Analysis of Enzyme Specificity by Multiple Substrate Kinetics†

Volker Schellenberger,‡ Ronald A. Siegel,§ and William J. Rutter*,‡

Hormone Research Institute, Box 0534, and Departments of Biochemistry and Biophysics, Pharmacy, and Pharmaceutical Chemistry, Box 0446, University of California, San Francisco, California 94143

Received October 27, 1992; Revised Manuscript Received January 25, 1993

ABSTRACT: Multiple approaches for screening large sets of compounds for a specific function are of growing interest. The use of substrate mixtures to characterize the specificity of enzymes has been limited so far to compounds with similar kinetic parameters, because the data were analyzed by applying the kinetics of two competing substrates. In this study we introduce a statistical method for the analysis of reactions with many competing substrates which makes use of the specific features of multiple substrate kinetics. It is assumed that the relative concentrations of all substrates in a mixture can be monitored by high-performance liquid chromatography or a similar technique. Relative second-order rate constants, i.e., k_{cat}/K_M values, can be calculated for all substrates in the mixture from the resulting data set. The calculation uses the fact that there is a relationship between the concentrations of all pairs of substrates in the mixture. As a result, the precision of the calculated parameters is increased and the range of kinetic constants that can be obtained from one experiment is considerably expanded. Simulations demonstrate that the precision in the kinetic parameters increases with the number of substrates in the mixture. In fact, estimation of ratios of rate constants can be improved (or made possible) for substrates with order of magnitude differences in reactivity by adding "dummy" substrates with intermediate reactivities, even though the rate constants for dummy substrates are themselves of no intrinsic interest.

In recent years there have been many reports dealing with the screening of large mixtures of compounds for a certain biological function (Devlin et al., 1990; Houghten et al., 1991; Lam et al., 1991; Owens et al., 1991; Scott & Smith, 1990). Peptides are especially promising substrates as methods have been developed which allow the efficient synthesis of peptide mixtures of well-defined variability [see Zuckermann et al. (1992) and references therein]. Mixtures of substrates have been used to study the specificity of enzymes (Berman et al., 1992; Birkett et al., 1991; Petithory et al., 1991; Rangheard et al., 1989), but so far, the kinetic rules developed only for *pairs* of competitive substrates were applied for data analysis. As a consequence, kinetic constants obtained from a substrate mixture were within 1 order of magnitude.

We introduce here a statistical method which takes full advantage of the kinetics of reactions with multiple competing substrates. It is assumed that decreasing substrate concentrations can be followed by high-performance liquid chromatography, gas chromatography, or a similar technique. On the basis of the resulting data, relative second-order rate constants are calculated for all substrates in the mixture. Advantages and limitations of the method have been studied by computer simulation.

THEORY

Kinetics of Reactions with Competing Substrates. We consider a class of reactions involving a single enzyme and a number (n) of competing substrates S_i, where $i = 1, ..., n$. Substrates are consumed according to second-order kinetics:

$$\frac{d[S_i]}{dt} = -k_i[E][S_i] \qquad i = 1, ..., n \tag{1}$$

where [E] is the concentration of the free (unbound) enzyme, $[S_i]$ is the concentration of the ith substrate, and k_i is the second-order rate constant (k_{cat}/K_M for Michaelis–Menten kinetics). Because multiple substrates are competing for the enzyme, and since the substrate concentrations change with time, [E] is difficult to know at any time if the k_i's are yet unknown. Also, circumstances may exist in which the *total* (bound and unbound) enzyme concentration in the reaction mixture is unknown. Thus it may be difficult or impossible to determine the absolute k_i values. However, we will presently show that *ratios* of second-order rate constants can be obtained.

Consider two substrates S_i and S_j. For each substrate, eq 1 will apply. Direct inspection of eq 1 leads to the conclusion that $d[S_i]/d[S_j] = k_i[S_i]/k_j[S_j]$. This relation can be rearranged and integrated to yield

$$\frac{1}{k_i} \ln \frac{[S_i]_t}{[S_i]_0} = \frac{1}{k_j} \ln \frac{[S_j]_t}{[S_j]_0} \tag{2}$$

where $[S_i]_0$ and $[S_i]_t$ denote substrate concentrations at times zero and t, respectively. Again rearranging eq 2, we can determine the ratio k_i/k_j by measuring the two substrate concentrations at times zero and t.

