Enzymes in context

Kinetic characterization of enzymes for systems biology

The kinetic behaviour of enzymes is typically observed in conditions appropriate for studying questions of mechanism of action, but these are not necessarily the most appropriate for studying their physiological roles, because they are often too far from those that exist in the living organism. Enzymes therefore need to be studied with natural substrates in the presence of all of the other small molecules likely to affect the activity *in vivo*, including the reaction products, so that the reverse reaction is not artificially prohibited. As complete reversible rate equations are often unmanageably complicated, especially for cooperative kinetics, care needs to be taken in choosing simpler equations that preserve the properties that are relevant in physiological conditions.

by Athel Cornish-Bowden (Centre National de la Recherche Scientifique, Marseilles, France) **and Jan-Hendrik S. Hofmeyr** (University of Stellenbosch, South Africa)

The rapid growth of systems biology has accentuated a problem that has usually been ignored, though it has always existed. Nearly all of the kinetic information that is now available for many enzymes was gathered in experiments designed as if to shed light on mechanisms of action. Establishing mechanisms of action is, of course, an entirely legitimate reason for studying enzymes, and it has led to a large part of the biochemistry that we now know. It is not the *only* legitimate reason, however, and gathering information about mechanisms is not necessarily the best or only way to understand enzyme physiology. Nonetheless, even experimenters with a primary interest in physiological questions have usually characterized enzymes in terms more appropriate for mechanism: assays are typically made with purified enzymes, in the absence of all metabolites that might influence the behaviour apart from any inhibitors or activators added specifically to answer specific mechanistic questions; no more than one substrate concentration is varied at a time; products are not present unless added specifically to study product inhibition; the pH and temperature are chosen for reasons of stability or convenience rather than to reproduce conditions in the cell, and so on.

To some degree these restrictions on the ways enzymes are studied have been necessary, and very little progress could have been made in understanding biochemistry if enzymes had never been purified and assays were always done on whole cells. Nonetheless, we have arrived at a state in which much of what we know about the kinetic behaviour of enzymes derives from experiments done in conditions very different from those that exist in the living organism. Moreover, mechanistic questions are often most easily answered by exploring behaviour at very high concentrations of substrates or inhibitors that accentuate slight differences between different mechanisms, whereas more light would be shed on physiological questions by exploring behaviour in the neighbourhood of physiological concentrations, not only of substrates but also of products and other effectors.

Perhaps the most important point is that a kinetic equation is always much simpler if written for irreversible conditions, because it then usually contains only a single numerator term and often becomes linear when written as an expression for the reciprocal rate. By contrast, a reversible equation always contains at least two terms in the numerator, at least one positive and at least one negative; it cannot be linearized and is less convenient to handle. With computer-based methods of data analysis this has become a trivial problem, however, because non-linear regression methods do not require a linearized equation.

To illustrate these points, the equation for an irreversible reaction obeying a compulsory-order ternary-complex mechanism takes the following form, in which v is the rate at concentrations a and b of the two substrates A and B and the other symbols represent kinetic constants[1]:

$$v = \frac{N_1 ab}{1 + D_1 a + D_2 b + D_5 ab} \quad (1)$$

This has just four parameters that are needed for calculating the rate at any combination of substrate

Key words: enzyme kinetics, Hill equation, Michaelis constants, systems biology.

concentrations, and the expressions for $1/v$, a/v and b/v are all linear. In the presence of the product that binds first, it acquires three additional parameters:

$$v=\frac{N_1ab}{1+D_1a+D_2b+D_3p+D_5ab+D_6ap+D_7abp} \quad (2)$$

However, this is still easily manageable, as it has just one term in the numerator and still gives linear expressions for $1/v$, a/v and b/v. The fully reversible form, however, is much more complicated:

$$v=\frac{N_1ab-N_2pq}{1+D_1a+D_2b+D_3p+D_4q+D_5ab+D_6ap+D_7bq+D_8pq+D_9abp+D_{10}bpq} \quad (3)$$

Not only are there now twelve parameters that in principle need to be known, but the presence of two terms in the numerator removes any possibility of writing the equation in a linear form. All of this is just for the simplest cases that usually need to be considered: when there are more than two substrates or products, or the enzyme shows cooperativity, the equations become much more complicated.

As a result, enzymes are almost never studied under fully reversible conditions, and when product inhibition is studied, irreversibility is still usually ensured by adding only one product at a time, so that there is not a complete reaction mixture for the reverse reaction. (Reactions with only a single product, such as those catalysed by isomerases, present special difficulties for the study of product inhibition, but we shall not consider these here.) However, this immediately creates an almost unbridgeable gap between enzymes in the spectrophotometer and enzymes in the cell, because physiological conditions are always reversible: all products are always present in the steady state, as well as numerous other metabolites that are normally excluded from reaction mixtures *in vitro*.

