

Estimating the turnover number in enzyme kinetic reactions using transient and stationary state data

Sibel Uludag-Demirer¹, Jorge Duran², Robert D. Tanner^{3*}

¹Villanova University, Civil and Environmental Engineering, Villanova, USA, ²INCYTH/CTUA, Buenos Aires, Argentina,

³Vanderbilt University, Department of Chemical and Biomolecular Engineering, Nashville, USA

Substrate and product concentration data obtained by simulating enzyme-substrate reaction rate equations were used to test two proposed kinetic rate constant estimation techniques in this study. In the first technique, the turnover number, k_3 , was calculated using early transient time domain data, which are difficult to obtain experimentally. The technique used an iterative approach to calculate k_3 with a pair of data and the value of k_3 could be retrieved with 35% error. The second technique calculated k_3 using stationary domain data and the value of k_3 could be retrieved with less than 5% error. This second technique also offered internal consistency in the calculation of k_3 by calculating k_3 both from the intercept and the slope of the linear plot derived in this study. A series of sensitivity analyses was conducted to understand the robustness of the second technique in estimating k_3 from simulated data to the changes in the reaction rate constants (k_1 , k_2 , and k_3) and the initial concentration of enzyme used for simulation. It was found that the second technique generally worked well in the estimation of k_3 , except for the simulated data for fast substrate conversions such as in the large k_3 and $[E]_0$ cases. This latter method, thus, shows promise for the use of late time experimental substrate/product concentration data to obtain k_3 . Exclusively using late time data avoids the need for difficult and expensive rapid early time measurement techniques for estimating k_3 . Once a reasonable estimate for k_3 is obtained, the initial enzyme value can easily be determined from the maximum velocity constant established from fitting the Michaelis-Menten or Briggs-Haldane equations to substrate and product stationary state domain (late time) data. While the first technique can estimate k_3 with only one point in the transient domain, it is suggested that the second method generally be favored since it only requires late-time stationary domain data and appears to be more accurate.

Uniterms: Enzyme. Kinetics/rate constant. Enzyme reactions.

Dados de concentração de substrato e de produto obtidos por simulação de equações de velocidade de reação enzima-substrato foram usados neste estudo para testar duas técnicas para estimar constantes cinéticas. Na primeira técnica, a constante, k_3 , foi calculada utilizando os dados de domínio do tempo inicial de transição, que são difíceis de serem obtidos experimentalmente. A técnica usou uma aproximação iterativa para calcular k_3 , cujo valor pôde ser estimado com erro de 33%. A segunda técnica calculou k_3 usando dados de domínio no estado estacionário e o valor de k_3 pôde ser estimado com erro de 5%. Esta segunda técnica também ofereceu uma consistência interna no cálculo de k_3 , por calculá-lo tanto pela intersecção quanto pela inclinação da reta derivada deste estudo. Uma série de análises de sensibilidade foi realizada para avaliar a robustez da segunda técnica na estimativa de k_3 utilizando dados que foram simulados quanto às mudanças nas constantes de taxa de reação (k_1 , k_2 e k_3) e na concentração inicial de enzima. Foi encontrado que a segunda técnica, em geral, proporcionou boa estimativa de k_3 , exceto para os dados simulados para as conversões rápidas de substrato, como no caso de valores elevados de k_3 e de $[E]_0$. Este último método, portanto, mostra ser promissor quando se usam dados experimentais tardios da concentração de substrato/produto para obter k_3 . O uso dados de tempo tardio evita a necessidade do uso de técnicas difíceis e caras na medida de tempo iniciais para estimativa de k_3 . Uma vez que é obtida uma estimativa razoável de k_3 , o valor inicial da enzima pode ser facilmente determinado a partir da constante de velocidade máxima estabelecida por ajuste das equações de Michaelis-Menten ou de Briggs-Haldane e partir de dados de substrato e de produtos no estado estacionário (tempo tardio). Enquanto a primeira

*Correspondence: R. D. Tanner. Chemical and Biomolecular Engineering Department, Vanderbilt University, Nashville, TN 37235, USA. e-mail: Robert.d.tanner@vanderbilt.edu.

técnica pode estimar k_3 com somente um ponto no regime transiente, sugere-se que o segundo método seja melhor uma vez que ele requer apenas dados do estado estacionário tardio e parece ser mais preciso.

Unitermos: Enzimas. Cinética/constantes. Reações enzimáticas.

INTRODUCTION

The reactions between an enzyme and a substrate are usually analyzed by using the substrate and product concentrations measured over time under isothermal conditions. The Michaelis-Menten or Briggs-Haldane model equations are applied to describe such time varying data for the simplest enzyme substrate interactions with the limitation that they exclude the initial (transient) time profiles. The simplest and most often used enzyme substrate model is the reaction scheme with one substrate, one intermediate, and one product with the recovery of the enzyme (Equation 1) and this model reaction is sufficient to describe the behavior of many enzyme-substrate systems (Bernhard, S., 1968).



