METHODS 24, 181–190 (2001)
doi:10.1006/meth.2001.1179, available online at http://www.idealibrary.com on IDEAL®

Detection of Errors of Interpretation in Experiments in Enzyme Kinetics

Athel Cornish-Bowden[1]

Institut Fédératif "Biologie Structurale et Microbiologie," Bioénergétique et Ingénierie des Protéines, Centre National de la Recherche Scientifique, 31 chemin Joseph-Aiguier, B.P. 71, 13402 Marseille Cedex 20, France

Although modern statistical computing will often be the method of choice for analyzing kinetic data, graphic methods provide an important supplement that ought not to be neglected. Residual plots, or plots of differences between observed and calculated values against variables not expected to be correlated with these differences, permit a rapid judgment of whether data have been correctly interpreted and analyzed. The rapid increase in the frequency with which artificially modified or mutated enzymes are studied is making it less and less safe to assume that enzymes are stable under assay conditions, and there is thus an increased need for methods to check for enzyme stability, and a method for doing this is briefly described. Finally, the Scatchard plot (together with the Eadie–Hofstee plot) is used as an example to discuss the dangers of publishing derived information unaccompanied by any primary data. © 2001 Academic Press

Key Words: residual plot; error detection; Scatchard plot; enzyme inactivation; enzyme assay; data analysis.

This article discusses methods of detecting errors in the interpretation of kinetic data. The methods are not in general new, and some are not specific to enzyme kinetic experiments, but as they are little used in the modern literature and seem to be little known, while being at the same time easy to apply and rich in information, it appears useful to bring them to the attention of a wider range of biochemists.

Model testing as such is, of course, within the domain of statistical analysis, and there are a number of sources of information for applying it to enzyme kinetic data, beginning with the influential articles of Cleland (1). However, this approach is not discussed here as it has not significantly changed over the years; the main points necessary for modern users are to be aware (at least qualitatively) of the assumptions that underlie the statistical analysis and to be aware also that they are not always true (2). By contrast we shall be concerned with methods that require no detailed mathematics and can readily be applied to published graphs (or even to graphs displayed on a screen during a lecture) without access to the numerical coordinates of the observations. More generally, this article argues for more reliance on common sense and graphic analysis and less on automatic analysis performed by computer: an example (out of many that could have been chosen) is given below to illustrate how automatic analysis by computer can lead to results that are obviously wrong. As this approach runs strongly counter to the trend in biochemical practice over the past 20 years, the next section illustrates that it has support from some of the main authorities in modern statistics.

RESIDUAL PLOTS

Statistical Background

Senior biochemists with some experience in the use of computational methods for analyzing data of all kinds often respond to the advocacy of graphic methods with the opinion that the graphic approach has been completely outmoded by the universal availability of computers, and that established statistical methods provide the informed experimenter with all of the information necessary to proceed. Widespread as this opinion is among scientists, it is important to realize that it is not shared by the people who revolutionized statistical practice and theory in the second half of the 20th century.

[1] To whom correspondence should be addressed at CNRS-BIP, 31 chemin Joseph-Aiguier, B.P. 71, 13402 Marseille Cedex 20, France. Fax: + 33 491 16 45 78. E-mail: athel@ibsm.cnrs-mrs.fr.

1046-2023/01 $35.00
Copyright © 2001 by Academic Press
All rights of reproduction in any form reserved.

For example, Chambers and colleagues (3) wrote that "There is no single statistical tool that is as powerful as a well-chosen graph. Our eye–brain system is the most sophisticated information processor ever developed, and through graphical displays we can put this system to good use to obtain deep insight into the structure of data."As this statement appeared in a book about graphic methods it could perhaps be regarded as biased, but it is echoed in other authoritative sources, for example: "Graphs are important" (4), or "Mechanizing . . . seems dangerous. The user needs some contact with what is going on" (5). This last example is perhaps not easy to understand out of context, but the point being made (which underlies the whole book from which it is quoted) is that data analysis demands constant interaction between analyst and computer; major decisions cannot be left to the machine.

The author common to the two sentences just quoted was also the originator of the word "bit" (in the sense of binary digit) and inventor of the fast Fourier transform, something that ought to convince even the most skeptical reader that a belief in the usefulness of graphs is not incompatible with long experience with computing and deep understanding of its principles.

Residual plots constitute the major focus of this paper. These are plots of the differences between observed and calculated values (of a rate, for example) against the calculated value or some other convenient variable, which in enzyme studies may often be elapsed time or the concentration of substrate. They are widely advocated in the statistical literature, but although they have sometimes been recommended to biochemists (e.g., 6) they remain very little used by them. As I argue in this article, residual plots allow various types of error in the design and analysis of experiments to be eliminated before it is too late. In addition, it is useful to cultivate the habit of imagining what a published graph would have looked like if the authors had displayed residuals instead of measured values; this often allows a rapid assessment of the credibility of the interpretation offered.

For the statistical background it is useful to consult an article in which Anscombe (7) gave an example of a series of four experiments that when analyzed by classic methods give numerically identical estimates of the fitted parameters together with numerically identical estimates of their variances, even though the underlying error structures are grossly different. Only one of these experiments would look at all odd in an inspection of the tabulated numerical data; the other three would look more or less the same if they could only be compared in the form of three tables of numbers, such as one might obtain from computer analysis. Nonetheless, as Anscombe pointed out, the differences leap out at the eye in even the most cursory graphic comparison of the residuals. His example was constructed for straight-line data, but I have used it as the basis for a similar comparison between results that one might obtain in a simple kinetic study of an enzyme (pp. 6–8 in Ref. 2), at the same time extending it to illustrate a fifth type of error structure to add to the four considered originally.

