Perspectives in Science (2014) 1, 121–125

Available online at www.sciencedirect.com

ScienceDirect

www.elsevier.com/locate/pisc

REVIEW

Analysis and interpretation of enzyme kinetic data ☆

Athel Cornish-Bowden

CNRS-BIP, 31 Chemin Joseph-Aiguier, B.P. 71, 13402 Marseille Cedex 20, France

Received 10 August 2011; accepted 4 November 2013; Available online 27 March 2014

KEYWORDS
Experimental design;
Model discrimination;
Parameter estimation;
Least squares

Abstract
Analysis of enzyme kinetic data to obtain valid information requires attention to two details that are often given less attention than they need. The first is an experimental design that ensures that the variables treated as independent are truly independent, that different interpretations can be distinguished, and that parameter values can be estimated. The second is that authors should be aware of the statistical assumptions that are implicit in the fitting programs that they use, whether commercial or not.

Contents

☆This article is part of a special issue entitled "Reporting Enzymology Data—STRENDA Recommendations and Beyond". Copyright by Beilstein-Institut.

E-mail address: acornish@imm.cnrs.fr

http://dx.doi.org/10.1016/j.pisc.2014.02.010

Introduction

The title of this chapter suggests a textbook account of enzyme kinetics, but that would not be appropriate here. Instead I shall concentrate on three aspects closer to the aims of STRENDA. How should kinetic experiments be designed if they are to yield results that allow analysis? How should kinetic parameters be deduced from kinetic measurements? What information needs to be provided in reporting the results of a kinetic experiment in such a way that they can be confirmed by other workers? Several textbooks are available for readers who need a more pedagogical account (Fersht, 1999; Copeland, 2000; Bisswanger, 2002; Marangoni, 2002; Cook and Cleland, 2007; Alberty, 2011; Cornish-Bowden, 2012).

Experimental design

The principles of experimental design are sufficiently obvious that they ought not to require discussion. They are often violated in published work, however, so apparently they are not perceived as obvious. The essential point is that an experiment should be capable of supplying the information that the experimenter is seeking to extract. The necessary design, therefore, must depend on the context in which the experiment is being used. If the aim is to obtain kinetic parameters to be used for elucidating an enzyme mechanism, the conditions need to be varied in ranges in which the results vary with the parameter of interest. If the aim is to understand the physiological role of an enzyme it needs to be studied in conditions that do not depart more than necessary from physiological conditions. All this is simply common sense, but it is useful to consider it in a little more detail.

Independent variables must be independent

This is a point that arises when there are two or more independent variables—two different substrate concentrations, for example, or a substrate and an inhibitor concentration. Put in words it is indeed obvious: if two variables are not independent then they are not independent! However, in practice it may not be obvious without an understanding of what independence means. This is easy to define for a linear regression model: it is sufficient to require that two independent variables x_1 and x_2 must not satisfy any linear equation $x_2 = a + bx_1$, where a and b are any constants. It is also easy to illustrate the consequences of violating this requirement in a linear regression. Virtually none of the equations considered in enzyme kinetics lead to linear models if properly analysed,[1] but in practice it is not difficult to ensure that the independent variables are indeed independent even in a non-linear regression: in essence, it means that knowledge of the values of one independent variable must not allow the values of another to be calculated. In the simplest case, concentrations must not be varied in constant ratio, or with a constant sum.

[1]Linear transformations of equations like the Michaelis–Menten equation exist, of course, but if properly weighted these do not make the model itself linear.

This does not of course exclude the possibility that one may want to remove the independence between two or more variables. For example, the method of Yagi and Ozawa (1960) for analysing multiple inhibition involves using linear combinations of the concentrations of two or more inhibitors, and that proposed much more recently by Cortés et al. (2001) for assessing whether two competing substrates bind at the same site involves linear combinations of the two substrate concentrations. In these sorts of experiments one is deliberately suppressing differences between the effects of the two variables in order to shine more light on some effect of the two together, and as long as this is understood there is no objection to the use of linear combinations of concentrations.