In Equation (1), k_1 , k_2 , and k_3 are the reaction rate constants. In general, the initial substrate concentration ($[S]_0$) is known while the initial (and total) enzyme concentration ($[E]_0$) in the reaction system is unknown. Therefore, $[E]_0$ behaves as another parameter to be estimated from the model. Based on the reaction scheme given in Equation (1), the rate expressions of each species follow from the law of mass action:

$$d[S]/dt = -k_1[E][S] + k_2[ES]; [S] = [S]_0 \quad (2)$$

$$d[E]/dt = -k_1[E][S] + (k_2 + k_3)[ES]; [E] = [E]_0 \quad (3)$$

$$d[ES]/dt = k_1[E][S] - (k_2 + k_3)[ES]; [ES] = 0 \quad (4)$$

$$d[P]/dt = k_3[ES]; [P] = 0 \quad (5)$$

where $[S]$, $[E]$, $[ES]$, and $[P]$ indicate the molar concentrations of the substrate, enzyme, enzyme-substrate complex and product, respectively, in the solution at any reaction time t . Initial concentrations of $[ES]$ and $[P]$ are zero. Typical simulations of $[S]$ and $[P]$ are shown in Figure 1. There are also two independent linear molar conservation equations, as follows:

$$[E] + [ES] = [E]_0 \quad (6)$$

$$[S] + [ES] + [P] = [S]_0 \quad (7)$$

Equation (6), an algebraic molar balance on the en-

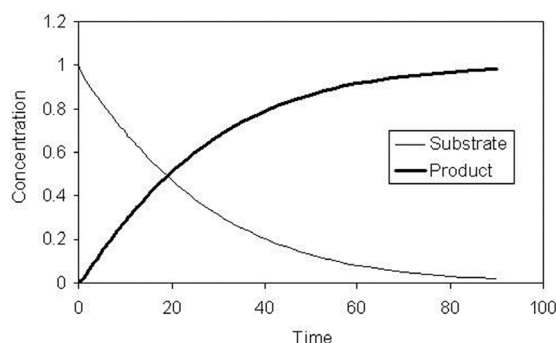


FIGURA 1 - Simulated concentrations of substrate and product for a reaction medium containing $[S]_0 = 1$, $[E]_0 = 0.1$ relative normalized molar units with rate constants of $k_i = 1$ ($i = 1, 2, 3$).

zyme concentration in its two forms, can be obtained by summing Equations (3) and (4) and then integrating, while Equation (7), an algebraic molar balance on the substrate and its products, follows from summing Equations (2), (4), and (5) and then integrating.

The Briggs-Haldane model equation, Equation 8, can be used to fit late time (stationary-state) experimental data ($[S]$ and $[P]$) to calculate the model parameters: the Michaelis constant (K_m) and the maximum velocity term (V_{max}).

$$(d[P]/dt) = v = V_{max} [S] / (K_m + [S]) \quad (8)$$

where $K_m = (k_2 + k_3)/k_1$; $V_{max} = k_3[E]_0$ and v = product formation rate or velocity

Equation (8) is developed from the stationary-state assumption of $d[ES]/dt = 0$, and Equation (6). The other widely used correlation equation is the Michaelis-Menten Equation, where the K_m term is replaced by $K_e = k_2 / k_1$. That equation is derived by using the equilibrium between $E+S$ and ES , neglecting the subsequent reaction to P and E , along with Equation (6). It turns out that the Briggs-Haldane equation provides a lower bound for $d[P]/dt$ as a function of $[S]$ and the Michaelis-Menten equation an upper bound as a function of $[S]$ (Rutherford, B.J. *et al.*, 2007). To calculate the model parameters, a nonlinear regression analysis can be performed. Typically, a linear arrangement of reciprocal v (product formation rate) plotted against reciprocal $[S]$ in the form of a Lineweaver-Burk plot is generally preferred for recovering the model parameters. To calculate the initial concentration of enzyme, $[E]_0$, k_3 should be estimated separately.

The turnover number, k_3 , can be determined directly from Equation (5), if both $[ES]$ and $[P]$ are measured at specified time intervals. Usually, however, $[ES]$ is not measured, and k_3 must then be determined from the early time (transient) data from measurements of $[S]$ and $[P]$ at specific time intervals. $[ES]$ is generally not very accurate if obtained from $[S]$ and $[P]$ measurements and $[S]_0$ using Equation (7), since it is the small difference between large numbers with relatively large measurement error. Assuming that at very early reaction times, $[P]$ is negligible with respect to $[S]_0$, i.e., $[P] \rightarrow 0$, then Equation (7) reduces to:

$$[ES] = [S]_0 - [S] \quad (9)$$

Now, replacing $[ES]$ in Equation (5) with (9), Equation (10) can be obtained as previously shown (Tanner *et al.*, 1977).

