statistical fluctuations complicate these relations either as errors of the input or output values (the former are usually better controlled) or as genuine statistical variations of the parameter values which are influenced by the experimental conditions (temperature, homogeneity of enzyme preparation, etc.). Such fluctuations are the so-called *stochastic* components of the model.

The construction of a model has usually three aspects: design, discrimination and parameter estimation. The experimental *design* concerns the choice of an optimal set of input concentrations to get the maximum of information from the experiment. *Discrimination* means to distinguish between different candidate models (for instance, between a hyperbola and a sigmoid curve), and *parameter estimation* aims to define the correct parameter values subject to the condition that a specified model is acceptable for the enzyme in question.

Traditionally, the design problem is considered as a matter of experience rather than of mathematical skill. The classical solution for the remaining problems, i.e. model choice and parameter estimation, has been *graphical transformation*: The data are plotted on transformed coordinates. If a model is acceptable, a straight line or some other salient picture is obtained, and the parameter values can be estimated from intercepts, slopes and secondary plots of the transformed data. It is important to note that in more involved cases such a parameter estimation is *sequential* by nature. A primary investigation gives some parameter values which are inserted as if they were established constants into the graphs which follow.

3. An ambiguous model

Before more recent estimation techniques are reviewed in detail, I present, as a precautionary warning, an illustration of the uncritical use of classical transformations. Its purpose is to show how practical difficulties necessitate the development of more sophisticated methods.

Sigmoid kinetics are currently in fashion because of their importance in regulation problems. An almost universally accepted measure for the cooperativity (steepness of transient) of a kinetic curve is the Hill coefficient n, the sigmoid curve being described by the Hill approximation:

$$v = V_{max} \times \frac{S^n}{K + S^n} ,$$

where S and v are substrate concentration and enzyme activity, respectively, and K is an apparent affinity constant which usually has no physical meaning and serves for fitting purposes only. The left hand part of fig. 1 shows two special cases of this formula (one with $n = 3.5$, $V_{max} = 1.02$ and $K = 273$, and the other one with $n = 4.8$, $V_{max} = 0.89$ and $K = 1469$). The curves are nearly identical, and it would be very ambitious to try to distinguish between them by an experiment. Nevertheless, the apparent cooperativity, as expressed by the Hill n, is clearly distinct, and this fact is reflected in the conventional linear Hill plot (right hand part of fig. 1). This paradox is explained by the V_{max}-values which differ by 10%, a difference which is barely significant when it is remembered that V_{max} is usually obtained by a rather uncertain extrapolation to infinite S. Thus, the degree of cooperativity is found to be statistically dependent on the choice of a parameter which by itself has no relevance for the steepness of the curve. Moreover, this conclusion has been obscured by the uncritical use of a linear plot which involves the difference between the stochastic variables ($V_{max} - v$), and is therefore itself mathematically suspect.

The situation just described is a typical example of an ambiguous model. A model is said to be *ambiguous* or *non-unique* when the information content of the data does not suffice to determine the model parameters uniquely. A characteristic symptom is that completely different parameter sets may more or less satisfactorily fit the data (as sketched in fig. 2b), but as soon as one or more of the parameters are assumed to be established (such as V_{max} in the Hill plot in fig. 1), the rest of them seem to be well determined by the data. Thus, the usual sequential estimation of parameters is liable to give an impression of certainty when in fact no certainty exists.

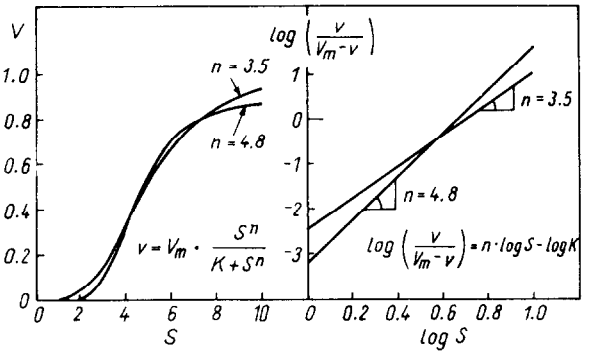

Fig. 1. An ambiguous model. The left hand part shows two numerical examples of the Hill equation, one with $n = 3.5$ ($V_m = 1.02$; $K = 273$), and the other one with $n = 4.8$ ($V_m = 0.89$; $K = 1469$). Note that the two curves are nearly identical despite the marked difference in their cooperativity (as expressed by n). The right-hand part shows Hill plots of both curves whose slopes are clearly distinct, in contrast to the direct plot (see text).