Examples from the Literature

Many graphs continue to be published that illustrate the reasons for plotting residuals. Recent examples of lines that manifestly fail to fit the observations that accompany them can be found in the major journals of biochemistry (8–11). Consider Fig. 1a, for example, which shows a Scatchard plot from one such paper redrawn as a generic plot of a dependent variable y as a function of an independent variable x. Even in the original coordinates it is obvious that the observations do not agree with the implied model, not only quantitatively but also qualitatively, given that the observations clearly demand a curve although a straight line is drawn. When the points are replotted as residuals (Fig. 1b) the same conclusion emerges even more clearly, as the points follow an obvious trend and are very far from being distributed at random (quite different from those in Fig. 2a below). In general a residual plot always magnifies deviations from expectation, and thus always makes it easier to judge how satisfactory the proposed model is. For this reason it is often helpful to imagine how the equivalent of Fig. 1b would look when all that is immediately available is the equivalent of Fig. 1a.

The line shown in Fig. 1a was obtained by fitting the data by regression analysis, but it is clearly not correct. Many similar examples could be given from the current literature, and all of them illustrate the important point

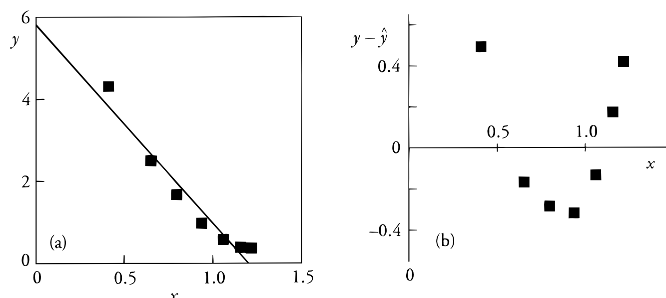

FIG. 1. Example of a straight line drawn through points that lie on a curve. The plot in (a) is based on data for binding of diphtheria toxin to an antigen from monkey (7). When redrawn as a residual plot of differences between observed and calculated values against the same abscissa, as in (b), the systematic trend becomes far more obvious. According to the original article "the data were fitted by regression analysis."

that computation alone provides no guarantee against reaching an obviously incorrect conclusion. The results of statistical analysis always need to be inspected by eye. Nonetheless, it is less easy to find current examples than it was 10 or more years ago, but it would be unwise to conclude from this that the problem illustrated in Fig. 1 now occurs less frequently than it did. On the contrary, it may well be much more frequent than it used to be, but is becoming less visible as a consequence of the increasing (and disquieting) practice of presenting derived kinetic parameters without any primary data to illustrate them. Such parameter values are often supported by only very brief and uninformative indications of how they were calculated, such as "all the kinetics of hydrolysis were fitted to the hyperbolic Michaelis–Menten equation" (12) or "steady-state kinetic parameters were calculated fitting the data using the GRAFIT program" (13).

Even more disquieting, it is increasingly common that papers not only report no primary data but also report no secondary data either. For example, in a recent study of a sialyltransferase (14) the parameters were obtained by analysis of progress curves, but no progress curves (the primary data) are illustrated; not only that, no rate data (the secondary data) are illustrated either; all that are shown are numbers derived from rate data derived from progress curves by application of commercial graph-drawing software that happened to include a curve-fitting function. There was nothing to suggest the use of reliable methods for analysis of progress curves, such as those discussed elsewhere in this issue (15). If one compares not the gross numbers of clearly faulty graphs published but their proportions to the total amount of primary data published there is no suggestion of a gradual improvement. We return to this point in the discussion of the Scatchard plot at the end of this article.

When assessing derived results that are already published without any primary data there is little the critical reader can do apart from exercising healthy skepticism about any conclusions drawn. When primary data are available, however, residual plots provide a powerful tool for reaching a rapid (but of course preliminary) opinion about the validity of the interpretation.

Choice of Coordinates

The ordinate in a residual plot is normally the weighted difference $w^{0.5}(y - \hat{y})$ between the observed value y of some quantity (such as an initial rate) and the value $\hat{y}$ calculated from the model that is considered to describe the data. The weight $w^{0.5}$ is the square root of whatever weight w was considered appropriate for y when fitting the data by least squares. If the y values were equally weighted then the simple difference $(y - \hat{y})$ can be plotted as the ordinate variable. This choice is appropriate even if the data are weighted if one is trying to judge what weighting system should be used (see the description of Fig. 2b below).

The abscissa variable can be any variable that is not expected to be correlated with the ordinate variable. With a correctly chosen model correctly fitted there ought to be no relationship between the residual $w^{0.5}(y - \hat{y})$ and the true value of y, a point that is amplified in the next paragraph. This true value is unknown except in a simulated experiment, but it can be represented by the calculated value $\hat{y}$, and so $\hat{y}$ is often a good choice for the abscissa variable. It is an especially good choice for assessing the validity of the weighting scheme, but it is not the only choice. For example, to check for any unwanted dependence of the results of an experiment on the progressive deterioration of a stock solution it is appropriate to choose the time elapsed after preparation of the stock solution as the abscissa variable. Note that in general there is no requirement for the abscissa variable to be one that played any role in fitting the data.

If a correlation does exist between $w^{0.5}(y - \hat{y})$ and the abscissa variable, and if the abscissa variable is $\hat{y}$ or one of the independent variables used for fitting the data, the correlation will inevitably be nonlinear; i.e., any trend in the residual plot will follow a curve. Even if a completely (and even obviously) wrong model has been fitted the mechanics of least-squares regression will eliminate any linear correlation and ensure that the mean displacement of the points from the abscissa axis is zero. This is exactly true if a linear model has been fitted by least squares, but it is also likely to be true within the limits of detectability by eye if a nonlinear model has been fitted or if the fitting has been done by a method other than least squares. Inspection of a residual plot is normally therefore a matter of looking for nonlinear trends.