$$d[P]/dt = k_3 ([S]_0 - [S]) \quad (10)$$

With the measurements of product and substrate concentrations corresponding to the transient period, it is possible to determine k_3 for $[S] \leq [S]_0$ from the slope of the expression in Equation (10) by plotting $d[P]/dt$ against $([S]_0 - [S])$ for a positive slope or against $[S]$ for a negative slope, as pictured by the line (qualitatively) near the peak in Figure 2. The estimated value of k_3 from Equation (10) increases as $[S] \rightarrow [S]_0$, so that the tangent at $[S]_0$ gives the most accurate value for k_3 . However, it is difficult to obtain the early time product formation rates since enzyme-substrate reactions are generally too fast, and there are typically only one or two data points taken (at best) in the transient period. Measurement errors come into play as well. Since $[S]_0 - [S]$ is small in early time, the variance associated with $[S]$ dominates this difference and becomes controlling. It is noted that, without the assumption of small $[P]$, Equation (10) would be replaced by: $d[P]/dt = k_3([S]_0 - [S] - [P])$. This rigorous replacement equation will be invoked in the iterative method which is described in next section.

Turnover Number Estimation Using The Maximum Product Formation Rate

Numerically integrating the model given by Equations (2) through (5), employing $k_i = 1$ ($i = 1, 2, 3$), $[E]_0 = 0.1$, $[S]_0 = 1$ in dimensionless molar units and using a linear, triangular, implicit integration scheme for a step size of 0.01 time units up to 90 time units leads to the set of simulated data presented in Figure 1 (Tanner, 1972). The simulated data, accurate to three decimal points, can be divided

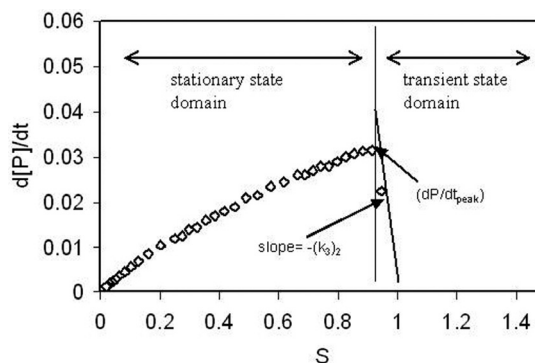


FIGURE 2 - A cross plot of the data in Figure 1 for the estimation of the turnover number (K_3) using the maximum product formation rate by iteration. The subscript 2 refers to the second iteration in the proposed estimation method.

into two domains: **transient** and **stationary state** (Figure 2). The transient domain extends from the initiation of the reaction ($t = 0$; $[S] = [S]_0$), to the moment ($t = t_{\text{peak}}$) when the maximum value is registered for the product generation rate ($d[P]/dt_{\text{peak}}$ at $[S] = S_{\text{peak}}$). The stationary state domain includes the data from $t = t_{\text{peak}}$ to $t = \infty$, i.e., from $d[P]/dt_{\text{peak}}$ to the end of the reaction, $d[P]/dt = 0$ as $t \rightarrow \infty$ and $[S] \rightarrow 0$ (Figure 2). Typically, there is only a single data point simulated in the transient state domain when $[E]_0$ equals $1/10$ of $[S]_0$ if the point at the boundary between transient and stationary state is excluded (Tanner, 1972).

In this section, the estimation of k_3 with only one pair of data points is illustrated using an iterative technique. Since special equipment like the stopped flow apparatus is generally required for multiple measurements of $[S]$ and $[P]$ in the transient state domain, it is not always possible to use Equation (10) to estimate k_3 . In addition, it is difficult to precisely define the maximum rate of product formation $d[P]/dt_{\text{peak}}$ with a corresponding S value of S_{peak} to define the transient state domain precisely and then to accurately use it with the initial time point, $[S] = [S]_0$ and $d[P]/dt = 0$ to estimate k_3 . Therefore, an iterative procedure should be developed to determine $d[P]/dt_{\text{peak}}$ and k_3 . A point near $d[P]/dt_{\text{peak}}$ and S_{peak} is assumed to be the largest estimated $d[P]/dt$ value and is used as the starting point for the iteration. The pair of $d[P]/dt_{\text{peak}}$ (pointed out in Figure 2) and S_{peak} measurements can then be used to draw the line connecting the peak and the origin (below the slope as the transient line shown in Figure 2) and k_3 can be calculated from the slope of the line. The first iteration in this example provided a k_3 estimate of 0.45 in reciprocal time units compared to 1, the value originally employed in the simulated model. In an effort to improve the accuracy of this calculation, an iteration procedure is developed to sharpen the estimate of the value of the slope. Since the most accurate slope in

the transient domain is the tangent at $[S]=S_0$ and since this is equal to k_3 , the negative slope has to be steeper than that of the line connecting the peak and the initial value, such as the slope of the line in the transient domain pictured in Figure 2. This slope has a value of $-(k_3)_2$ for the second iteration, in which the subscript stands for the number of iterations, as shown in Figure 2. The subscript or k_3 refers to the second iteration. To enhance the convergence of the results, an “overshoot segment” is defined as the product of k_3 and P_{peak} and it was added to the $d[P]/dt|_{\text{peak}}$ previously employed to determine k_3 . It is hypothesized that this additive term will bring the value of $d[P]/dt|_{\text{peak}}$ closer to the desired $d[P]/dt$ value on the initial tangent line. The theoretical development of the method and the iteration sequence now follows (in the six indicated parts):

