Error sensitivity and optimization of steady-state kinetic parameters using multidimensional chemical kinetic analysis

Joseph S. Eskew[1], Christopher G. Connell[2], and Jared C. Cochran[1,]*

[1]Department of Molecular & Cellular Biochemistry, Indiana University, Bloomington, IN 47405, USA.
[2]Department of Mathematics, Indiana University, Bloomington, IN 47405, USA.

[*]Corresponding author: Jared C. Cochran, Department of Molecular and Cellular Biochemistry, Indiana University, 405C Simon Hall, 212 S. Hawthorne Dr., Bloomington, IN 47405 USA, Tel: 812-855-6935, Fax: 812-856-5710, Email: jcc6@indiana.edu

Short Title: Multidimensional chemical kinetics analysis

Abbreviations: [1] MM, Michaelis-Menten; E, enzyme; S, substrate; P, product; errA, additive error; P_i, inorganic phosphate; C_{est}, estimated constant from nonlinear least squares fitting; C_{exp}, the expected constant; v_f, final instantaneous velocity; v_o, initial velocity; CI, confidence interval (95%); AIC, Akaike information criterion; BIC, Bayesian information criterion; MT, microtubule.

Abstract

Enzyme behavior has been described using the Michaelis-Menten mechanism. The analysis of extended time domains provides a means to extract the Michaelis-Menten constants through direct fitting of raw data. We have developed a scheme for determining Michaelis-Menten rate constants by appropriate fitting of multidimensional experimental data sets to the closed form of the Michaelis-Menten model. We considered how varying parameters in experimental data affect the accuracy of the remaining parameter estimates. We determine how to improve experimental design to achieve a given accuracy, relative to the amount of intrinsic or external error. We analyze this scheme on data sets built around 20 hypothetical and 2 natural enzymes (kinesin and apyrase) to test error sensitivity in different parameter regimes. Overall, we provide evidence that our data fitting regime will tolerate significant experimental error in the raw data and still converge on the four Michaelis-Menten constants.

Introduction

Enzymes are proteins that catalyze specific biochemical reactions in living cells to generate needed products for metabolic and/or replicative processes. Scientists have been studying enzymes for decades and a wealth of information is available for many different types of enzymes that exist in all domains of life (Eukarya, Bacteria, and Archaea). The Michaelis-Menten (MM)[1] mechanism has been the foundation for studying the steady-state kinetic behavior of enzymatic reactions for over 100 years [1]. In the basic MM scheme (Equation 1), a substrate (S) binds reversibly with a free enzyme (E) to form a non-covalent enzyme-substrate complex, which reacts irreversibly to yield free product (P) and the regenerated free enzyme. One can easily demonstrate the hyperbolic relationship between the initial reaction velocity and substrate concentration through mathematical derivation under a given set of basic assumptions [2]. However, these assumptions are not always valid for enzymatic systems and/or experimental design regimens. This is especially true for extended time domains that allow for both substrate exhaustion and product inhibition of enzyme activity through binding competition. Therefore, more complex mechanisms must be employed to consider the biochemical steps for chemistry (ES⇌EP) and product release and re-binding to free enzyme (EP⇌E + P).

MM steady-state experiments focus on capturing data for product formation during the initial velocity time regime with variant substrate concentrations. These raw data are subsequently fit to a linear function and the slopes that define the initial reaction velocities are plotted against the substrate concentration to be fit to the hyperbolic MM equation. This process leads to extraction of the k_{cat} and K_S (also called the K_m) parameter for the enzyme-catalyzed reaction. Due to restrictions that are directly related to the experimental methodologies (e.g. linear range for product detection) or to the nature of the enzyme under investigation (e.g. an enzyme that binds its substrate very tightly but has a very slow maximal rate of substrate turnover), these experiments often do not accurately define the intrinsic rate and equilibrium constants. Also, when these experiments do work as intended, the enzymologist would only yield two of the multiple parameters that govern the overall enzymatic cycle.

A real challenge in obtaining reliable steady state values centers on designing experiments to minimize error and maximize information about the enzyme mechanism. Since