For simplicity in discussing examples of residual plots we can assume that unless otherwise indicated the ordinate variable is $w^{0.5}(y - \hat{y})$ and the abscissa variable is $\hat{y}$.

Characteristic Residual Plots

Figure 2 shows a selection of characteristic appearances that a residual plot may have. In each of the first eight cases (Figs. 2a–2h) the gray shading imposes an interpretation of trend, if any, in the data points. The subjectivity or otherwise of this interpretation is discussed below in the context of Fig. 2h. For the moment we shall assume that the gray shading provides a valid picture of the underlying trend.

In Fig. 2a the points are scattered in a parallel band symmetrically placed about the horizontal axis and evenly distributed in both dimensions. Although a small proportion of points may lie outside the main band none of them is far outside it. In an experiment with a very large number of observations one expects the points to be more concentrated near the axis, but this may not be obvious with fewer than about 50 observations. This plot illustrates the ideal: it does not of course prove that the data are correctly interpreted and the correct model fitted, but it provides no immediate reason to doubt this.

The parallel nature of the band in Fig. 2a suggests that the observations were correctly weighted during fitting. Figure 2b shows a typical result of inappropriate weighting, for example, if $w^{0.5}$ were set to 1 when it would have been appropriate for it to be a decreasing function of the abscissa variable. There is nothing here to suggest that the model itself is inappropriate, but it should still be checked after fitting the data again with appropriate weights, because any systematic trend will be more visible in the residual plot if there is no accompanying problem with the weighting.

Figure 2c is perhaps the most important case for discussing model discrimination, and is typical of the results one obtains when the wrong model has been fitted. Although one may feel that the trend, and hence the lack of fit, is so obvious that one hardly needs a

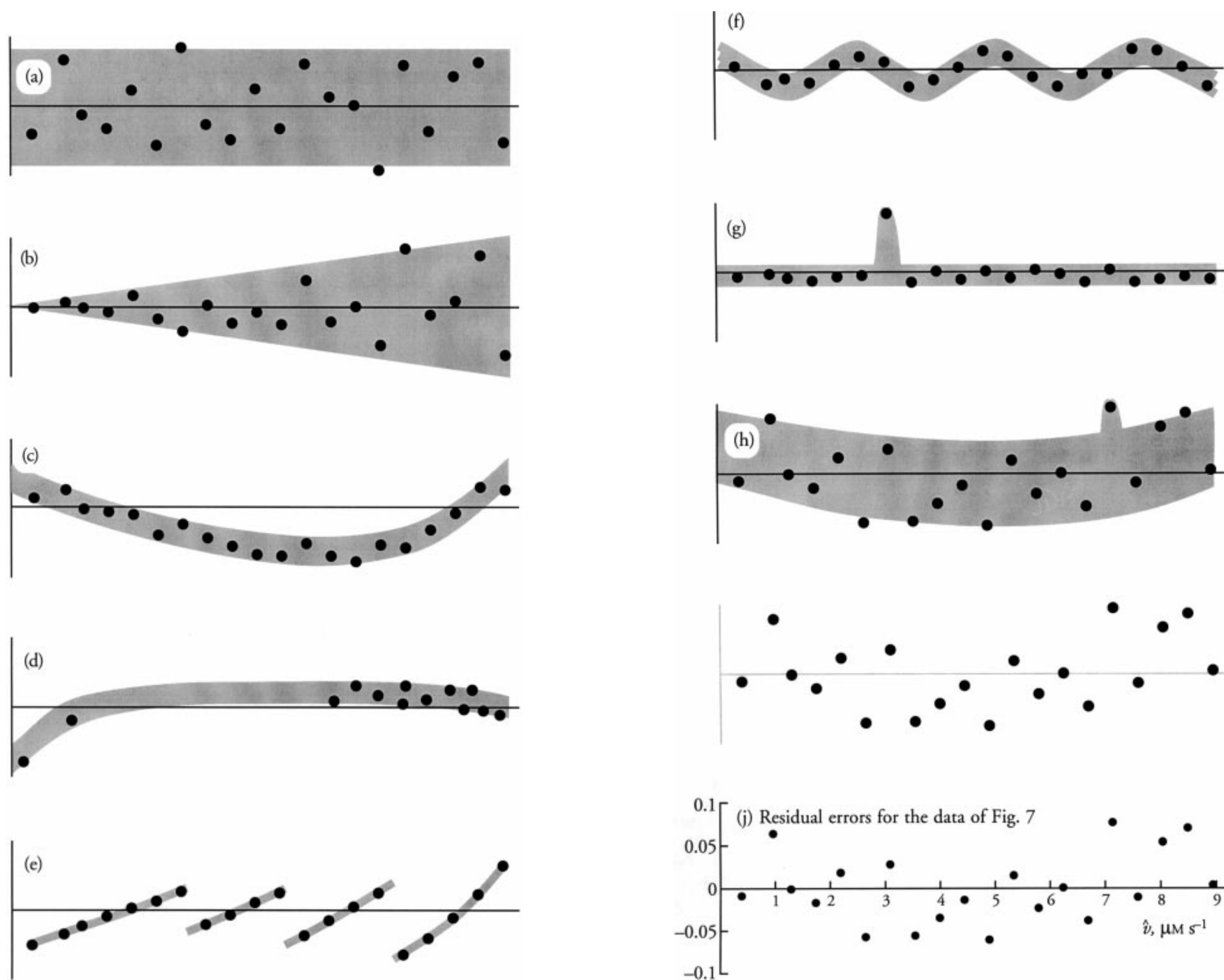

FIG. 2. Representative examples of residual plots in the presence of various kinds of anomaly. In plots with shading the shading represents an interpretation of the arrangements of points. (a) Ideal case with no anomaly; (b) effect of inappropriate weighting; (c) systematic trend resulting from fitting an incorrect model (cf. Fig. 1b for a real example); (d) effect of using a poor experimental design incapable of supplying the information expected; (e) effect of excessive rounding of the values in the primary observations before analysis; (f) effect of a periodic disturbance; (g) effect of the presence of an outlier, i.e., a single observation with very large error; (h) effect of a moderate outlier in the presence of a weak systematic trend. (i) and (j) show the same points as (h): in (i) the shading is omitted to avoid imposing an interpretation on the data, and the label is omitted and the axes shown in gray to place essentially all the emphasis on the data; in (j) the presence of a large amount of extraneous information interferes with inspection of the arrangement of points.