- Initial estimate of k_3

Initial estimate of k_3 starts with using Equation (10). This equation was derived by assuming that $[P]$ in Equation (7) is small with respect to $[S]_0 - [S]$. $(d[P]/dt)_0$ represents a real experimental value, which is typically at the largest $[S]$ value in the stationary state domain, and it is equal to the observed $(d[P]/dt)_{\text{peak}}$. It is the starting rate in the beginning iteration. The value of $[S]_{\text{peak}}$, which is the substrate concentration corresponding to the maximum product formation rate, will remain the same at each step of the iteration. It therefore follows that the first estimate for k_3 is:

$$(k_3)_1 = (d[P]/dt)_0 / ([S]_0 - [S]_{\text{peak}}), \quad (11)$$

where the subscript 1 indicates the first iteration.

Note that the slope is negative for the tangent line in Figure 2 since the $d[P]/dt$ value is zero at the point of initial substrate. The two $[S]$ terms in the denominator of Equation (11) are switched from the usual order for a tangent line in order to make (k_3) positive in Equation (11).

- Calculation of the overshoot segment using Equations of (5) and (7):

Since: $(d[P]/dt)_1 = (k_3)_1([S]_0 - [S]_{\text{peak}}) + (k_3)_1[P]_{\text{peak}}$, the overshoot segment for the calculation of $(d[P]/dt)_1$ is equal to $(k_3)_1 [P]_{\text{peak}}$, where $[P]_{\text{peak}}$ is the concentration of product at the maximum product formation rate. $[P]_{\text{peak}}$ like $[S]_{\text{peak}}$ remains constant throughout the iteration.

- Definition of the overshoot point or $(d[P]/dt)_2$:

The first iterative equation thus becomes:

$$(d[P]/dt)_1 = (d[P]/dt)_{\text{peak}} + (k_3)_1 [P]_{\text{peak}} \quad (12)$$

- Calculation of the next estimate of k_3 :

$$(k_3)_2 = (d[P]/dt)_1 / ([S]_0 - [S]_{\text{peak}})$$

- Calculation of the new overshoot segment:

$$(\text{overshoot segment})_2 = (k_3)_2 [P]_{\text{peak}}$$

- Calculation of the new overshoot point:

$$(d[P]/dt)_2 = (d[P]/dt)_{\text{peak}} + (k_3)_2 [P]_{\text{peak}}$$

In general, we have two iteration equations as follows:

$$(k_3)_n = (d[P]/dt)_{n-1} / ([S]_0 - [S]_{\text{peak}}) \quad (13)$$

$$(d[P]/dt)_n = (d[P]/dt)_{\text{peak}} + (k_3)_n [P]_{\text{peak}} \quad (14)$$

where n is the iteration number. Here, $(d[P]/dt)_{\text{peak}}$ was a fixed number. This iterative procedure is performed until the value of k_3 appeared to converge. The iteration procedure was stopped when the retrieved value of k_3 was within a difference of ± 0.01 .

This procedure was applied to the simulated data (Figure 2) and the value of k_3 appeared to converge to 0.65 after five iterations, but clearly more iterations would have been in order since the known outcome of 1 should be reached. This procedure retrieved the value of k_3 with a 35% error. Combining Equations (13) and (14) and generalizing for the n 'th case gives: $(d[P]/dt)_n = (d[P]/dt)_{\text{peak}} [1+x+(x)(x)+\dots] = (d[P]/dt)_{\text{peak}} (1-x)$ for x less than 1, in this geometric progression. Here, $(x = P_{\text{peak}} / ([S]_0 - [S]_{\text{peak}}))$, which is less than 1, since $[P] / ([S]_0 - [S])$ is less than 1 in general from (Equation (7) because $[ES]$ is positive). Thus, $(d[P]/dt)_n$ converges so that Equation (14) converges for large n . It is noted that the iterative value of k_3 in this procedure used only a single data point. Stopping at five iterations was in keeping with another iterative procedure in the transient domain in which three to five iterations typically resulted in the desired outcome. That procedure was based on the use of orthogonal polynomials and Picard's iteration technique and used essentially the same simulated data (Tanner, R.D., 1972). Since there is only one data point in this method, another technique was explored for the estimation of k_3 using the much more accessible stationary state domain data (See Figure 2).