From eq 2 it is apparent that no knowledge of enzyme concentration, bound or free, is needed to determine the ratios k_i/k_j. This is convenient since, as noted before, [E] is often impossible to determine. In addition, eq 2 can be generalized to include measurements of substrate concentrations at any two distinct times. Moreover, time itself does not enter explicitly into the calculation of k_i/k_j; the actual time value is unimportant so long as measurements for all substrates are taken simultaneously. Eq 2 depends upon accurate measurements of substrate concentrations in order to calculate k_i/k_j. In practice, it is useful to measure substrate concentrations at several time points and to utilize statistical estimation procedures, one of which will now be described.

Statistical Model. We assume that substrate concentrations are measured at times t_q, $q = 1, ..., m$. The measured concentration of substrate i at time t_q will be denoted $[S_i]_q$.

† This work was supported by a postdoctoral fellowship of the Deutsche Forschungsgemeinschaft to V.S.

‡ Hormone Research Institute and Department of Biochemistry and Biophysics.

§ Departments of Pharmacy and Pharmaceutical Chemistry.

0006-2960/93/0432-4344$04.00/0 © 1993 American Chemical Society

These measured concentrations will contain errors. Thus we rewrite eq 2 as

$$\Theta_q = \frac{1}{k_1}\ln\frac{[S_1]_q}{[S_1]_0} + \epsilon_{1,q} = \frac{1}{k_2}\ln\frac{[S_2]_q}{[S_2]_0} + \epsilon_{2,q} = \ldots = \frac{1}{k_n}\ln\frac{[S_n]_q}{[S_n]_0} + \epsilon_{n,q} \qquad q = 1, \ldots, m \quad (3)$$

where Θ_q is a parameter to be estimated at each time point t_q and $\epsilon_{i,q}$ is the error in estimating Θ_q that is associated with the measurement $[S_i]_q$. Since we do not wish to give preference to any particular time point, including $t = 0$, we also take $[S_i]_0$ to be a parameter to be estimated. Finally, we recall that the scheme permits only ratios of rate constants to be estimated, so that one of the k_i's must be set to an arbitrary value, say $k_1 = 1$. Also, since time does not explicitly enter into the equations, the $t = 0$ point is also arbitrarily placed, so that at least one $[S_i]_0$ or Θ_q must be fixed to an arbitrary value. We choose to fix $\Theta_1 = 0$.

Fitting this model to the data requires estimation of the parameters $k_2, \ldots, k_n$, $[S_1]_0, \ldots, [S_n]_0$, and $\Theta_2, \ldots, \Theta_m$, i.e., $2n + m - 2$ parameters from mn measurements. Parameter estimates are obtained by minimizing the objective function[1]

$$\Omega = \sum_{q=1}^{m}\sum_{i=1}^{n}\left(\frac{[S_i]_q}{[S_i]_0} - e^{k_i\Theta_q}\right)^2 w_{i,q} \quad (4)$$

where $w_{i,q}$ is a weighting function that may depend on measurements and/or parameter estimate values. Our simulations gave excellent results with $w_{i,q} = 1$.

The $[S_i]_0$'s and Θ_q's are "nuisance" parameters, i.e., parameters whose values are not of direct concern but which must be estimated along with the desired parameters (here the k_i's). The definition of $[S_i]_0$ is clear. It is interesting to observe that Θ_q is actually an estimate of the negative integral of free enzyme concentration between times 0 and t_q. This is seen by noting that, in the absence of measurement error and for any substrate S_i,

$$\Theta_q = \frac{1}{k_i}\ln\frac{[S_i]_q}{[S_i]_0} = \frac{1}{k_i}\int_{[S_i]_0}^{[S_i]_q}\frac{d[S_i]}{[S_i]} = -\int_0^{t_q}[E]\,dt \quad (5)$$

where use is made of eq 1 to obtain the last equality.

EXPERIMENTAL PROCEDURES

The minimization procedure described in the previous section will be useful for estimating relative rate constants for reactions with multiple substrates. In order to demonstrate its efficiency, the procedure has been coded in Turbo Pascal (Borland, CA) and tested with simulated data (program will run on any PC clone; available on request).

Computation of Kinetic Constants. The required minimization of the objective function Ω was carried out using the Davidon–Fletcher–Powell algorithm (Press et al., 1989). The initial estimates of the $[S_i]_0$ parameters were set to the simulated $[S_i]_1$ values (see below). (With real data, measured $[S_i]_1$ values should be used.) In order to prevent the algorithm from being trapped in a local minimum, a two-stage minimization was utilized. In the first stage, Ω was minimized with the $[S_i]_0$'s fixed at their initial values. In the second stage the minimization was over all parameters.