residual plot to illustrate it, there are two reasons why this is too optimistic. First, many plots are published where the trend is even more obvious than it is here; for example, in Fig. 1 all of the dispersion is systematic, with no random scatter even when the residuals are plotted, and one can only suppose that the question of whether the right model had been used was not even asked. The advantage of the residual plot here is that it converts a strong suggestion into a result that cannot be missed. The second point is that sometimes one may have very accurate data that appear in the primary plot to fit the model very precisely, even though there is a systematic trend; in such a case, discussed further below in relation to Fig. 3, a trend that may pass unnoticed in the primary plot becomes obvious in the residual plot.

Figure 2d may also illustrate a case of a systematic trend, but here it is largely obscured by a poor experimental design. If the two points at lowest $\hat{y}$ are ignored the others could be taken to suggest a weak decreasing dependence of residual on $\hat{y}$, but if they are included they suggest a more complex dependence. The question can be resolved only by repeating the experiment with some observations between the two groups. Although in such a case one could argue that the poor design ought to be detected in the primary plots, the residual plot again converts a strong suggestion into something that cannot be ignored. Here even a primary plot represents a major advance over a table of numbers, which can usually be arranged with little difficulty in such a way as to conceal any deficiencies in the data.

Figure 2e illustrates a much less common problem, but one that is still worth mentioning not for its own sake but to make the point that residual plots magnify and make obvious even anomalies that are completely unexpected and not described in standard textbooks. In this case the distribution of points along a curve with several saltuses is characteristic of a residual plot in which overaggressive rounding before analysis provided the major source of error (16). It also illustrates the danger of assuming that the software built into a scientific instrument is of a quality to match the engineering precision of the instrument itself.

Figure 2f illustrates a periodic effect on the residuals. If the abscissa variable is $\hat{y}$ it may not be obvious how a direct periodic effect of $\hat{y}$ on the residuals could arise, but it could still happen if the experiment was designed in such a way that the measured y values decreased (or increased) monotonically during the course of the experiment and the periodic effect was caused by some environmental variable that fluctuated with time.

Figure 2g illustrates that if one outlying residual is numerically much larger than all the others it becomes possible to include it in the plot only if the scale is chosen so that most of the points are compressed close to the axis, thereby obscuring any trend that they may follow. In this case the first essential is to try to find a cause for the large residual. If it can be assigned to a simple mistake when recording the data, such as omitting a decimal point when writing down a number, then the mistake can be corrected. However, unless such a cause is definitely established one should never discard the possibility that the residual contains the principal interesting information in the experiment. Perhaps there is a genuinely important departure from simple behavior over a limited range of conditions that ought to be investigated more closely. In the case of Fig. 2g there is no suggestion of anomalous behavior in the neighboring points, so such an interpretation may seem

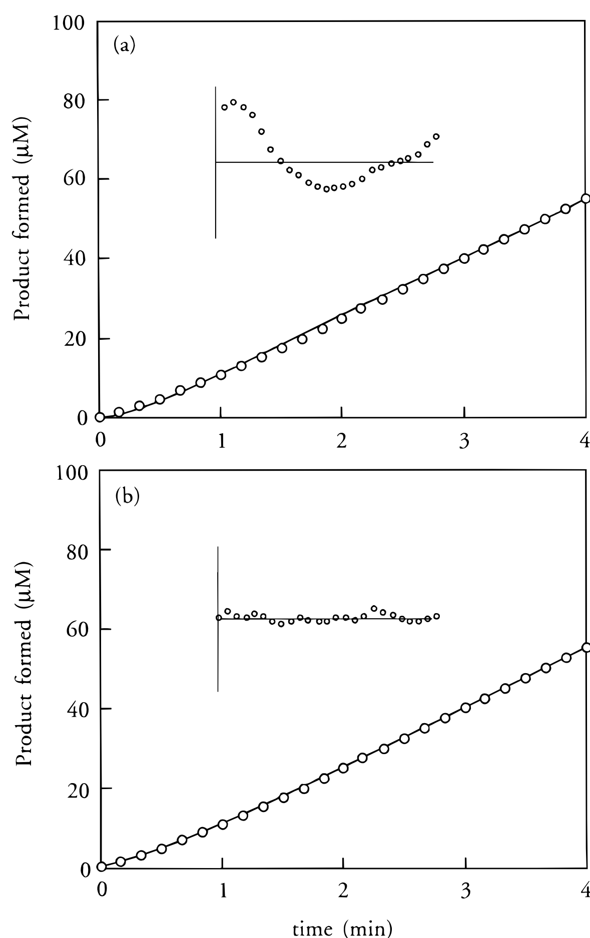

FIG. 3. Model discrimination with the aid of residual plots. Two models for the time course of product release in the reaction catalyzed by phosphoribulokinase are almost indistinguishable when compared as ordinary plots, but for the model treating complexed phosphoribulokinase as inactive (a) the residual plot shows clear evidence of systematic error, whereas for the model assuming the complex to be active (b) it does not. Reproduced, with permission, from Fig. 1 of Lebreton and Gontero (17).

unlikely; more generally it is unlikely in simple enzyme kinetics because nothing in what we know about how enzymes behave leads us to expect large deviations from simple models over small ranges of conditions.