In an ideal world one retains as many independent variables as may be relevant to the behaviour one is seeking to explain, but in practice that advice may be difficult to follow. For example, in studying an enzyme with activity dependent on $MgATP^{2-}$ it is possible to vary the total concentrations of ATP, $MgCl_2$ and the pH in such a way that the concentrations of all relevant ions and molecules vary independently, so that effects due to the different ones can be separated. It is much easier, however, to follow a design in which the total $MgCl_2$ concentration is kept at a constant level (typically 2 mM or 5 mM) in excess over the total ATP concentration (Storer and Cornish-Bowden, 1974). This ensures that a high and almost constant proportion of ATP exists as MgATP, and that the concentration of ATP^{4-} is low enough not to interfere with the analysis. On the other hand it makes it difficult or impossible to isolate effects due to ATP^{4-}. In an instructive example, Mannervik (1981) examined four designs for varying the concentrations of glutathione and methylgloxal for distinguishing between models for glyoxalase I. He showed that maintaining one or other constant, or varying them in constant relation to one another, showed poor discriminatory power, but varying them independently was very powerful.

Experimental design for model discrimination

In the preceding discussion there has been an implied assumption that the purpose of data analysis is model discrimination rather than parameter estimation as such. In a study to establish an enzyme mechanism this is certainly true at some level. For distinguishing between two possible explanations of observed behaviour it hardly matters whether the true value of a parameter such as a catalytic constant is $100\,s^{-1}$ or $1000\,s^{-1}$, though it may certainly be important for understanding the physiological role of an enzyme, or for comparing the properties of enzymes from different sources. Within the mechanistic context it becomes important for understanding the variation of the parameter in question with the conditions, such as the pH or the concentration of an inhibitor. In practice, therefore, one cannot avoid designing for effective parameter estimation regardless of the ultimate aim, but in any case few experimenters would want to do that.

Lack of fit and pure error

Textbooks of regression such as that of Draper and Smith (1981) typically distinguish between *lack of fit*, the

deviations from calculated behaviour that result from fitting the wrong model, and *pure error*, the deviations from calculated behaviour that are independent of the model fitted. Although both sources of error normally contribute to the sum of squares of deviations from a model, they can be separated: inconsistencies between replicate observations are unaffected by the choice of model and thus allow calculation of how much of the total sum of squares is due to pure error, and from this one can calculate the contribution of lack of fit. My purpose here is not to describe how to do that, but to emphasize that any experimental design involves a trade-off between lack of fit and pure error. For any given total number of measurements, the more information one obtains about pure error (the more observations made in replicate) the less information one can obtain about lack of fit (the fewer design points). If model discrimination is the principal objective, as assumed in the preceding section, it is sensible to have many design points, covering a wide range of relevant conditions, but have enough replicate observations to have at least some idea of the pure error. In fact, a measure of pure error is necessary even if one is looking at just one model (rather than comparing two or more), because comparison of the contributions of lack of fit and pure error to the sum of squares allows an assessment of whether the fitted equation is reasonable.

Experimental design for parameter estimation

It is possible to design an experiment to yield the maximum possible information about parameter values at the expense of all information about model discrimination, and Duggleby (1979) has explained how to do that. One must assume that the correct equation to be fitted is known without any possibility of error, and then choose exactly the same number of design points as there are parameters to be estimated, the exact design points (and the number of replicates at each one) being calculated to be optimal. For mechanistic studies this approach is clearly not a good idea, but even for other purposes it seems unwise, as not only does it eliminate any possibility of knowing whether the right equation has been fitted, but it also eliminates any information about failure of the equation. Even if the parameters are required only for predicting the behaviour of an enzyme in different conditions it is a risky approach, because it takes no account of the possibility that the assumed equation is insufficiently accurate if it is applied to conditions different from the design points.

A more realistic general approach is to follow similar principles of design to those used for model discrimination, taking account of which parts of the design space contribute most to the estimate of each parameter of interest. In some cases these are obvious: estimating the catalytic constant k_{cat} requires some observations at high substrate concentrations; estimating a competitive inhibition constant K_{ic} requires observations at low substrate concentrations, because a competitive inhibitor is most effective at low substrate concentrations; conversely, estimating an uncompetitive inhibition constant K_{iu} requires observations at high substrate concentrations. In other cases the requirements are less obvious: the value of the Michaelis constant K_m depends both on k_{cat} and on the specificity constant k_{cat}/K_m, and needs a design that defines both of these precisely. However, although k_{cat}/K_m is sensitive to variations in the rate at very low substrate concentrations, it does not necessarily require the concentrations to extend as low as possible. This is because if the observed rates have uniform standard deviations the requirement for the rate to be zero in the absence of substrate can be treated as a fixed point that gives as much information about the value of k_{cat}/K_m as measurements at less than $0.4K_m$ (Endrenyi, 1981).