Turnover Number Estimation Using Stationary State Data

A combination of Equations (5) and (7) can lead to an alternative expression which, when suitably arranged, can be employed to estimate k_3 directly (without iteration) using “ordinary” $[S]$ and $[P]$ data in the stationary state

domain, and without requiring the difficult-to-obtain transient data, although such additional data would enhance the fitting technique. Starting with Equation (5), the [ES] complex concentration can be expressed in terms of [P] and [S] employing Equation (7) to obtain:

$$\frac{d[P]}{dt} = k_3([S]_0 - [S] - [P]) \quad (16)$$

After division of Equation (16) by [P] the following mathematical expression is obtained:

$$\frac{1}{[P]} \frac{d[P]}{dt} = -k_3 + k_3 \left\{ \frac{([S]_0 - [S])}{[P]} \right\} \quad (17)$$

This form of the product formation rate equation is very helpful in estimating the k_3 constant, because it bears a linear relationship between $1/[P] d\{P\}/dt$ and $([S]_0 - [S])/[P]$ and k_3 can be calculated both from the intercept and the slope of the plot. Linearity in Equation (17) may also be useful in directly testing the original premise shown in Equation (1), as stated in Equations (2)-(5). An example of such a linear plot is shown in Figure 3 for [S] and [P] simulated data from Figure (1) with 3 significant figures. Note that as $t \rightarrow \infty$, the line goes to the point (1,0). This is seen, in general, for the “x” coordinate going to one as [S] goes to zero and [P] goes to $[S]_0$ as $t \rightarrow \infty$ and for the “y” coordinate going to zero as [P] goes to $[S]_0$ and $d[P]/dt$ goes to zero as $t \rightarrow \infty$. Internal consistency of these two estimated k_3 values serves as a check on the scatter and the time spacing of the [S] and [P] data as well as the integrity of the differenced value, $\Delta[P]/\Delta t$, which is used to estimate $d[P]/dt$. The linear regression analysis of the plot shown in Figure 3 resulted in four place k_3 values as 0.9883 using the slope and 0.9875 using the intercept. The goodness of fit of the method could be evaluated in two ways: (a) Accuracy, the difference with the known parameter value used to simulate the data, and (b) Consistency, the internal difference between the slope and intercept estimates of k_3 . The first one is related to the accuracy of the method that is the average of the k_3 values, 0.988 (in keeping with the three place simulated data), was found to be very close to 1, the value used to create [S] and [P]. The difference with the known value for this particular regression was measured considering the percent difference between the average k_3 calculated and the known k_3 value and estimated as 1.24%, which is close to the desired 0% for highest accuracy. The internal difference was evaluated by calculating the percent difference between the slope and the intercept values for k_3 and estimated as 0.08%, which is negligibly small as compared to a desired 0% for desired internal consistency.

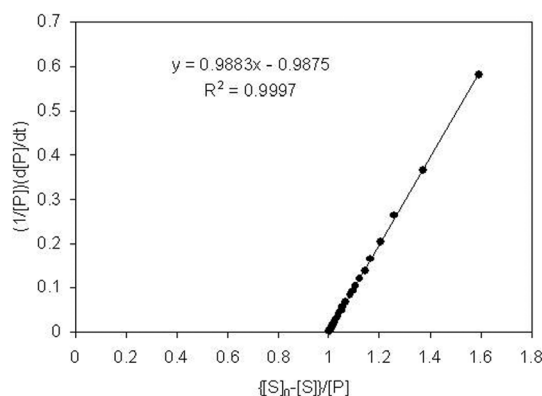


FIGURA 3 - The estimation of the turnover number, K_3 , from the slope and intercept using three place stationary state domain data obtained from the simulation underlying Figures 1 and 2. The smallest time point is closest to the upper right corner.

The two parameters in the Briggs-Haldane model equation, Equation (8), V_{\max} and K_m , are generally estimated using stationary state data with the Lineweaver-Burk expression, Equation (18). This equation is the reciprocal of Equation (8) and is used to formulate a linear form to use as an alternative to the nonlinear regression of the hyperbolic Equation (8):

$$\frac{1}{\left(\frac{dP}{dt}\right)} = \frac{K_m}{V_{\max}} \frac{1}{[S]} + \frac{1}{V_{\max}} \quad (18)$$

In estimating the K_m and V_{\max} constants with the Lineweaver-Burk plot, linear regression analysis of the straight line was performed (Figure 4). The values of the parameters are calculated as 0.968 concentration units and 0.0588 moles per unit time for K_m and V_{\max} respectively. Since $V_{\max} = k_3[E]_0$, the initial concentration of the enzyme, which is typically unknown in the system, was calculated as 0.06 concentration units which is different by 40% from the product of the original parameter values used to