To this point we have been assuming that only one effect at a time will cause a residual plot to differ in appearance from Fig. 2a, but the reality will usually be more complicated and several effects may be simultaneous and superimposed. Moreover, the systematic trend underlying the data may be less obvious than it is in Fig. 2c (and hence much less obvious than it is in Fig. 1b). Figure 2h thus illustrates a combination of a weak suggestion of a nonlinear trend with a weak suggestion of an outlier. In such a case the appearance of the plot is much more suggestive than conclusive, suggesting further experiments to be done before a definitive interpretation is possible. A weak suggestion of a systematic trend in the residuals indicates that the experimental design should be revised so as to accentuate the trend, if it exists, for example, by extending the range of the experiment or by making more observations in the region of design space where the trend is most evident. Increasing the experimental precision will always help, of course, but this is rarely a useful suggestion as normally the experiments will have been done as accurately as possible already.

In each of Figs. 2a–2h the gray shading represents an interpretation that will not be visible in the data when first plotted, and although such shading is useful in the context of a methodological discussion like this one or an educational article, in the research context it is usually better to omit shading as it strongly biases the eye. Compare Figs. 2h and 2i, for example, which show exactly the same data with and without shading. The relatively weak effects visible in Fig. 2h appear even weaker in Fig. 2i. They appear weaker still in Fig. 2j, which again shows exactly the same points, but here this is because the data are accompanied by a large amount of extraneous information that just confuses attempts to discern the pattern of points around the axes. In general all of the extra information is redundant in a residual plot; if one needs to record it somewhere for future reference or for publication, it is better to place it in a separate legend and not superimposed over the data.

Thus Fig. 2i shows how to display the data with a minimum of irrelevance. Even the label (i) is omitted, so one recognizes it as Fig. 2i by the fact that it is placed between Figs. 2h and 2j. More generally one can use expressions like "bottom right" rather than labels. If possible, and as also illustrated in Fig. 2i, it is best to show the axes in light gray or pale blue, so that they do not dominate the data but are visible if one wants to know where they are.

A practical example of the use of residual plots to discriminate between models that generate very similar curves (17) is illustrated in Fig. 3. The two curves are almost indistinguishable by eye, and in the absence of the residual plots one might well be inclined to accept the curve in Fig. 3a as a good fit to the data. However, the inset residual plot reveals a very obvious systematic trend, whereas the corresponding residual plot in Fig. 3b shows a slight suggestion of the effects of rounding error but no clear systematic trend. The residual plot allowed the authors to make a convincing case for a significant contribution of an enzyme–enzyme complex to the total reaction observed even though this contribution was too small to be detected in the conventional plot.

Moderating the Effect of an Outlier

An outlier such as that in Fig. 2g (reproduced without the shading as Fig. 4a) presents the problem of how to display the other residuals in such a way that any trend is apparent, because if a linear scale is used the nonoutliers will be compressed against the abscissa axis in such a way that it becomes difficult to perceive how they are distributed. If the scale is expanded so as to reveal the arrangement a large amount of space needs to be wasted if the outlier is to remain within the range of the graph. The most obvious solution is to place the outlier off-scale (effectively omitting it), as in Fig. 4, but this may be dangerous in case the outlier contains real information, and in general it is best to include all the data in any plot. An alternative is to use a "stabilizing transformation" that smoothly adjusts the largest values while having very little effect on most of them. In a residual plot the arctangent function has an appropriate stabilizing effect, because if x is measured in radians, arctan (x) is within a few percent of x if x is in the range -0.6 to $+0.6$, but remains within the range -1.6 to $+1.6$ even if x approaches infinity (Fig. 4c). This means that even very large residuals can be kept on-scale without greatly distorting the pattern of the others.

The main complication in this approach is that the residuals need to be scaled before calculating their arctangents, as otherwise the results will be arbitrarily dependent on the units of measurement. Any scaling that ensures that nearly all the values are in the range -0.6 to $+0.6$ will be satisfactory; a practical choice is to define μ as the mean of the absolute values $|x|$ of a series of residuals x and plot arctan ($x/2\mu$) in the residual plot. The effect of this on the data of Fig. 4a is shown in Fig. 4d; note that the largest residual remains on-scale (and would do so even if its value were much