If taken too literally Endrenyi's analysis suggests that there is nothing to be gained by extending the range of substrate concentrations below $0.4K_m$. However, there are in fact two reasons not to take it too literally. First, it will rarely be certain that the observed rates have uniform standard deviation, and if, for example, they have uniform coefficient of variation (which may be more likely: see the discussion below of the assumptions in least squares), the ideal lower limit is zero, not $0.4K_m$ (Endrenyi, 1981). Secondly, it will often not be safe to assume that there is no "blank rate", i.e. that the rate is zero in the absence of substrate, and measurements at very low substrate concentrations will provide an indication of this.

Other experimental conditions

An appropriate design of an experiment for kinetic characterization of an enzyme involves more than just choosing appropriate substrate and effector concentrations. Even if no pH or temperature dependence studies as such are being made, it is still necessary to choose appropriate pH, temperature, ionic strength, etc., and to choose an appropriate buffer. If the results are intended to have physiological meaning (including use for metabolic modelling, these conditions should be as close to physiological as possible, but for mechanistic studies they can be varied to supply the particular kind of information sought. In either case it is important to use a buffer appropriate for the pH to be used, with a pK_a no more than 1 pH unit from the desired pH, and preferably less, so an acetate buffer ($pK_a=4.64$) would be ineffective as a buffer at pH 7, for example. One must also take care that the buffer does not react with the enzyme or interfere with the assay: for example, glycylglycine is typically inappropriate for use with peptidases, and Hepes and numerous other buffers interfere with the Lowry method of protein analysis. When it is desirable to simplify the mixture as much as possible, the pHstat allows the pH to be maintained constant without any chemical buffer.

Estimating enzyme kinetic parameters

Assumptions in least-squares analysis

It is no more realistic in 2014 to suggest that biochemists should write their own computer programs to analyse their kinetic data than it would be to suggest that they should prepare their own ATP. So far as molecular biology is concerned it is clear that we live in an age of kits, and if that is less true of enzymology than of molecular biology it is mainly because enzymology is a less fashionable subject for which manufacturers do not find it worth their while to develop kits on the same scale. Nonetheless, parameter estimation has become almost entirely a matter of using

commercial programs as if they were black boxes, without any idea of how they work or what they are assuming about the input data, in other words using them as kits. This may be an advance on the days when nearly all kinetic parameters were derived from hand-drawn double-reciprocal plots that were interpreted by eye, but it is not as large an advance as commonly assumed. The question is akin to the use of buffers to control the pH: on the one hand it may be sensible to leave the preparation of the buffer to a technician, but one still has to know what buffer is appropriate for a particular pH, and how one can check whether it does in fact supply the intended pH.

It is important to realize also that most users use a commercial data-processing packages with their default options. So even if they offer the possibility of selecting a more appropriate weighting scheme than the default that is of little value if it is used straight out of the box. The popular program SigmaPlot (version 11.2) can fit Michaelis-Menten data very easily, but if used in its default state it incorporates assumptions that

(1) The errors in the observed rates are subject to a normal (Gaussian) distribution.
(2) The substrate concentrations are known exactly.
(3) All of the rates have the same standard deviation, which is not the same as assuming that they all have the same coefficient of variation (standard error expressed as a percent).
(4) The errors are uncorrelated, which means that the magnitude of the error in one rate implies nothing at all about the magnitude the error in any other.
(5) There is no systematic error: the Michaelis-Menten equation is the right equation to fit.

Extremely few studies have been made to check whether any of these assumptions are likely to be true, and those studies are either old (Storer et al., 1975; Askelöf et al., 1976) or very old (Lineweaver et al., 1934), and thus tell us rather little about error behaviour in modern conditions. The last assumption is very important, but it is also the easiest to check, for example with the use of residual plots.

Tukey and McLauglin (1963) suggested many years ago that the "normal" distribution is actually so rare that it might be better be called the "pathological" distribution, going on to say that "the typical distribution of errors and fluctuations has a shape whose tails are longer than that of a Gaussian distribution". In practice deviations from the normal distribution severe enough to produce substantial errors in estimated parameters are likely to be obvious in residual plots. For example, a clear outlier is easily recognized in a residual plot: once recognized, a careful experimenter must assess whether it reflects an unexpected failure of the assumed model, and undertake additional experiments to find out, or whether it reflects a mistake in carrying out the experiment, such as use of the wrong stock solution, or a numerical error such as omission of a decimal point when entering the data in the computer. However, not all deviations from normality are easy to recognize. Minor deviations will have a negligible effect on the parameter values estimated, but they may still have a major effect on the precision estimates.