results from steady-state experiments are typically the first insight into the mechanistic character of an unknown enzyme, having a more thorough analysis gives the enzymologist a better understanding for designing more complicated kinetic and/or thermodynamic experiments. Obtaining microscopic rate constants are typically validated by numerous orthogonal chemical and physical studies, such as isotope and viscosity effects [3-9]. More elaborate single turnover and transient state kinetic experiments can be employed to define intrinsic rate constants in the minimal enzyme mechanism [10-15]. Together, data from multiple kinetic approaches can be globally fit based on numerical integration of the rate equations using a host of kinetic simulation software packages including KINSIM/FITSIM [16, 17], DYNAFIT [18], COPASI [6, 19], KinTek Global Explorer [20, 21], and Berkeley Madonna™. Current computation speed has made these algorithms very effective, especially in one and two dimensions, for recovering best-fits for rate constants from a given sampling of data without the need to explicitly solve the underlying model. However, as the number of simultaneously unknown rate constants increases, corresponding to the dimension of numerical integration, e.g. via the Levenberg–Marquardt algorithm, the relative speed gain from using explicit model solutions increases. Also by their very nature, such integration fitting algorithms are unable to easily exploit the specific features of analytic solutions to achieve the goal of minimizing the data collection needed to obtain reasonable precision of the rate constants. We propose to improve optimized collection through selection of data points based on direct, yet fully automated, analysis of features in analytical solutions of the specific model. This allows us to achieve surprising accuracy in fitting even three or more rate constants simultaneously from a sparing number of experimental data points.

Since the seminal work of MM [22], reaction kinetics have been monitored beyond the initial linear phase into a phase where substrate deletion and product accumulation results in slowing the instantaneous velocity of the reaction, approaching a velocity of zero at infinite time (or at equilibrium). The analysis of these long time domains through fitting of progress curves has provided an alternative method for ascertaining the parameters that govern the basic MM steady-state mechanism [23-27]. Nevertheless, the sensitivity of the fitting regimes has not been well defined [28], especially along the axes for experimental conditions that can be varied for any given enzymatic system. We attempt to address these issues with a fitting scheme outlined in this manuscript. We analyze for accuracy in recovering known rate constants from data in the presence of various error regimes. Parameter redundancy is elucidated explicitly, allowing us to both avoid redundant sampling on the one hand, and to test its effects on poorly chosen data as well. Several statistical comparisons of the efficiency of different experimental designs are performed. These especially focus on contrasting the effectiveness of finding unknown parameters starting from different known parameters and how sample size and the dimension of the data collected affects accuracy.

Materials and Methods

Experimental conditions and proteins used.

The DNA construct coding for the motor domain of human kinesin-5 (Eg5; residues 1–368) was cloned into a modified pET16b (Novagen) vector for bacterial expression to achieve an N-terminally 6x His tag and TEV cleavage sequence. (N-terminus: MGSSHHHHHH-SSGENLYFQGSH-^{1}MASQ). Expression and purification of nucleotide-free Eg5 was performed

as previously described [29-32]. Kinetic experiments for Eg5 were performed in assay buffer 1 (20 mM HEPES pH 7.2 with KOH, 5 mM magnesium acetate, 0.1 mM EGTA, 0.1 mM EDTA, 50 mM potassium acetate, 150 mM sucrose, 1 mM DTT) at 298 K as described previously [30, 31, 33]. Microtubules (MTs) were prepared using an aliquot of purified bovine brain tubulin that was thawed and cycled, and the MTs were stabilized with 20 μM paclitaxel (Sigma-Aldrich). Kinetic experiments for *Solanum tuberosum* apyrase (Apy; Grade VII; Sigma-Aldrich) were performed in assay buffer 2 (10 mM HEPES pH 7.2 with KOH, 15 mM $MgCl_2$, 1 mM EGTA, 100 mM NaCl, 150 mM sucrose, 0.02% v/v TWEEN 20) at 298 K with the concentrations reported as final after mixing. Protein concentrations were determined using the Bradford assay (Sigma-Aldrich) with bovine serum albumin as standard.

Kinetic assays.

For microtubule-stimulated ATPase measurements, Eg5 at 0.2 μM was complexed with paclitaxel-stabilized microtubules (MTs) at 2 μM and reacted with varying concentrations of ATP substrate (95–310 μM) and added ADP product (0 or 1500 μM). Samples of the reaction (10 μL each) were quenched in 1 N HCl (2.5 μL of 5 N stock) followed by addition of malachite green reagent (150 μL; room temperature; [34]). After thorough mixing by vortexing, 150 μL of this mixture was transferred to a 96-well plate for endpoint absorbance (λ=650 nm) using a BioTek Epoch microplate spectrophotometer. A standard curve from 10–400 μM Na_2HPO_4/NaH_2PO_4 (pH 7.2) was used to convert A_{650} to inorganic phosphate (P_i) concentration. For each progress curve, P_i produced by Eg5 ATPase (total P_i minus background P_i) was plotted against time.

For ADPase measurements, Apy at 0.13 μM was reacted with varying ADP substrate (40–260 μM) and added AMP product (0 or 10 mM). P_i product was monitored using the malachite green assay as described above. For each progress curve, P_i produced by Apy ADPase was plotted against time. Stock ATP and ADP concentrations were experimentally determined using absorbance (λ=260 nm) with molar extinction coefficient (15,400 L mol^{-1} cm^{-1}).