The development of systems biology has made this gap far more evident than it used to be. As long as enzymes were considered one at a time it was comparatively unimportant if most of the knowledge about their behaviour was unphysiological, but as soon as one starts to regard them as components of systems, it becomes essential to use data that refer to conditions that are at least approximately physiological, and the possibility of the reverse reaction is not excluded. Moreover, combining data for several enzymes into a single metabolic model requires the data to refer to comparable conditions. Ideally, therefore, one would like to have data in which all of the enzymes have been studied

1. under reversible conditions (with all substrates and products present);
2. in the presence of all metabolites likely to affect the activity *in vivo*;
3. at the same temperature, pH, ionic strength, degree of macromolecular crowding, etc. (ideally those that apply *in vivo*);
4. by the same research group (to minimize uncontrolled variations).

In practice, hardly any of these requirements are satisfied, even approximately, but when most of them are, as with data used to model metabolism in *Trypanosoma brucei*[2] and yeast[3], a metabolic model can accurately reproduce behaviour observed for complete cells. More usually, as in models of metabolism in *Dictyostelium* developed by Wright and colleagues[4], one may need to assume very large differences (factors of 100 or more) between the kinetics measured *in vitro* and those assumed to apply *in vivo* in order to make the model fit.

This is clearly not satisfactory if systems biology is to advance, and minimal criteria need to be defined for the kinetic characterization of enzymes to be used in models. Similar conclusions have been reached at the Beilstein Institute, and this short article was stimulated by an initiative of the Institute to improve matters, and, in particular, to make data stored in databases such as BRENDA (http://www.brenda.uni-koeln.de/) more appropriate for metabolic models.

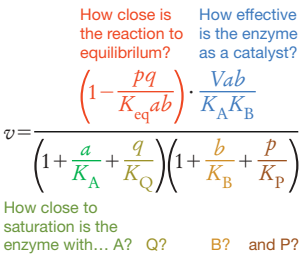
Eqn (3) is in a sense the minimal equation for representing a reaction with two substrates and two products in a metabolic model. As it stands, however, it presents two practical difficulties: as a reversible equation it cannot be written in a linear form, and with twelve adjustable parameters it demands much more experimental information than is likely to be readily available.

There is nothing we can do about the first of these difficulties: reactions in living systems are reversible, and we need to live with it. The second, however, is not so hopeless. We do not need an equation capable of accounting for the fine differences in kinetic behaviour predicted by different mechanisms, so it is unlikely to matter that with some mechanisms the term in *abp* is predicted to be zero and in others it

is not. What matters is that the equation can give an adequate prediction of the rate in physiologically realistic conditions, and that it predicts a zero rate at equilibrium. Incorporating the equilibrium constant explicitly in the rate equation will fully take care of this last point, and writing the denominator as a set of saturation terms multiplied together may well deal adequately with the rest. In other words, we need to write the equation in a way that completely separates thermodynamic and kinetic considerations, as in the following example:

$$v = \frac{\left(1 - \frac{pq}{K_{eq}ab}\right) \cdot \frac{Vab}{K_A K_B}}{\left(1 + \frac{a}{K_A} + \frac{q}{K_Q}\right)\left(1 + \frac{b}{K_B} + \frac{p}{K_P}\right)} \quad (4)$$

As this contains only half as many parameters as eqn (3) it obviously cannot express the full range of kinetic behaviour represented by eqn (3), but it can express the major features, and, most important, it insists on adherence to thermodynamic constraints.



It can be analysed into six conceptually distinct properties, as illustrated in the colour-coded version of the equation in Figure 1.

The first numerator factor is a pure thermodynamic term, containing no kinetic information, and provided that the correct equilibrium constant is used for K_{eq} it ensures that the rate is zero at equilibrium and in the correct direction otherwise. The second numerator factor defines the kinetic activity of the enzyme, V being the limit of the forward rate approached at saturation with A and B and in the absence of products. Each of the four denominator factors expresses the degree of saturation with one of the four reactants.

The parameters K_A etc. in these terms are apparent Michaelis constants, but they are not real Michaelis constants because they do not refer to saturating concentrations of co-substrates and zero concentrations of products. Ideally they should be measured at appropriate physiological concentrations of the other reactants, but this immediately raises the problem that the physiological concentrations of metabolites are not constant but vary according to the metabolic state. Nonetheless, even if "one-size-fits-all" concentrations are chosen to represent typical physiological conditions, that is already an advance on the strict definitions, because if there is one thing we can be certain of about the physiological concentrations of metabolites it is that they are not infinite.