4. An incompatible model

Let the experimental data be slightly sigmoid, and the experimenter tries to adapt a hyperbola to the data (as did generations of biochemists before the importance of "allosteric" kinetics was recognized). However close the approximation by the hyperbola might be, there will always remain some non-random fluctuation around the theoretical curve (see fig. 2c): One group of the points lie above, another below the fitted curve. This is one reliable symptom of an *incompatible* or *inconsistent* model. The information content (or variability) of the data is greater than the model can account for. In contrast to the ambiguous model, the incompatible model is often recognized in transformation graphs. A Lineweaver-Burk plot shows a distinct deviation from the expected linear curve.

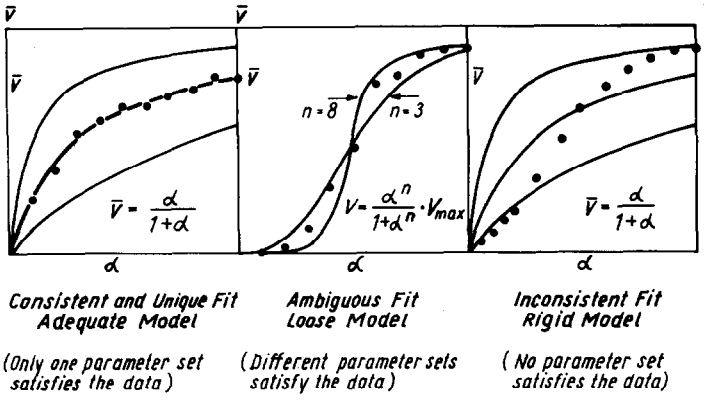

Fig. 2. Consistency and uniqueness of models. The examples are artificially generated. The respective models are given as explicit formulae; $\bar{v}$ is normalized velocity, α is normalized substrate concentration. Points refer to simulated experimental measurements and solid lines to specified models with different parameter sets.

5. Statistical investigation of the adequacy of a model

So far I have given only graphical illustrations for inappropriate kinetic models. I hope to have convinced the reader that a more sophisticated treatment is desirable in all but the most simple cases, because inspection of graphs is not a very satisfactory way of investigating a mathematical fact. The following sections describe possible mathematical expressions for the adequacy of models and parameter values.

6. Best-fit criteria

When a formula is tentatively applied to the data, one first needs a criterion which defines the parameter values providing the best prediction of the data. It is clear that such criteria make use of the difference between the predicted value η and the observed value y. The better the fit, the smaller are the differences. The four candidates for such a criterion differ in the weight of the individual differences:

(1) The least-square or maximum-likelihood principle assumes that the best parameter fit minimizes the sum of the squared deviations between all predicted and observed values:

$$\sum_{i=1}^{n} (\eta_i - y_i)^2 = \text{Min!}$$

(2) The minimum-χ^2-principle [1, 2] minimizes one of the following two expressions which are asymptotically identical with certain assumptions:

$$\sum_i \frac{(\eta_i - y_i)^2}{y_i} \quad \text{or} \quad \sum_i \frac{(\eta_i - y_i)^2}{\eta_i} = \text{Min!}$$

(3) A third principle would be to minimize the relative deviations:

$$\sum_i \left(\frac{\eta_i - y_i}{y_i}\right)^2 \quad \text{or} \quad \sum_i \left(\frac{\eta_i - y_i}{\eta_i}\right)^2 = \text{Min!}$$

(4) Davies [3] has discussed a fourth expression to be minimized: the sum of the absolute deviations between fit and measurements:

$$\sum_i |\eta_i - y_i| = \text{Min!}$$

In many cases these criteria lead to very similar results. Most authors prefer the maximum-likelihood criterion [4–6]. Their argument is based on statistical considerations which are difficult to explain in a few sentences. Their basic assumption is that the correct prediction is biassed by a small experimental error which is (i) normally distributed; (ii) has zero mean; (iii) whose variance does not depend on the size of η. Then, by mathematical derivation, principle 1 turns out to be the best.