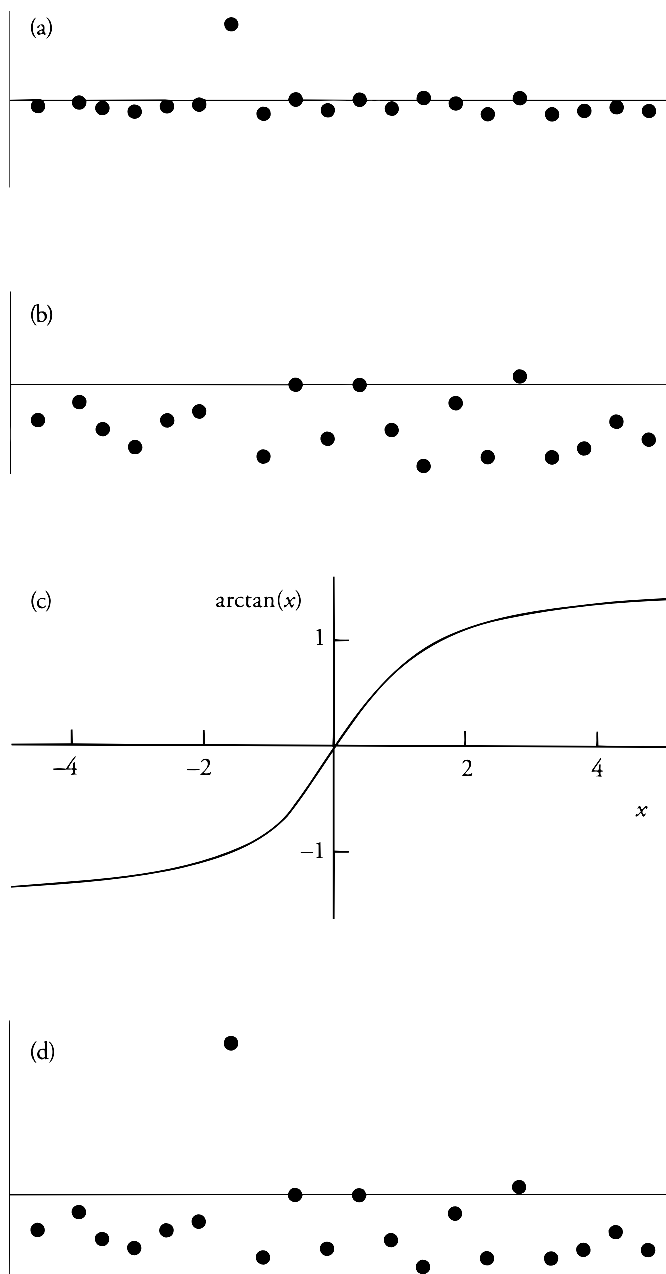

FIG. 4. Effect of an outlier on a residual plot. (a) The top part of the figure reproduces the data of Fig. 2g with the shading omitted. The presence of one large residual means that it can be plotted on a linear scale only if all the others are compressed against the abscissa axis, making any systematic pattern difficult to detect. (b) If the scale is expanded a large amount of space would be wasted if the most deviant value were not omitted from the plot. (c) The function arctan (x) has the property that it differs very little from x if x is in the range -0.6 to $+0.6$ but remains finite and small even if x is very far outside this range. (d) As a result, the arctangent transformation defined in the text allows all the residuals to be plotted while allowing any information contained in the smaller residuals to remain visible.

larger than that used), but the arrangement of the others differs little from that seen with a linear scale, as in Fig. 4. Note also that the outlier remains easily recognizable as an outlier, i.e., that the large distortion in its magnitude does not incur any danger that it will be confused with the other observations.

The arctangent transformation may be tedious to apply manually, of course. However, this presents no problem for a computer program that draws residual plots automatically (as any serious data-fitting program ought to do), as the computing effort is trivial.

Residual Plots with Sparse Data

It often happens, especially when trying to assess the credibility of data in the literature, that the number of observations in each plot is too small to permit a meaningful conclusion. Even such an obvious trend as that illustrated in Fig. 1, with only seven points, could conceivably be due to random scatter and not to a real lack of fit. It would be sufficient for just one point to be in a significantly different position (e.g., moving the third point in the residual plot to its mirror position about the x axis) for the trend that appears so unequivocal in the plot as it stands to be called into question.

Consider for example the three sets of points in the upper part of Fig. 5, which is based on results from a kinetic study of the proteolytic enzyme napsin (11). The fact that all of the triangles lie on the same side of the line that is supposed to fit them is noticeable, but by itself it means very little, as there is a probability of 12.5% that four random values from a symmetrical distribution will have the same sign. However, when all of the 12 residuals are plotted together, as in the lower part of Fig. 5, it is seen that now we have 11 negative deviations and one of about zero, an arrangement that is much more difficult to dismiss as a chance fluctuation and makes it clear that the lines drawn could not have been best-fit lines.

DETECTING ENZYME INACTIVATION

During the period when the basic methods for studying enzymes were being developed, during the first three-quarters of the 20th century, most work was done with natural enzymes or with enzymes that had been subjected to minimal modifications. Under these conditions it was reasonable to assume that enzyme structures had been selected during evolution to be sufficiently stable in the cell to fulfill their physiological functions (though not necessarily under conditions that can easily be produced in the assay). This does not

mean, of course, that all natural enzymes are as stable as, say, ribonuclease, which is certainly not the case, but it does mean that with care one can hope to find conditions under which there is little or no loss of activity during the course of an experiment.

In the increasingly frequent studies of artificial mutant variants of enzymes no such comforting assumptions have much validity. These proteins have not experienced natural selection for stability, and there is no basis for thinking that they can survive the conditions of an assay without loss of activity. On the contrary, mutant enzymes are often less stable than their wild-type counterparts. For example, recent studies by circular dichroism of mutant forms of a mouse protein related to carbonic anhydrase showed considerable variations in stability and three-dimensional structure among proteins of similar catalytic activity (18).

It follows that testing for progressive inactivation should be a routine procedure in any study of a mutant or modified enzyme, especially when establishing correlations between structure and function. For such studies to be valid it is important to know whether a lower measured activity is due to a genuinely decreased activity or simply to problems due to inactivation during the assay that might be trivial or nonexistent for the native enzyme on which the assay was developed.