Good estimates of the parameters are prerequisite for convergence. In the Appendix a semi-ad hoc algorithm for calculating initial estimates of $k_2, \ldots, k_n$ and $\Theta_2, \ldots, \Theta_m$ is presented.

Simulation of Experimental Data. We have tested the algorithm for a set of competing reactions that are first order in substrate concentrations. For simplicity we set all $[S_i]_0 = 1$, but we tried various sets of rate constants k_i. We then chose one of the substrates, $i = 1$, as the reference substrate whose rate constant was set to $k_1 = 1$. Substrate concentrations were then generated by the formula

$$[S_i]_q = \max(e^{-k_i t_q} + \delta_\beta, 0) \quad (6)$$

where t_q is the qth time point ($t_1 = 0$) and δ_β is a random normal deviate with mean zero and standard deviation β. In other words, the true values of the substrate concentrations were assumed to decay exponentially, and the measurement corruption was assumed to be of a constant level. Also, negative measurements were set to zero. It will be noticed that in the absence of the random term δ_β, eq 2 holds for all pairs of substrate concentrations.

The choice of time points was motivated by two considerations. First, we desire that the relative changes in at least two concentrations be sufficiently large that they can be detected over the measurement noise. Thus, we choose t_2 such that for at least two substrates, $[S_i]_2/[S_i]_1 < 0.9$. Second, substrate concentrations should not be too low. We adopt the rule that data will be practically meaningless if its value lies below 3 standard deviations of the measurement noise (3β). Thus, we take the last time point t_m to be the time at which all but two concentrations are expected to fall beneath 3β. Between t_2 and t_m we chose logarithmically spaced time points:

$$t_q = t_2\left(\frac{t_m}{t_2}\right)^{(q-2)/(m-2)} \qquad q = 3, \ldots, m \quad (7)$$

This choice of intermediate time points is based on the observation that substrate concentrations decrease exponentially, so that increasingly longer time intervals between successive time points are needed if significant changes in concentration are to be observed.

For a given set of "true" kinetic constants, 500 simulated data sets were generated. For each data set, values of the estimates that minimize Ω (eq 4) were calculated. In the present work the weighting factors $w_{i,j}$ were set to unity. Means and standard deviations of the parameter estimates were calculated using well-known formulas.

RESULTS

Number of Samples. An example of the time course of a reaction with five competing substrates is shown in Figure 1. True $[S_i]_q$ values are displayed. A 3% measurement noise level was assumed as illustrated by the bands around the true data. Figure 2 shows the range of the computed relative kinetic constants which are obtained for the reaction shown in Figure 1 as a function of m, the number of time points. It can be seen that there is virtually no bias in the estimation scheme. The precision of the kinetic parameters increases with the

[1] An alternative and perhaps more direct objective function based on eq 3 would be $\Omega' = \sum_{q=1}^{m}\sum_{i=1}^{n}[\Theta_q - (1/k_i)\ln([S_i]_q/[S_i]_0)]^2 w'_{i,q}$, where the of the weights $w'_{i,q}$ depends on assumptions made about the variances of the error terms $\epsilon_{i,q}$. If all terms had equal variance, then we would have $w'_{i,q} = 1$. Such would be the case if the error was proportional to the measurement, the proportion being the same for all substrates. In the present work, error was assumed to be at a constant absolute level for all substrates and for all measurements; i.e., error was assumed to be dominated by baseline noise (see below). In this case, the objective function chosen in eq 4 works extremely well with the weighting functions set to unity, as shown in the results section. For this reason, we have chosen to stick with eq 4.

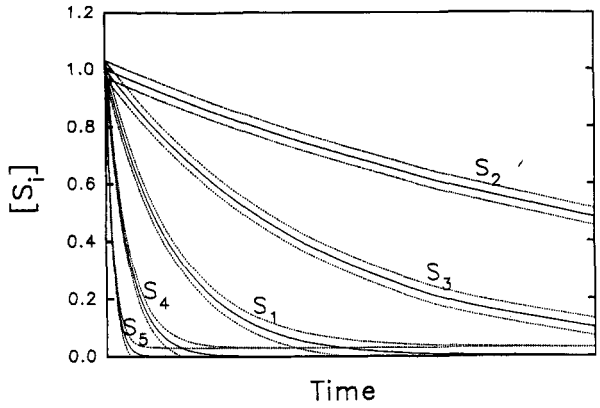

FIGURE 1: Time course of a reaction with five competing substrates. The reaction was assumed to be first order with respect to the substrate concentrations. Dotted lines indicate intervals of 3% standard deviation that were used in most simulations. $k_1 = 1$, $k_2 = 0.1$, $k_3 = 0.32$, $k_4 = 3.2$, and $k_5 = 10$.