Of the other assumptions, the one most likely to create problems is the third, the assumption of uniform standard deviation, because at least some investigations (Storer et al., 1975; Askelöf et al., 1976) suggest that a uniform coefficient of correlation will be likely to be closer to reality; this is relatively easy to incorporate into a fitting procedure, but only if one is aware that it needs to be done. SigmaPlot allows it to be done, but not easily (calling it an "advanced topic", rather as if knowing what buffer system to use at pH 8.5 is an advanced topic), and not precisely correctly. One may guess that other popular programs have similar characteristics, but I am not aware that any systematic testing has been done, so using any of them involves investing more confidence in the competence of the programmer than is wise.

It is possible to decrease the dependence on assumptions about whose truth or otherwise there is little or no information, either by using distribution-free methods (Cornish-Bowden and Eisenthal, 1974) or by using internal evidence in the data to suggest the most appropriate weighting scheme for least-squares analysis (Cornish-Bowden and Endrenyi, 1981). The former approach is easy to apply to Michaelis–Menten data, but computationally inconvenient for equations of more than two parameters; the latter is readily generalizable. However, neither of them, as far as I know, has been incorporated into commercial programs in current use.

I have discussed these questions in more detail elsewhere (Cornish-Bowden, 2012).

What does a non-linear least-squares regression program actually do?

For any set of observed rates v, a program finds the parameter values for which the weighted sum of squared differences $\sum w(v-\hat{v})^2$ between the observed values v and the calculated values $\hat{v}$ is a minimum. If the rates are known to have uniform standard deviation then each weight w is set to 1; if they are known to have uniform coefficient of variation they should be weighted according to the true values, but as these are always unknown one must use calculated values as the best estimates, i.e., $w = 1/\hat{v}^2$. intermediate and other weighting schemes are also possible, but commercial programs usually make no provision for these.

In introducing proper methods of statistical analysis to enzymology, Wilkinson (1961) used the following series of (a, v) pairs to illustrate the method he proposed: (0.138, 0.148); (0.220, 0.171); (0.291, 0.234); (0.560, 0.324); (0.766, 0.390); (1.46, 0.493). This example allows a simple test of whether a commercial program actually calculates what it is claimed to calculate. For a uniform standard error it should give $\hat{K}_m = 0.59655$, $\hat{V} = 0.69040$ and for a uniform coefficient of variation it should give $\hat{K}_m = 0.51976$, $\hat{V} = 0.64919$ (the circumflexes indicate that these are least-squares values). SigmaPlot 11.2 gets the first calculation correct, for example, but for the second it gives $\hat{K}_m = 0.5519$, $\hat{V} = 0.6632$ which is not correct. The discrepancy is within experimental uncertainty, of course, and has little practical importance, but it still illustrates the important point that one cannot assume that the authors

of commercial programs really understand what they are trying to do. I have discussed elsewhere (Cornish-Bowden, 2012) how they could have obtained such a result. It would be interesting to make similar studies of the results given by other widely used commercial programs, but as far as I know this has not been done.

Presentation of kinetic results

The third question posed in the introduction referred to the information to be provided in reporting the results of a kinetic experiment in such a way that they can be confirmed by other workers. As this is in a sense the theme of this entire book, it is dealt with in other chapters, but a brief summary can be given here (see also Tipton et al., 2014).

Any report of a kinetic investigation should specify how many complete independent experiments were carried out, and should include estimation of the precision of the parameters obtained. For oligomeric enzyme it should be clear whether the values are relative to one subunit or for one molecule. If the enzyme molarity is known (as will usually be the case for well characterized enzymes today), the catalytic constant k_{cat} should be reported, but otherwise the limiting rate V. Ideally, kinetic values for both the forward and reverse directions of reaction should be reported, especially if the equilibrium constant is such that the reverse reaction can be expected to be significant. It is especially important to report data for the reverse reaction if the results are intended for metabolic modelling, but they can also provide valuable mechanistic information.

The method used for calculating the kinetic parameters should be specified, together with the assumptions made about error distribution. The criterion used for choosing a particular equation to fit should be given. For example, if parameters are reported for competitive inhibition, what criteria were used to decide that any uncompetitive component in the inhibition could be neglected? If the inhibitor concentration for 50% inhibition is reported (not recommended in serious kinetic studies, but commonplace in pharmacological studies), appropriate mechanism-based inhibition constants should also be reported.