Measuring Michaelis constants in reversible conditions, with all substrates and products present, may be difficult, but they can often be adequately approximated by apparent Michaelis constants measured with only one product present: for example, K_A could be estimated as the apparent Michaelis constant for A with B and P at appropriate concentrations and Q absent; K_P could be measured as the apparent Michaelis constant for P with Q and A at appropriate concentrations and B absent, and so on.

Eqn (4) can be generalized in an obvious way to allow for reactions with more than two substrates, or more than two products, or both.

In contrast to more "mechanistic" equations, there is no combinatorial explosion in this case, as each reactant added to the model just implies one additional parameter. Reactions subject to inhibition by metabolites present in the model, but not substrates or products of the reaction considered, are normally straightforward to include, as each such inhibitor I just implies an additional term i/K_I in the denominator.

Matters become more complicated when we need to consider cooperative effects, whether of substrates and products or of allosteric effectors. Nonetheless, the case is not hopeless if we recognize that much of the disagreement 30 years ago about which of the classical models of cooperativity was the right one for a particular set of data was due to the fact that the predictions made by equations that look very different are almost indistinguishable in many experimental contexts. For irreversible rate equations there has always been a simple solution to this, as it has long been realized that the Hill equation, with just three parameters, V, K_A and h,

$$v = \frac{Va^h}{K_A^h + a^h} \quad (5)$$

can usually describe the behaviour over the middle and high ranges of saturation just as accurately as mechanistically more realistic equations.

However, there is just as much need to take account of reversibility in this case as there is for non-cooperative enzymes, and the Hill equation can be generalized to allow for reversibility and allosteric effects[5]. Various forms are possible, depending on the particular case considered. Here we shall just give the example of a reaction with one substrate A, one product P and an allosteric

Figure 1. Information content of a kinetic equation. An equation can be written in such a way that the different kinds of information contained in the different terms are clearly separated. In this colour-coded version of eqn (4), the first numerator term (red) is purely thermodynamic, and expresses how far the reaction is from equilibrium and its sign defines the direction in which it can go; the second numerator term (blue) defines how effective the enzyme is as a catalyst; the various denominator terms (other colours) express how close to saturation with each substrate or product the enzyme is.

inhibitor I, for which the rate can be expressed as follows (shown for simplicity for the case where substrate and product cannot bind to molecules that have inhibitor bound):

$$v=\frac{\left(1-\frac{p}{K_{eq}a}\right)\cdot\frac{Va}{K_A}\cdot\left(\frac{a}{K_A}+\frac{p}{K_P}\right)^{h-1}}{1+\left(\frac{a}{K_A}+\frac{p}{K_P}\right)^{h}+\frac{i^h}{K_I^h}} \quad (6)$$

Like eqn (4), this can be analysed into thermodynamic, activity and saturation terms, with, in addition, a term that expresses the degree of cooperativity.

As the thermodynamic term is written without any exponent *h* this ensures that the rate is zero at equilibrium and in the right direction otherwise, and the equation simplifies to the ordinary Hill equation if *p* and *i* are zero.

Metabolic models are fortunately rather robust to faults in the kinetic data included in them, at least if they are just used for calculating metabolic fluxes.

This mean that even quite large errors in kinetic parameters for most of the enzymes have very little effect on the calculated flux. This is a natural consequence of the flux summation property, which leads to an expectation that most flux control coefficients are numerically small, and thus the importance of any errors in the kinetic parameters is decreased as each enzyme is added to the model. However, there are important exceptions, as some enzymes do make major contributions to flux control, and in any case metabolite concentrations are much more sensitive than fluxes to errors. In general, therefore, both the extent and quality of data used in metabolic models need to be improved.

References

1. Cornish-Bowden, A. (2004) Fundamentals of Enzyme Kinetics (3rd edn), Portland Press, London
2. Bakker, B.M., Michels, P.A., Opperdoes, F.R. and Westerhoff, H.V. (1997) J. Biol. Chem. **272**, 3207–3215
3. Teusink, B., Passarge, J., Reijenga, C.A. et al. (2000) Eur. J. Biochem. **267**, 5313–5329
4. Wright, B.E., Butler, M.H. and Albe, K.R. (1992) J. Biol. Chem. **267**, 3101–3105
5. Hofmeyr, J.-H.S. and Cornish-Bowden, A. (1997) Comput. Appl. Biosci. **13**, 377–385

Athel Cornish-Bowden is Directeur de Recherche in the CNRS Unit of Bioenergetics and Protein Engineering in Marseilles, France. His main interests are in enzyme kinetics in general, and their application to multi-enzyme systems in particular.

email: acornish@ibsm.cnrs-mrs.fr

Jan-Hendrik Hofmeyr is Professor of Biochemistry at the University of Stellenbosch, South Africa. His main interests are in metabolic regulation and metabolic control analysis.

email: jhsh@sun.ac.za