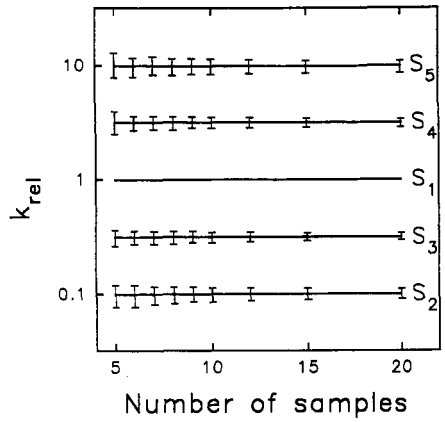

FIGURE 2: Influence of the number of samples on the precision of the kinetic parameters. Reaction conditions and error intervals were as shown in Figure 1. Samples were taken at the following time points: $t_1 = 0$, t_2 at $[S_4]_2/[S_4]_1 = 0.9$, t_m at $[S_2]_m/[S_2]_1 = 0.1$, and other reaction times as given by eq 7. Solid lines indicate the true values of the kinetic constants. Displayed are error bars representing 1 standard deviation from the mean rate constants observed in 500 simulations.

number of analyzed samples. Also, for all reactions simulated here it was observed that 8–10 samples provided sufficient information to calculate all relative rate constants.

Range of Kinetic Constants. Consider a mixture of two substrates with a large difference in their rate constants. In this case the concentration of the poorer of the two substrates only begins to decline when the concentration of the other substrate is close to zero. Consequently, relative rate constants calculated on the basis of eq 2 will be prone to large errors. The larger the ratio between the kinetic constants of two substrates, the lower is the precision of the calculated estimates. Figure 3 shows how the standard deviation of the calculated rate constant increases with the ratio between these rate constants. Obviously, relative rate constants can be calculated with good precision even from poor raw data (standard deviation of 5% in the measured substrate concentrations) as long as both substrates have similar rate constants. Rate constants differing by 1.5 logarithmic units can be determined if substrate concentrations are measured with a standard deviation of 1%.

Accuracy Increases with the Number of Substrates. It follows from the arguments given in the previous paragraph that no relative rate constants can be reliably estimated for a mixture of two substrates differing by two or more

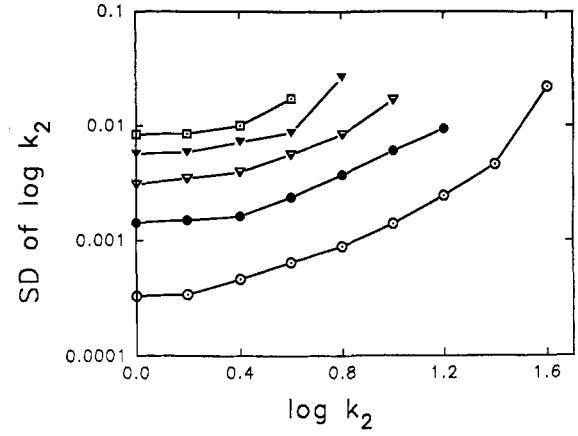

FIGURE 3: Precision of the calculated rate constants in dependence of the error in the substrate concentration. Reactions with two competing substrates were simulated. The standard deviations of the calculated ratios of second-order rate constants is plotted versus the ratio of the *true* kinetic constants. Ten samples were analyzed from the reactions with $t_1 = 0$, t_2 at $[S_1]_2/[S_1]_1 = 0.99$, t_m at $[S_1]_m/[S_1]_1 = 0.01$, and other reaction times as given by eq 7. Standard deviation of the noise that was added to the true data: ⊙, 1%; ●, 2%; ▽, 3%; ▼, 4%; ⊡, 5%. For values of log (k_2) larger than given in the figure, the minimization failed to converge at least once in 500 simulations.