My experience suggests that this decision is not a matter of mathematical deduction, but of biochemical

reasoning. It is by no means clear why a hypothetical "epsilon" of the recording instrument should be made the scapegoat for all undesired fluctuation in the data. It is very probable that the parameter population rather than the final measurement introduces most of the statistical variation. To give an example: If in a test tube the enzyme population cannot be controlled to absolute homogeneity, the parameters, say K_m or the allosteric constant L, which enter non-linearly into the rate equation, are much more liable to give rise to variation than a well controlled spectrometer. In such a case, all the afore-mentioned statistical assumptions are definitely wrong. A careful practical determination of the very nature of the fluctuation is a better guide for a suitable choice among the criteria. When the percentage error of each measurement is constant (as with some colorimetric methods), take principle 3; when the absolute error is identical, irrespective of the value of the measurement (as with some spectrometers), take principle 1; when the errors lie somewhere between both these situations, take principle 2; when outliers are expected, principle 4 might be the most suitable, because squaring of the large deviations is avoided. At any rate, exact mathematical considerations are useless unless the stochastic structure of the particular model has been clarified.

7. Uniqueness criterion

As Box [7] has pointed out, uniqueness should be tested for by constructing *all* sets (a whole realm) of parameter values which give a reasonable fit within a confidence limit. This is a difficult task with highly non-linear models, although the linearization error may be determined [8, 9]. The actual evaluations amount to arithmetically rather ill-conditioned problems (an eigensystem transformation or, at least, an inversion of the so-called information matrix [7]), requiring access to fast computers. It should also be pointed out that such studies do not by themselves reveal the reason for the non-definiteness of a model which may be either "bad" data with too much fluctuation or an unfavourable mathematical model structure. The first case is self-evident; the second is exemplified in the context of fig. 1, where the formula itself produced an interdependency between V_{max} and n which is present even with error-free data. In summary, the problem of uniqueness is far from being formalized.

8. Consistency criterion

Fig. 2c illustrates how inconsistent models are recognized by systematic trends in the residual deviations $\eta_i - y_i$ between model prediction and observed values. A reasonable approach to quantify this, which was introduced by Anscombe and Tukey [10, 11] and adopted by Haarhoff [12] is therefore to draw regression lines through these residuals. If these curves are significantly different from the zero coordinate, the model is suspect. A second approach, also exemplified by Haarhoff, seems to me to be less promising. It is based on the following argument: When the model is appropriate, the variance of the residuals should be in the range of the experimental sampling variance. The validity of the derivation, however, depends on the three assertions on error structure which have been stated and criticized in the section "best-fit criteria". Moreover, it is usually not easy to measure the sampling variance precisely enough.

A rather different approach has been devised by Gardiner and Ottaway [13]. They expected the Michaelis formula to be valid, so that an Eadie transformation of the data ought be a straight line. To test this, they instead fit a parabola to the transformed data. If the curvature is insignificant, the original linear model may be accepted, otherwise rejected. It should be pointed out that the variance analysis can only be applied when the sum of squared deviations of both models (quadratic as well as linear hyperbola) is known. Two computer runs are therefore always necessary (which has not been stated clearly by these authors). For details of this, see Mandel's comprehensive account [14]. In practice, Gardiner's method is very reliable, but restricted to a particular class of alternative hypotheses.

Lasch [15], in a recent paper, has recommended simply to take the standard deviation of the resulting parameters (obtained by Cleland's program [6]) as a measure of consistency; the higher the deviations, the more suspect is the model. We have experience with this criterion tested on models with up to 10 parameters (allosteric systems), and can confirm that it is reliable and sensitive. It cannot distinguish, however, between inconsistent and ambiguous models.

9. Modeling strategies in enzyme kinetics: recent trends

So far criteria for good models have been discussed, but the technical question of how it can be obtained was omitted. The progress of the last years has shown that fast computers become indispensable for this work. But again, the methods are by no means perfect and require further improvement. This discussion is directed towards design, discrimination and estimation, as introduced earlier.

10. Design of experiments

A number of papers [16–18] treats design optimization in non-linear models. All of them follow the general line of an ingenious idea originated by Box [7]. Its description without algebraic formalism would take too much space in this survey. Furthermore, the techniques, in their present state of development, are so difficult to apply to enzyme kinetics that this brief mention must suffice. But it is evident that with more and more expensive experiments (rare substances, unstable enzymes etc.) the question of design optimization becomes more urgent. The development of computer simulation programmes as pioneered by Garfinkel, Hess and E.M.Chance [19–21] is also a step in the direction of design improvement and should be complemented by automatic strategies making use of the results of Box and his coworkers from the industrial planning field.