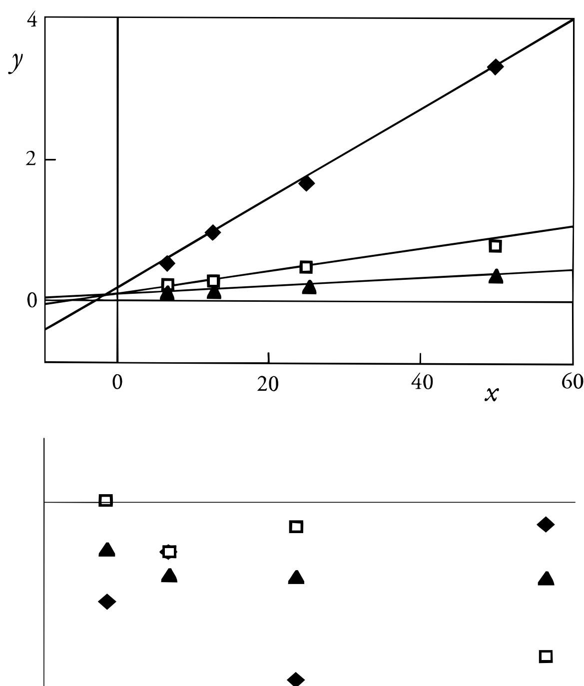

FIG. 5. Pooling data from multiple sources into a single residual plot. The plot shown at the top is based on a study of inhibition of napsin by pepstatin (10), and as each line contains only four data points it is difficult to draw any conclusions if they are considered one at a time. Combining all the data into a single residual plot, shown below, makes it clear that virtually all of the deviations are negative.

One type of information about stability and homogeneity may come from determination of the active site molarity of an enzyme, as described by Brocklehurst and colleagues in this issue (19), but here we are concerned with simple techniques that can be applied routinely without requiring special equipment or knowledge.

Fortunately an easy test for inactivation has been known for many years, though rather infrequently applied. Building on observations from the beginning of the 20th century (20, 21), Selwyn (22) developed the modern form of the test, after pointing out that under normal assay conditions with a large excess of substrates over enzyme most mechanisms can be described by a rate equation of the type

$$\frac{dp}{dt} = e_0 f(p), \qquad [1]$$

in which p is the concentration of product after time t, e_0 is the total enzyme concentration, and f is a homogeneous function of p that can (in principle) be derived from the rate equation. The fact that it may be difficult to derive and rather complicated (because all other concentrations need to be expressed in terms of p by means of stoichiometric relationships) is not important. The equation can be integrated as follows:

$$e_0 t = \int \frac{dp}{f(p)}. \qquad [2]$$

The solution to the integral on the right-hand side does not need to be known. It is sufficient to know that it does not depend on e_0, and thus has exactly the same dependence on p at all values of e_0. Experimentally, this means that plots of p against $e_0 t$ should be superimposable if the starting assumptions are true. An example, based on results of Michaelis and Davidsohn (21), is illustrated in Fig. 6. Even though the progress curve appears not to pass through the origin (though the reason for this should be clear from inspection) and to show some sigmoidicity, the points obtained at three different enzyme concentrations all lie on the same curve.

As Selwyn (22) discussed, there are various reasons why the results from such an experiment might differ from those in Fig. 6. The simplest is that the enzyme loses (or gains) activity during the assay, which would make it invalid to treat e_0 as a constant during the integration. More complex problems would arise if the rate was not strictly proportional to the total enzyme concentration, which would happen, for example, if the

enzyme existed in two or more states of association (monomer, dimer, and tetramer, for example) in equilibrium with one another and with different catalytic activities. Deutscher (23) described an example of this, and his results are reproduced in a discussion of Selwyn's test elsewhere (24). From the point of view of checking the validity of an assay and for the stability of the enzyme the essentials are to check that the points from different enzyme concentrations fall on the same curve, as in Fig. 6, and to investigate possible causes if they do not.

Here we have been concerned with one of the simplest kinds of information available from the time course of an enzyme-catalyzed reaction. As Duggleby (15) discusses elsewhere in this issue, however, quantitative analyses of time courses can provide considerably more.

A CASE STUDY IN FAULTY GRAPHIC TREATMENT

Scatchard and Eadie–Hofstee plots

Earlier this article referred to the increasing tendency in the literature to conceal primary data from the reader by giving only the final results of the data processing. Of course, the literature would be impossibly unwieldy if all primary data were published, and

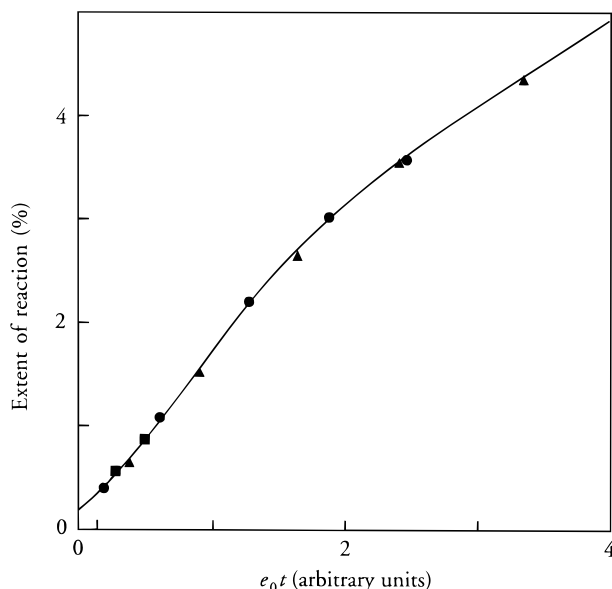

FIG. 6. Checking for enzyme inactivation during an assay. If the assumptions needed for an enzyme assay to be valid are true, plots of accumulated product against time multiplied by the total enzyme should yield points that lie on a single curve regardless of the enzyme concentration at which the assays were carried out. The example is based on data for invertase (19) that satisfy this criterion, for three invertase concentrations in the ratio 0.4(■):1(●):2(▲).

all modern journals rightly refuse to allow this, but there is a large difference between including all of the primary data in a paper and including one or two examples to illustrate the quality of the experiments and to allow the reader to form an impression of whether the analysis was correctly done.