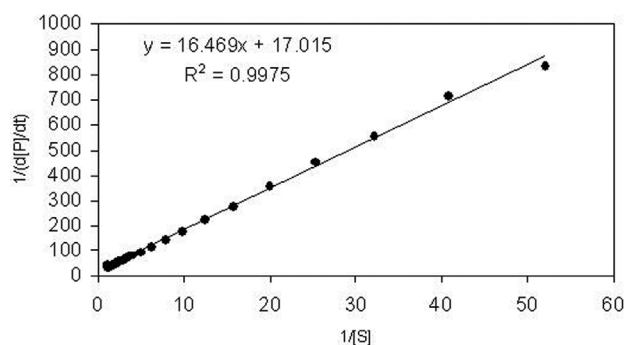


FIGURA 4 - Lineweaver-Burk Plot of the simulated data. The early time points are close to the lower left corner of this graph.

simulate in the differential equations. The K_m estimate was off by 50%. These are large errors when compared to those incurred in estimating k_3 from Equation (17).

Sensitivity Analysis

The enzyme-substrate single intermediate model previously introduced in Equation (1) contains four typically unknown parameters; k_1 , k_2 , k_3 , and $[E]_0$, given $[S]_0$. Here, the accuracy of the parameter estimation method comparing Equation (17) to simulated data in determining k_3 and its sensitivity to the variation of the different parameters used in simulating the data is evaluated. Also considered in this analysis is the effect of the number of significant figures employed in the linear regression procedure.

The computer software, Polymath, was employed to simulate Equations (2) through (5) in order to determine the evolution of substrate and product concentration trajectories with time by using an implicit modified Euler method. These simulated data were then used to estimate the value of k_3 both from the slope and intercept using Equation (17). The estimated value of k_3 for a particular fit was then calculated by averaging the k_3 values obtained from the slope and from the intercept. Only the data corresponding to the stationary state domain, following the $(d[P]/dt)_{\text{peak}}$ for late times, were utilized in the proposed fitting Equation (17). This limitation to only stationary state data was followed even though Equation (17) applies in general to include the early time transient as well as the stationary data, because in practice only stationary data are readily available.

The goodness of fit of the method was evaluated in terms of accuracy and consistency. The absolute difference with the known value, k_3 , for a particular regression was determined to develop the accuracy using the percent difference between the average k_3 calculated and the known k_3 . The internal difference was evaluated by calculating the percent difference between the slope and the intercept values for k_3 , to determine the consistency of the method.

Effect of k_1 variation on the estimate for k_3

Values of 0.01, 0.1, 1, and 10 for k_1 , which were used to simulate three place data for the estimation process, were selected in this part of the study to estimate the effect of k_1 variation on k_3 estimation. The remaining parameters, k_2 , k_3 , $[E]_0$, and $[S]_0$ were kept equal to 1, while the initial product concentration $[P]_0$ was kept at zero throughout these case studies. Note that $[E]_0$ is now 1 rather than 0.1 as in Figure 1.

When k_1 was larger than 1, the percent absolute difference of the k_3 value determined with the linear regression (Equation 17) with respect to the known value was generally unacceptable (above 10% error). However, for k_1 equal to 0.1 or less, that difference decreased to levels less than 10% (Figure 5). Equation (7) itself is useful in helping us understand why the absolute value of the estimated k_3 increases as the value of k_1 increases. Equation (7), when divided through by $[P]$ yields Equation (19):

$$\frac{[ES]}{[P]} + 1 = \frac{([S]_0 - [S])}{[P]} \quad (19)$$

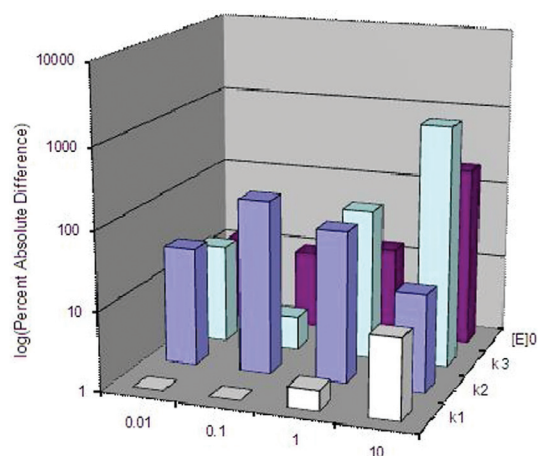


FIGURA 5 - Percent absolute difference of the k_3 value estimated using Equation (17) and the known given values of k_1 , k_2 , k_3 and E_0 (shown above) used to simulate the data used for the estimation process.

It can be seen, at least for late time data using Equation (8), that as $K_m = (k_2 + k_3)/k_1$ increases goes towards zero and (dP/dt) increases. This means that the rate at which (P) is formed increases, up to $(P)=1$ at $t=\infty$. This is when $[P]$ goes to $[S]_0=1$, when (ES) and (S) go to zero referring to Equation (7). As t approaches ∞ , the term (dP/dt) in Equation (17) eventually tends towards zero since $d[P]/dt = k_3 [ES]$ and $[ES]$ goes towards zero while $[P]$ goes to $[S]_0$. Thus, the term $([S]_0 - [S])/[P]$ drives to 1. This is also observed in Figure 3. and can serve as an internal check step for fitting Equation (17) as seen in Figure (3).