11. Model discrimination

There is no need to stress the importance of the distinction between different mechanistic models to describe an enzyme at the molecular level. Recent publications seem to shift the weight from intuitive to numerical concepts of model choice [18, 22–26]. The danger of false conclusions with the semiquantitative techniques (of the type: Model A is correct when plot B gives straight lines intersecting at a point) is obvious. But the problem of the "blind" numerical concepts is their instability. Even small biasses of the data (say, an excess substrate inhibition which is easily recognized on a graph) can wreak havoc in the building of goodness-of-fit criteira. A further drawback is that such expressions always require knowledge of the best parameter set under the asserted validity of a model. This leads to endless computer iterations when the fit is necessarily poor, i.e. when the model just tried is inappropriate.

12. Parameter estimation

The pioneer in the field of enzyme parameter estimation by computer techniques is Cleland [27]. The Bethesda programme by Berman [28] has been devised for compartment models, but may also be used for enzymological studies. Other authors have studied more specialized algorithms [29, 30]. All these programmes are more or less well suited for a certain model class of isosteric kinetics, and get into difficulties when highly non-linear models are encountered, especially when the validity of a model is not known in advance. This seems to exclude allosteric kinetics in all its variants from automatic parameter estimation. Swann [31] has described, in a detailed review, the algebraic difficulties and possible solution paths for the problem of nonlinear parameter "optimization" as he would call it (this term stems from chemical engineering).

With regard to traditional graphical estimation techniques, which many authors have tried to improve [32–36], it seems reasonable to demand of everyone who devises new or old plots to submit clear evidence (analytical or by Monte-Carlo methods *) of the statistical consequences of his procedure. Good examples are Wilkinson's and Lumry's papers [37, 38]. It is probable that with such critical use plots will remain a versatile tool in the hands of the experimenter for some time to come.

13. Envoi

In the present state of the science, the conclusion

* Monte-Carlo experiments are theoretical studies (mostly on computers) in which artificially generated random numbers are superimposed on the model. The effect of such stochastic fluctuations on the data and on plots may be directly evaluated from such simulation runs.

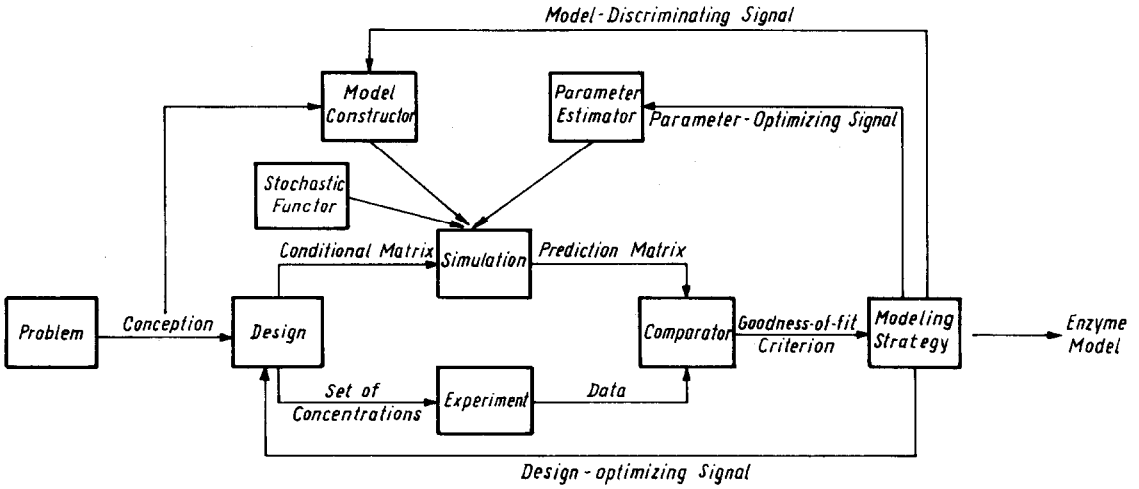

Fig. 3. Computer-oriented modeling and simulation system for kinetic experiments. The operators (boxes) act on input signals providing output signals; they may be computer subroutines controlled by the experimenter. The formulation of the problem determines a conception which defines the outline of the design. The design operator gives a signal which is a set of test concentrations, in abstract form represented as a matrix of conditions. The experiment produces data, while the simulation, fed by model and parameter functors gives abstract, simulated data ("prediction matrix"). A comparator gives a goodness-of-fit signal after comparison of simulation and experiment (for example, the sum of squared deviations). The model strategy ("monitor") evaluates numerical indices for better fit, better parameter values or better models to be fed back into the system, or, finally, decides that a model is either appropriate or inappropriate.

that kinetic modeling is a rather intuitive and unsystematic activity is hardly avoidable. Nevertheless, it seems probable that the difficulties and inconsistencies outlined in the preceding paragraphs can be overcome. Our goal should be a general computer-controlled automatic programme for kinetic experiments, where the measurement itself becomes part of the modeling instead of providing a constant flow of raw data. I have outlined such a programme in fig. 3; its supervisor is the monitor box "modeling strategy" which decides whether an investigation is to be iterated or finished. I am convinced that automatic experiments can greatly enhance our capacity of understanding and operating enzyme systems even at high levels of metabolic integration.