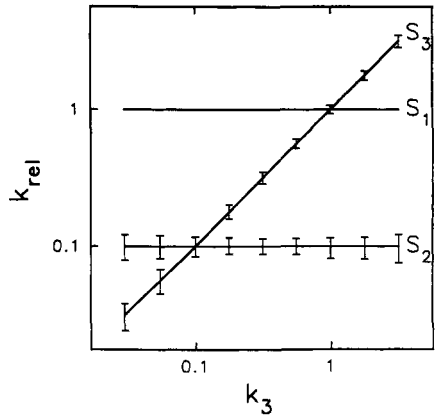

FIGURE 4: Effect of a third substrate S_3 on the precision in estimating the relative rate constant k_2. True value of $k_2 = 0.1$. Displayed are error bars representing 1 standard deviation from the mean value of estimated k_2 and k_3 as a function of the true k_3. Ten samples were taken, and 3% measurement error was assumed. Other conditions were as in Figure 2.

logarithmic units in their true rate constants (unless substrate concentrations can be quantified with extremely high precision). On the other hand, relative rate constants could be determined for the reaction whose kinetics are shown in Figure 1, even though two substrates are included which differ by a factor of 100 in their rate constants. Clearly, the data measured for the substrates with intermediate rate constants permit calculation of the rate constants for the best and worst substrate in the mixture.

We simulated reactions of two substrates S_1 and S_2 differing in their rate constants by one logarithmic unit. Assuming a standard deviation of 3% in the raw data, the ratio of kinetic constants could be determined with a standard deviation of 0.018 logarithmic units. We then repeated the simulation including a third substrate S_3. Figure 4 shows that the calculation of the ratio k_2/k_1 is significantly improved when the value of k_3 is between k_1 and k_2. In this case the substrates S_1 and S_2 can be determined indirectly by comparing them both with S_3. The standard deviation of k_2 was reduced to

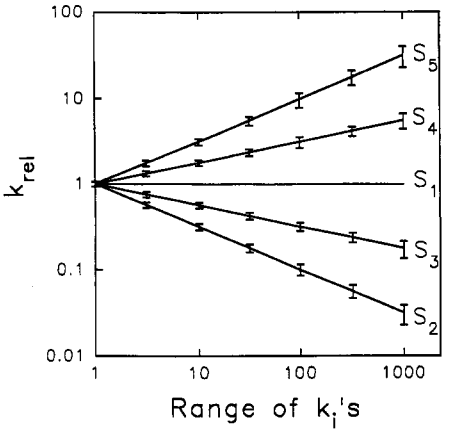

FIGURE 5: Precision of the calculated rate constants in dependence of the range of true kinetic constants. Rate constants were spaced logarithmically over the range. Ten samples were taken from reactions with five substrates and 3% error in the substrate concentrations. Other conditions were as in Figure 2.

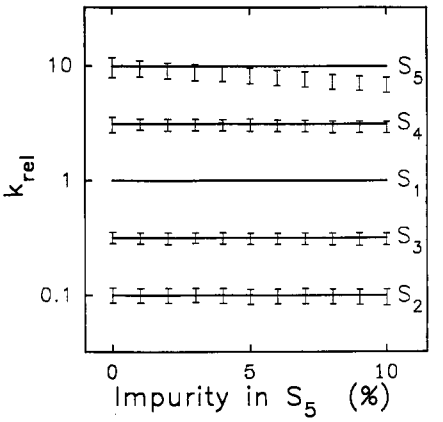

FIGURE 6: Influence of an impurity in the substrate mixture on the calculated rate constants. The reaction conditions were as shown in Figure 1 but a constant value was added to the true concentration of the best substrate (S_5, k_5 = 10) in the mixture. Error (3%) was added to the substrate concentration, and 10 samples were taken from the reactions. Other conditions were as in Figure 2.

0.013 log units when k_3 was exactly between k_1 and k_2.[2] These findings clearly demonstrate the advantage of our statistical approach, which calculates relative rate constants using the whole set of substrate concentrations. If the relative rate constant of a substrate is calculated by correlating its concentration only with the concentration of the standard substrate, then the information about the concentrations of all other substrates is wasted.

Obviously, increasing the number of substrates in the reaction mixture not only increases the amount of data gained per experiment but also improves the precision of the calculated rate constants. Figure 5 shows that a mixture of five substrates with kinetic constants spanning a range of 3 orders of magnitude can be analyzed with good precision when the substrates' rate constants are spread logarithmically over the whole range. In many cases it will be useful to include a number of "dummy" substrates with intermediate rate constants into a mixture of substrates with large differences in their rate constants, even though the reactivities of the dummy substrates are of no intrinsic interest.