Since higher k_1 values correspond to higher rates of substrate consumption, the number of spread out data points available for a reasonable least squares fit of the data decreases (in both the substrate and product domains) affecting the accuracy of the linear regression. For very large k_1 , for example, the points tend to collapse to one point (zero for $[S]$ and one for $[P]$). With fewer data points in hand, it can be seen that it would be more and more difficult to obtain both accurate slope and intercept values

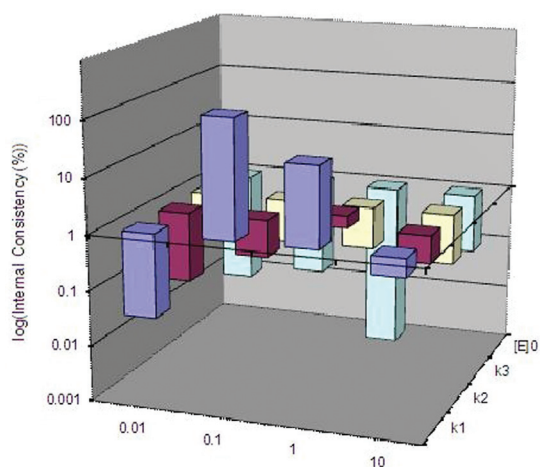


FIGURA 6 - Internal consistency of the turnover estimation method using Equation (17) and the known given values of k_1 , k_2 , k_3 , and E_0 (shown above) used to simulate the data used for the estimation process. Note that the z axis scale is much less than that in Figure 5. The internal consistency was above 100 for the $k_2 \geq 0.001$ and 0.1 values.

from Figure 3, since the early time high $[S]$ and low $[P]$ values tend to define the entire fitting domain. Interestingly, the internal consistency between the k_3 estimated from the slope and the k_3 from the intercept becomes more acceptable as k_1 approaches 10 (Figure 6).

Effect of k_2 variation on the estimate for k_3

Four simulations were performed with k_2 values at 0.01, 0.1, 1 and 10 while k_1 , k_3 , $[E]_0$, and $[S]_0$ were kept equal to 1, keeping $[P]_0$ at zero. The percent absolute differences of the estimated k_3 values with the known value, 1, were very high for the k_2 range tested in the study (Figure 5). The simulation performed with the lowest k_2 value (0.01) yielded only two data points to estimate k_3 using Equation (17) due to the fast conversion of substrate to product in the early transient domain with this small back reaction term. Moreover, there was very poor internal consistency in the calculation of k_3 from the slope and the intercept of Equation (17). Therefore, the percent absolute difference is not shown for this case in Figure 6. Interestingly as the reverse reaction became more important (larger k_2), there was a decrease in the absolute difference of k_3 from the known value as the value of k_2 increased, from 0.1 to 10 (Figure 5). This decrease was due to the increase in the number of data points spread over the entire stationary state domain. At the same time, the internal difference values of k_3 obtained from the slope and the intercept of the Equation (17) became insignificant as shown in Figure 6.

Effect of k_3 variation in the equations used to simulate the data on the estimate for k_3 from those data

Again, four different values in the range of 0.01-10 for the turnover number, k_3 , were selected to test the effect of the magnitude of k_3 on the performance of the method in predicting the known value. The remaining parameters, k_1 , k_2 , $[E]_0$, and $[S]_0$ were kept equal to 1, while $[P]_0$ was made equal to zero. The absolute value of the difference between the k_3 determined with Equation (17) and the known value remained below 0.5 % for k_3 values equal or lower than 0.1, indicating good accuracy of the method in that range. However, the difference with the known value increased for higher k_3 values, reaching, for instance, 80 % for k_3 equal to 1 (Figure 5), showing that the method becomes significantly less accurate at high k_3 values. An explanation for this breakdown in the estimations is that there is the lack of data in the stationary state domain for these cases with high k_3 values and this leads to large fitting errors. This lack of data points is the direct result of having the conversion rate of the substrate become so fast that the product concentration reaches 1 concentration unit rapidly in the early transient domain with the reaction essentially reaching completion. Thus there are few remaining data points left for fitting in the stationary state domain for fitting when these early time points are not measurable in practice.

It is observed that the internal difference is negligible in the estimations of higher k_3 values (Figure 6). The internal difference is calculated as 27% for the case where the k_3 value is 0.01. This is the lowest value tested in the study, indicating that at low k_3 values (i.e. 0.01) the consistency of the method tends to break down leading to a high variance in the estimated value for k_3 .