Mechanisms.

Expansions of the Michaelis-Menten (MM) mechanism establish irreversible product formation/release (MM model; Equation 1) as well as competitive product inhibition (PI) of enzyme activity with irreversible chemistry step (PI model; Equation 2) and the fully reversible (FR) three-step mechanism (FR model; Equation 3).

$$\mathrm{E} + \mathrm{S} \underset{k_2}{\overset{k_1}{\rightleftharpoons}} \mathrm{ES} \overset{k_3}{\rightarrow} \mathrm{E} + \mathrm{P} \tag{1}$$

$$\mathrm{E} + \mathrm{S} \underset{k_2}{\overset{k_1}{\rightleftharpoons}} \mathrm{ES} \overset{k_3}{\rightarrow} \mathrm{EP} \underset{k_6}{\overset{k_5}{\rightleftharpoons}} \mathrm{E} + \mathrm{P} \tag{2}$$

$$\mathrm{E} + \mathrm{S} \underset{k_2}{\overset{k_1}{\rightleftharpoons}} \mathrm{ES} \underset{k_4}{\overset{k_3}{\rightleftharpoons}} \mathrm{EP} \underset{k_6}{\overset{k_5}{\rightleftharpoons}} \mathrm{E} + \mathrm{P} \tag{3}$$

Closed form of the model.

Using *Mathematica* 10 (Wolfram), it was verified by differentiation in time that the differential equation solution takes the closed form of Scheme 1, where ω represents the Lambert ω-function and t is time [35-44].

Scheme 1

$$[\mathrm{P}]_t = \alpha - \beta\,\omega\left(\frac{\alpha}{\beta}e^{\frac{\alpha-\gamma t}{\beta}}\right)$$

	Fully Reversible (FR)	Product Inhibition (PI)	Michaelis-Menten (MM)
α	$\frac{[\mathrm{S_o}]k_{cat}^F K_P - [\mathrm{P_o}]k_{cat}^R K_S}{k_{cat}^F K_P + k_{cat}^R K_S}$	$[\mathrm{S_o}]$	$[\mathrm{S_o}]$
β	$\frac{K_P K_S\{[\mathrm{P_o}](k_{cat}^R + k_{cat}^F) + [\mathrm{S_o}](k_{cat}^R + k_{cat}^F) + k_{cat}^F K_P + k_{cat}^R K_S\}}{(K_P - K_S)(k_{cat}^F K_P + k_{cat}^R K_S)}$	$\frac{([\mathrm{P_o}] + [\mathrm{S_o}] + K_P)K_S}{K_P - K_S}$	K_S
γ	$[\mathrm{E_o}]\frac{k_{cat}^F K_P + k_{cat}^R K_S}{K_P - K_S}$	$[\mathrm{E_o}]\frac{k_{cat}^F K_P}{K_P - K_S}$	$[\mathrm{E_o}]k_{cat}^F$

Asymmetric confidence estimates.

Since our simulation analyses tested different magnitudes of total degrees of freedom, we utilized mean squared error (MSE) from nonlinear least squared fitting to construct 3D contour plots for asymmetric confidence estimates based on a similar method of FitSpace Explorer [21]. Briefly, pairs of parameters were fixed and varied around their best fit value, and remaining constants were floated in the fitting scheme. The best fit MSE was divided by the MSE from each fixed pair of constants $\left(\mathrm{MSE}_{min}/\mathrm{MSE}_{x,y}\right)$ within a 14x14 matrix centered on the best fit values. Contour plots were constructed in *Mathematica* using these arrays of MSE ratios to linearly interpolate values to provide smooth contours at 0.1 intervals. Bounds of the asymmetric confidence estimates were defined at 11% and 25% increases in MSE $\left(\mathrm{MSE}_{min}/\mathrm{MSE}_{x,y} = 0.9 \text{ (red) and } 0.8 \text{ (orange), respectively}\right)$ as previously reported [21].

Error simulation.

To generate variably 'noisy' data sets that realistically represent a range of experimental observations, normally distributed random additive error (errA) was added to the quantity of product formed during the simulations using Scheme 1 and the following equation:

$$[\mathrm{P}]_t^{\sigma} = [\mathrm{P}]_t + \mathrm{errA}[\mathrm{S_o}]\kappa \tag{4}$$

where κ represents the fraction of $[\mathrm{S_o}]$ converted to product at equilibrium defined as:

$$\kappa = \frac{k_{cat}^F K_P}{\left(k_{cat}^F K_P + k_{cat}^R K_S\right)} \tag{5}$$

Time parameter threshold computation.