Impurities in the Substrate Mixture. Substrates, especially if synthesized by multiple methods, may contain impurities which are difficult to detect in complex mixtures. Therefore we decided to study the influence of a contamination in the substrate mixture on the calculated parameters.

Let us assume that one substrate contains an impurity. As a result the concentration of this substrate in the mixture would be lower than anticipated. This has no influence on the observed rate constants, so long as the substrate and the impurity are analytically separated, because only relative substrate concentrations are used in the calculation. Even if the impurity binds to the active site of the enzyme (as a competitive inhibitor or a competitive substrate) this has no influence on the calculated rate constants, as all reactions are slowed down. The observed kinetic parameters will only be affected by an impurity if this compound coelutes with a substrate in the mixture (for instance a nonreactive diastereomer of a substrate). In this case one would observe that the peak area of one substrate does not decrease to zero but to the value which is caused by the impurity. Figure 6 shows the effect of a nonreactive impurity that coelutes with the fastest substrate, S_5 (true k_5 = 10) on the estimation of the various rate constants. As the concentration of the impurity increases, the estimate of k_5 shows increasing downward bias. The estimate of k_4 is degraded by a lesser but still noticeable amount. The result for k_5 is explained by realizing that a constant added to a decaying exponential will be fit best by another exponential with a smaller decay constant. The result for k_4 in turn is attributed to the influence of the estimate of K_5 on the former. This influence will be much attenuated for the other substrates whose rate constants are far away from k_5, as is clear from Figure 6.

DISCUSSION

We have proposed a statistical method for the analysis of multiple substrate kinetics. Relative second-order rate constants can be calculated with good precision for all substrates in a mixture. The kinetics of very different substrates can be analyzed if substrates with intermediate rate contrasts are included in the mixture. These features should make multiple substrate kinetics a powerful approach to characterize the specificity of enzymes or to analyze the effect of point mutations on the specificity of an enzyme. In order to test the efficacy of the procedure introduced here, we generated stimulated data based on a model assuming first-order kinetics for all substrates. As noted from the outset, however, such kinetics do not always occur in real problems. Free enzyme concentrations often vary throughout a reaction because substrate concentrations are changing. We stress that the procedure does not rely on the type of kinetics, because time does not enter the calculation. The method should be applicable in any situation where eq 2 describes the set of reactions.

In the present work we have ignored the weighting factors, $w_{i,j}$ in eq 4. In many statistical analyses such weighting factors are utilized to "focus" a fit on time points where the data has the highest precision. Weighting functions should be chosen to minimize the bias of a parameter estimation. In the simulation results displayed in Figures 2 and 4–6, it appears that minimal bias occurs, except when an impurity interferes with the assay of one of the substrates. The latter cause cannot be fixed completely by changing weights. (One could add another nuisance parameter to account for the impurity if necessary, and the parameter could be estimated by observing

[2] A closer examination of Figure 4 reveals a slight but noticeable bias in the estimate of k_2/k_1 when k_3 exceeds k_1. Similarly, bias in k_2/k_1 is discerned when $k_3 < k_2$. Apparently, estimation for two substrates can be degraded by the presence of a third substrate if the first two substrates are far apart in their reactivity and the third substrate's reactivity does not lie in between.

the final value of the assay for the contaminated substrate.) Thus, except for the case of the nonseparable impurity, it appears that our unweighted minimization provides adequately unbiased estimates of the rate constants.

Reaction kinetics can be analyzed by measuring either product formation or substrate consumption. The latter variant should be preferred with multiple substrate kinetics whenever possible. Otherwise, exact quantitative information about the composition of the substrate mixture and calibration of the analytical method for the various products is required. Alternatively, $[S_i]_q$ in the objective function (eq 4) has to be replaced by $[S_i]_0 - [P_i]_q$, where $[P_i]_q$ is the concentration of the product which is formed from the substrate S_i at time t_q. This modified function can only be used if all substrates are almost completely consumed during reaction and it may be necessary to include an appropriate weight into the objective function.