Effect of $[E]_0$ variation on the estimate for k_3

The initial concentration of enzyme ($[E]_0$) was varied in four simulation runs, adopting values of 0.01, 0.1, 1, and 10 concentration units. The remaining parameters, k_1 , k_2 , k_3 , and $[S]_0$ were kept equal to 1, while $[P]_0$ was kept equal to zero. When $[E]_0$ had values equal to or smaller than 1, the percent difference of the k_3 value estimated with Equation (17) and the known value remained within 10% (Figure 5). At a high $[E]_0$ concentration (10 concentration units), that difference started to increase, reaching 191%. It seems that at very high initial enzyme concentrations the accuracy of the method deteriorates due to the very fast consumption of the substrate available in the system and lack of fittable data in the stationary-state (late time) domain. The resulting fitting is like putting a line through one point. The high

internal consistency of the method was found to be almost independent of the initial enzyme concentration (Figure 6).

Influence of the number of significant figures employed

Analytical results are often expressed with two significant figures rather than with the three significant figure data employed in the above sensitivity analyses. In order to observe the variation introduced in the previous $[E]_0$ sensitivity analysis employing more "standard quality" data, the values originally employed were rounded to two significant figures. As expected, the trends observed in absolute values were similar, but their magnitudes increased for each case. An absolute difference of 12.2% with the known value of k_3 was observed with the two significant figure substrate and product concentration data compared to the absolute difference of 11.6% with the known value of k_3 substrate and product concentration data with five significant figures.

CONCLUSIONS

Two kinetic rate estimation techniques were studied using simulated data of substrate and product concentrations. The data were obtained by integrating the standard enzyme-substrate nonlinear differential equations (Equations 2-5) with the rate constants (k_1, k_2, k_3) equal to 1 and $[S]_0=1 = [E]_0$. The turnover number, k_3 , was calculated in the first method using early transient domain data with an iterative method. k_3 was found to equal 0.65 in this method after five iterations. This value lies within the bound of $0.4 k_3$ 2.1 estimated in an earlier study using the same data and a simple polynomial fitting technique (Tanner, 1972). A second method for the estimation of k_3 was developed using stationary-state (late time) data only. The sensitivity of this method to changes in the reaction rate constants, including the turnover number, and the initial enzyme concentration was tested. It was clear that the method broke down in those cases of high rates of substrate consumption corresponding to very high enzymatic levels, a relatively unusual situation in traditional enzyme kinetic experiments. When all of the substrate is converted to product ($[P]_{\max} = 1, [S]_0=1$) in the transient domain, there are no or limited data in the subsequent in time stationary domain to be used in the linear regression calculations. However, the method yielded very good estimates of k_3 in these cases containing sufficient data in the stationary phase. By its generally good accuracy for all but the highest k_3 and very high enzyme level cases the applicability of this method now provides a good reason for measuring both substrate and product concentration data in experimental enzy-

me kinetic systems in order to estimate k_3 , the turnover number, directly. Once k_3 is established, the total enzyme concentration is readily calculated from the V_{\max} term of the Briggs-Haldane or the Michaelis-Menten equation.

Estimating k_3 from late-time data in this new method, described by Equation (17), obviates the need for difficult-to-obtain initial rate data. The required late-time data or substrate and product concentrations needs to be of such quality as to provide reasonably accurate product rates, $d[P]/dt$, and normalized product rates, $(1/[P])d[P]/dt$, as well as the substrate difference, $[S]_0 - [S]$ and its division by $[P]$. This new method for estimating k_3 not only provides a relatively simple and easy way to estimate $[E]_0$, but when k_1 and k_2 are estimated by other techniques, it offers the parameters to directly validate the original differential equations, Equations (2) and (5), against experimental data.

ACKNOWLEDGEMENT

We thank one of the reviewers of this paper for the comment:

"Most biochemists employ the method of initial rates to obtain steady-state data for parameter estimation (Briggs-Haldane equation). Complete progress curves are valuable for biochemists to generate, in addition to initial rate data, because comparison of the fitted parameters can reveal anomalies for further study."

REFERENCES

- BERNHARD, S. *The structure and function of enzymes*. New York: W.A. Benjamin, 1968. p.73.
- RUTHERFORD, B. J.; HAHS, D.; TANNER, R.D. Quantifying the position of the stationary state trajectory between the Michaelis-Menten and Briggs-Haldane equations as a function of the initial enzyme concentration. *Chem. Eng. Comm.*, v.194, n.7, p.867-888, 2007.
- TANNER, R.D. Estimating kinetic rate constants using orthogonal polynomials and Picard's iteration method. *Ind. Eng. Chem. Fundamentals*, v.11, n.1, p.1-8, 1972.
- TANNER, R. D.; LOO, A. C.; SHISLER, J. L.; REED, M. W.; ROWLETT, R. D.; MORRIS, J. W.; SCHLOSSNAGLE, G. W.; OVERLEY, J. R. Mapping the lag phase and bounding the growth phase in fermentation reactions. *AIChE Symp. Ser.*, v.78, n.167, p.55-65, 1977.

Received for publication on 12th June 2009.

Accepted for publication on 27th September 2009.