Acknowledgment

I am very much indebted to Dr. Bastien D. Gomperts, London, for a detailed and critical discussion of an earlier version of this paper.

References

[1] M.G.Kendall and A.Stuart, The Advance Theory of Statistics, vol. II (Griffin, London, 1961).
[2] M.Halperin and N.Mantel, J. Am. Statist. Assoc. 58 (1963) 611.
[3] M.Davies, J.Roy. Statist. Soc. B29 (1967) 101.
[4] M.Berman, M.F.Weiss and E.Shahn, Biophys. J. 2 (1962) 289.
[5] J.Mandel, The Statistical Analysis of Experimental Data (Interscience, New York, 1964).
[6] W.W.Cleland, Advan. Enzymol. 29 (1967) 1.
[7] G.E.P.Box, Ann. N.Y. Acad. Sci. 86 (1960) 792.
[8] E.M.L.Beale, J. Roy. Statist. Soc. B22 (1960) 41.
[9] I.Guttman and D.A.Meeter, Technometrics 7 (1965) 623.
[10] F.J.Anscombe and J.W.Tukey, Technometrics 5 (1963) 141.
[11] F.J.Anscombe, Technometrics 2 (1960) 123.
[12] K.N.Haarhoff, J. Theor. Biol. 22 (1969) 117.
[13] W.R.Gardiner and J.H.Ottaway, FEBS Letters Suppl. 2 (1969) 34.
[14] see reference 5.
[15] J.Lasch, Acta Biol. Med. Germ. 23 (1969) 747.
[16] G.E.P.Box and H.L.Lucas, Biometrika 46 (1959) 77.
[17] G.E.P.Box and G.S.Watson, Biometrika 49 (1962) 93.
[18] G.E.P.Box and W.G.Hunter, Technometrics 7 (1965) 123.

[19] D.Garfinkel and B.Hess, J. Biol. Chem. 239 (1964) 971.
[20] D.Garfinkel, Computers Biomed. Res. 2 (1968) 92.
[21] E.M.Chance, Computers Biomed. Res. 2 (1968) 28.
[22] A.C.Atkinson, Biometrika 56 (1969) 337.
[23] G.E.P.Box and W.J.Hill, Technometrics 9 (1967) 57.
[24] D.R.Cox, Proc. 4th Berkeley Symp. Prob. Math. Statist. 1 (1961) 105.
[25] D.R.Cox, J.Roy.Statist. Soc. B24 (1962) 406.
[26] W.G.Hunter and A.M.Reiner, Technometrics 7 (1965) 307.
[27] W.W.Cleland, Nature 198 (1963) 463.
[28] K.R.Hanson, E.A.Havin and R.Ling, Biochem. Biophys. Res. Commun. 29 (1967) 194.
[29] J.Kowalik and J.F.Morrison, Math. Biosciences 2 (1968) 57.
[30] M.Berman, Compartmental Analaysis in Kinetics, in: Computers in Biomedical Research, vol. II, eds. R.Stacy and B.Waxman (Academic Press, New York, 1965), ch. VII.
[31] W.H.Swann, FEBS Letters, Suppl. 2 (1969) 39.
[32] J.P.Changeux, J.C.Gerhart and H.K.Schachman, Biochemistry 7 (1968) 531.
[33] L.W.Nichol, G.D.Smith and D.J.Winzor, Nature 222 (1969) 174.
[34] P.P.Tchukniyski, Molekul. Biol. USSR 3 (1969) fasc. 2 197.
[35] A.Horn and H.Börnig, FEBS Letters 3 (1969) 325.
[36] K.Dalziel, Acta Chem. Scand. 11 (1957) 1706.
[37] G.N.Wilkinson, Biochem. J. 80 (1961) 324.
[38] G.Johansen and R.Lumry, Compt. Rend. Trav. Lab. Carlsberg 32 (1961) 185.