Intrinsically, most factors which reduce the precision in conventional kinetics (no competing substrates) have no influence on multiple substrate reactions. No information about the concentrations of the substrates or the enzyme is required. The results are not affected by inactivation of the enzyme which may occur during incubation. As already discussed, samples do not have to be analyzed at exactly defined time points. In fact, the samples taken from one reaction can be analyzed in a random order without any influence on the calculated parameters. Competitive inhibitors or substrates do not influence the result. Furthermore, errors resulting from sample handling (dilution, derivatization) can be avoided by addition of an internal standard.

Finally, we note that the simulations were performed for certain idealized cases in order to demonstrate the efficacy and potential limitations of the procedure. We suggest that the simulations can have a second function when real data are gathered. Once having minimized the objective function Ω to obtain relative rate constants using experimental data, one can use the simulations to obtain confidence limits on these parameter estimates by incorporating knowledge of the measurement noise characteristics. The presence of bias can also be detected by simulations.

ACKNOWLEDGMENT

We thank Professor Lewis B. Sheiner for helpful suggestions.

APPENDIX

Initial parameter estimates are provided by the following semi-ad hoc procedure. The measured values of $[S_i]_1$ were assigned to the $[S_i]_0$'s. Estimates of k_i were provided by minimization of eq 8, which directly evolves from eq 3:

$$F = \sum_{q=1}^{m} \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} \left(\ln \frac{[S_i]_q}{[S_i]_0} - \frac{k_i}{k_j} \ln \frac{[S_j]_q}{[S_j]_0} \right)^2 w_{q,i} w_{q,j} \quad (8)$$

where $w_{q,i}$ is defined as

$$w_{q,i} = \frac{\left| \ln \frac{[S_i]_q}{[S_i]_0} \right|}{\left(\frac{\sigma_{[S_i]_q}}{[S_i]_q} \right)^2 + \left(\frac{\sigma_{[S_i]_0}}{[S_i]_0} \right)^2} \quad (9)$$

The standard deviations $\sigma_{[S_i]_q}$ and $\sigma_{[S_i]_0}$ for the measured substrate concentrations were estimated as 3% of $[S_i]_0$. The weight of the first data point $w_{1,i}$ was calculated as the mean of the weights for the next two data points, $w_{2,i}$ and $w_{3,i}$. This model converges very stably using $k_i = 1$ as starting value for all rate constants and keeping $[S_i]_0$ fixed. Estimates of Θ_q can be calculated from the kinetic parameters assuming first-order kinetics by

$$\Theta_q = -\sum_{i=1}^{n} \frac{1}{k_i} \ln \frac{[S_i]_q}{[S_i]_0} w^2_{q,i} k_i / \sum_{i=1}^{n} w^2_{q,i} k_i \quad (10)$$

REFERENCES

Berman, J., Green, M., Sugg, E., Anderegg, R., Millington, D. S., Norwood, D. L., McGeehan, J., & Wiseman, J. (1992) *J. Biol. Chem. 267*, 1434–1437.

Birkett, A. J., Soler, D. F., Wolz, R. L., Bond, J. S., Wiseman, J., Berman, J., & Harris, R. B. (1991) *Anal. Biochem. 196*, 137–143.

Devlin, J. J., Panganiban, L. C., & Devlin, P. E. (1990) *Science 249*, 404–406.

Houghten, R. A., Pinilla, C., Blondelle, S. E., Appel, J. R., Dooley, C. T., & Cuervo, J. H. (1991) *Nature 354*, 84–86.

Lam, K. S., Salmon, S. E., Hersh, E. M., Hruby, V. J., Kazmierski, W. M., & Knapp, R. J. (1991) *Nature 354*, 82–84.

Owens, R. A., Gesellchen, P. D., Houchins, B. J., & DiMarchi, R. D. (1991) *Biochem. Biophys. Res. Commun. 181*, 402–408.

Petithory, J. R., Masiarz, F. R., Kirsch, J. F., Santi, D. V., & Malcolm, B. A. (1991) *Proc. Natl. Acad. Sci. U.S.A. 88*, 11510–11514.

Rangheard, M. S., Langrand, G., Triantaphylides, C., & Baratti, J. (1989) *Biochim. Biophys. Acta 1004*, 20–28.

Scott, J. K., & Smith, G. P. (1990) *Science 249*, 386–390.

Zuckermann, R. N., Kerr, J. M., Siani, M. A., Banville, S. C., & Santi, D. V. (1992) *Proc. Natl. Acad. Sci. U.S.A. 89*, 4505–4509.