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