Worries about lack of primary data in the modern literature might be unfounded if there were abundant evidence that the great majority of experiments were analyzed appropriately, but unfortunately this is not the case. The Scatchard plot for studying ligand binding to proteins is a revealing example. Essentially the same points apply to the Eadie–Hofstee plot for analyzing enzyme kinetic data, and in this account the term *Scatchard plot* is taken to apply to both.

Some years ago we examined a considerable number of papers published in major biochemical journals around 1993–1994 that presented Scatchard plots in which the data points could not be interpreted as straight lines. Of these, around 30% used computational methods and included sufficient evidence to suggest that they had been applied appropriately and correctly. The majority, however, used graphic methods in ways that were demonstrably incorrect and likely to produce significant errors both in the qualitative interpretation of the data and in the values of any binding or kinetic constants estimated. The commonest error was to draw a straight line through points that obviously demanded a curve (as in Fig. 1, which shows a more recent example of this error), but it was also common to draw two straight lines and then analyze these separately as if each had been obtained in a separate experiment in which the points lay on a straight line. We only encountered one paper (25) in which nonlinear Scatchard data were analyzed by means of a valid graphic method, that of Rosenthal (26). This method uses the property that when the observed binding is the result of binding at several different sites the effects in the Scatchard plot are additive along lines drawn through the origin, as discussed and illustrated in more detail elsewhere (24). By contrast, the approach of treating two straight lines as if they had been observed separately leads to gross errors.

The point here, however, is not to discuss in detail how the Scatchard plot should be used, but to use it as an example to illustrate the very frequent misuse of graphic methods in modern biochemical work and as a warning not to assume that conclusions derived from unseen primary data have been reached by a valid route.

CONCLUDING REMARKS

The near-universal use of computing in the modern laboratory has encouraged the view that any approach to data analysis other than a completely automated one is outdated. However, this ignores the fact, well recognized in the statistical literature, that the eye is much better at detecting anomalies than any computer program, which can test only for conditions that have been foreseen by the programmer. Some graphic processing of experimental data will therefore remain essential for the foreseeable future. Plotting residual errors rather than direct observations provides a very powerful way of focusing attention on the degree of agreement between an experiment and any proposed interpretation of it, and thus helps to ensure that significant discrepancies do not pass unnoticed.

REFERENCES

1. Cleland, W. W. (1967) *Adv. Enzymol.* **29,** 1–32.
2. Cornish-Bowden, A. (1995) Analysis of Enzyme Kinetic Data, Oxford Univ. Press, Oxford.
3. Chambers, J. M., Cleveland, W. S., Klein, B., and Tukey, P. A. (1983) Graphical Methods for Data Analysis, Wadsworth, Belmont, CA.
4. Tukey, J. W. (1977) Exploratory Data Analysis, p. 126, Addison–Wesley, Reading, MA.
5. Mosteller, F., and Tukey, J. W. (1977) Data Analysis and Regression, p. 388, Addison–Wesley, Reading, MA.
6. Ellis, K. J., and Duggleby, R. G. (1978) *Biochem. J.* **171,** 513–517.
7. Anscombe, F. J. (1973) *Am. Stat.* **27,** 17–21.
8. Cha, J.-H., Brooke, J. S., Ivey, K. N., and Eidels, L. (2000) *J. Biol. Chem.* **275,** 6901–6907.
9. Höld, K. M., Sirisoma, N. S., Ikeda, T., Narahashi, T., and Casida, J. E. (2000) *Proc. Natl. Acad. Sci. USA* **97,** 3826–3831.
10. Nilsson, J., Grahn, B., and Heby, O. (2000) *Biochem. J.* **346,** 699–704.
11. Schauer-Vukasinovic, V., Bur, D., Kitas, E., Schlatter, D., Rossé, G., Lahm, H.-W., and Giller, T. (2000) *Eur. J. Biochem.* **267,** 2573–2580.
12. Portaro, F. C. V., Santos, A. B. F., Cezari, M. H., Juliano, M. A., Juliano, L., and Carmona, E. (2000) *Biochem. J.* **347,** 123–129.
13. Stabile, H., Curti, B., and Vanoni, M. A. (2000) *Eur. J. Biochem.* **267,** 2720–2730.
14. Nash, P., Barry, M., Seet, B. T., Veugelers, K., Hota, S., Heger, J., Hodgkinson, C., Graham, K., Jackson, R. J., and McFadden, G. (2000) *Biochem. J.* **347,** 375–382.
15. Duggleby, R. G. (2001) *Methods* **24,** 168–174.
16. Cárdenas, M. L., and Cornish-Bowden, A. (1993) *Biochem. J.* **292,** 37–40.
17. Lebreton, S., and Gontero, B. (1999) *J. Biol. Chem.* **274,** 20879–20884.
18. Elleby, B., Sjöblom, B., Tu, C., Silverman, D. N., and Lindskog, S. (2000) *Eur. J. Biochem.* **267,** 5908–5915.
19. Brocklehurst, K., Resmini, M., and Topham, C. M. (2001) *Methods* **24,** 153–167.
20. Hudson, C. S. (1908) *J. Am. Chem. Soc.* **30,** 1564–1583.
21. Michaelis, L., and Davidsohn, H. (1911) *Biochem. Z.* **35,** 386–412.
22. Selwyn, M. J. (1965) *Biochim. Biophys. Acta* **105,** 193–195.
23. Deutscher, M. P. (1967) *J. Biol. Chem.* **242,** 1123–1131.
24. Cornish-Bowden, A. (1995) Fundamentals of Enzyme Kinetics, Portland Press, London.
25. Saunier, B., Pierre, M., Jacquemin, C., and Courtin, F. (1993) *Eur. J. Biochem.* **218,** 1091–1094.
26. Rosenthal, H. E. (1967) *Anal. Biochem.* **20,** 525–532.