In all reports the ranges of concentrations (substrate always, inhibitors etc. if relevant) used should be clearly stated, as should all other relevant conditions, including the pH, the type of buffer, and the temperature.

Conflict of interest statement

The author has no conflict of interest.

References

Alberty, R.A., 2011. Enzyme Kinetics: Rapid-Equilibrium Applications of Mathematica. Wiley, Hoboken, New Jersey.

Askelöf, P., Korsfeldt, M., Mannervik, B., 1976. Error structure of enzyme kinetic experiments: implications for weighting in regression analysis of experimental data. Eur. J. Biochem. 69, 61–67.

Bisswanger, H., 2002. Enzyme Kinetics: Principles and Methods, 2nd ed. Wiley-VCH, Weinheim.

Cook, P.F., Cleland, W.W., 2007. Enzyme Kinetics and Mechanism. Garland Science, London.

Cornish-Bowden, A., Eisenthal, R., 1974. Statistical considerations in the estimation of enzyme kinetic parameters by the direct linear plot and other methods. Biochem. J. 139, 721–730.

Cornish-Bowden, A., Endrenyi, L., 1981. Fitting of enzyme kinetic data without prior knowledge of weights. Biochem. J. 193, 1005–1008.

Cornish-Bowden, A., 2012. Fundamentals of Enzyme Kinetics, 4th ed. Wiley-VCH, Weinheim.

Cortés, A., Cascante, M., Cárdenas, M.L., Cornish-Bowden, A., 2001. Relationships between inhibition constants, inhibitor concentrations for 50% inhibition and types of inhibition: new ways of analysing data. Biochem. J. 357, 263–268.

Copeland, R.A., 2000. Enzymes: A Practical Introduction to Structure, Mechanism and Data Analysis, 2nd ed. Wiley-VCH, New York.

Draper, N.R., Smith, H., 1981. Applied Regression Analysis, 3rd ed. Wiley-Interscience, New York.

Duggleby, R.G., 1979. Experimental designs for estimating the kinetic parameters for enzyme-catalysed reactions. J. Theor. Biol. 81, 671–684.

Endrenyi, L., 1981. Design of experiments for estimating enzyme and pharmacokinetic parameters. In: Endrenyi, L. (Ed.), Kinetic Data Analysis. Plenum Press, New York, pp. 137–167.

Fersht, A., 1999. Structure and Mechanism in Protein Science. Freeman, New York.

Lineweaver, H., Burk, D., Deming, W.E., 1934. The dissociation constant of nitrogen-nitrogenase in *Azotobacter*. J. Am. Chem. Soc. 56, 225–230.

Mannervik, B., 1981. Design and analysis of kinetic experiments for discrimination between rival models. In: Endrenyi, L. (Ed.), Kinetic Data Analysis. Plenum Press, New York, pp. 235–270.

Marangoni, A.G., 2002. Enzyme Kinetics: A Modern Approach. John Wiley and Sons, Hoboken, NJ.

Storer, A.C., Cornish-Bowden, A., 1974. Concentration of $MgATP^{2-}$ and other ions in solution. Biochem. J. 159, 1–5.

Storer, A.C., Darlison, M.G., Cornish-Bowden, A., 1975. The nature of experimental error in enzyme kinetic measurements. Biochem. J. 151, 361–367.

Tipton, K., Armstrong, Richard N., Bakker, Barbara, Bairoch, Amos, Cornish-Bowden, Athel, Halling, Peter, Hofmeyr, Jan-Hendrik, Leyh, Thomas S., Kettner, Carsten, Raushel, Frank M., Rohwer, Johann, Schomburg, Dietmar, Steinbeck, Christoph, 2014. Standards for Reporting enzyme Data: the STRENDA Consortium: what it aims to do and why it should be helpful. Perspect. Sci. 1, 131–137.

Tukey, J.W., McLauglin, D.H., 1963. Less vulnerable confidence and significance procedures for location based on a single sample: trimming/Winsorization 1. Sankhyā A 25, 331–352.

Wilkinson, G.N., 1961. Statistical estimations in enzyme kinetics. Biochem. J. 80, 324–332.

Yagi, K., Ozawa, T., 1960. Complex formation of apo-enzyme, coenzyme and substrate of D-amino acid oxidase. I. Kinetic analysis using inhibitors. Biochim. Biophys. Acta 42, 381–387.