We used the closed form of the model to determine the near-optimal time parameter interval to maximize the efficiency of the fitting algorithms given a fixed sampling frequency. The most valuable portion of the curve is where the absolute value of the second derivative is bounded away from zero, since the remaining portions are nearly linear where the data points are quite redundant for the nonlinear parameter estimation. Toward this end, we compute the derivative of the product formula (Scheme 1). The derivative $\omega'(t)$ of the Lambert ω-function is:

$$\omega'(t) = \frac{1}{t\left(\frac{1}{\omega(t)} + 1\right)} \tag{6}$$

Hence $[\mathrm{P}]'(t)$ simplifies to:

$$[\mathrm{P}]'_t = \gamma - \frac{\gamma}{1 + \omega\left(\frac{\alpha}{\beta} e^{\frac{\alpha - \gamma t}{\beta}}\right)} \tag{7}$$

This quantity exponentially decays to 0 as $t \to \infty$. Given this derivative, we chose an appropriate scale to determine when it has become sufficiently small. Since the product $[\mathrm{P}]_t$ is almost exactly linear at the beginning of the time regime, we use the scale of the derivative at 0 to set the appropriate scale.

$$[\mathrm{P}]'_{t=0} = \frac{[\mathrm{E_o}]\left([\mathrm{S_o}] k_{cat}^{F} K_P - [\mathrm{P_o}]\, k_{cat}^{R} K_S\right)}{[\mathrm{S_o}]\, K_P + \left([\mathrm{P_o}] + K_P\right) K_S} \tag{8}$$

We set the stopping time t_{stop} to be the time when the derivative reaches $[\mathrm{P}]'[t_{stop}] = \theta * [\mathrm{P}]'_{t=0}$ for a given threshold parameter $\theta = v_f / v_o$ where v_f is the instantaneous final velocity and v_o is the initial velocity as $t \to 0$.

Design for data simulation.

Simulated progress curves were generated through calculating the product concentration (Scheme 1) with random error applied (Eqs. 4 and 5) over time under various conditions. SIM1 represents a data set at increasing $[\mathrm{S_o}]$ with no added $[\mathrm{P_o}]$ at $t = 0$. SIM2 represents a data set at constant $[\mathrm{S_o}]$ with increasing $[\mathrm{P_o}]$ that was added at $t = 0$. SIM3 represents a data set at increasing $[\mathrm{S_o}]$ and increasing $[\mathrm{P_o}]$ that was added at $t = 0$. All simulated data sets were designed for equal total degrees of freedom (i.e. equal total data points with 25 points per curve and 9 curves; Table A.1). All statistical evaluations were standard for NonlinearModelFit in *Mathematica.*

Scoring functions.

Judging the quality of the final nonlinear least squares fitting for our SIM1, SIM2, and SIM3 simulations required a scoring function that incorporated both accuracy (i.e. estimated

constant from curve fitting divided by the expected constant) and precision (i.e. symmetric 95% confidence interval). In general, more meaningful fitted constants correlate with higher accuracy (~1) and lower confidence intervals. A scoring function was developed to assess each fitted constant:

$$\text{Score} = e^{-\rho\left|\frac{C_{est}}{C_{exp}} - 1\right|^{\sigma} - \tau\left|\text{CI}_{95\%}\right|^{2\sigma}} \tag{9}$$

where C_{est} was the estimated constant from nonlinear least squares fitting, C_{exp} was the simulated (i.e. expected) constant, and $\text{CI}_{95\%}$ was the confidence interval (95%) for each measurement based on a simple multiple of the standard error. ρ, σ, and τ are weighting constants chosen to provide a score of 0.5–1 when normalized constants (C_{est}/C_{exp}) are 0.5–1.5 with narrow confidence intervals (< ±50%). For all score calculations presented in this manuscript, $\rho = 1$, $\sigma = 0.75$, and $\tau = 0.5$ (Fig S2*b*).

Results

Analysis of experimental kinetic data for Apy and Eg5

Experimental data for kinesin-5 (Eg5) and apyrase (Apy) were fit to either one of three mechanistic models, represented by the algebraic rate equations listed in Scheme 1. The three models, in order of increasing complexity, are designated as "MM", "PI", and "FR", respectively. Model MM represents the simplest Michaelis-Menten kinetic model, whereby the reaction product does not rebind and inhibit the enzyme and the reaction proceeds only in the forward direction (Equation 1). Model PI represents the product inhibition model, whereby the reaction product rebinds and inhibits the enzyme but the overall reaction can only proceed in the forward direction (Equation 2). Model FR represents the fully reversible kinetic model, whereby the reaction product not only rebinds and inhibits the enzyme, but also the overall reaction can proceed in either the forward or reverse direction (Equation 3).

Apy (type VII; EC 3.6.1.5) catalyzes the hydrolysis ADP to AMP and P_i [45]. To measure the ADPase kinetics of Apy, we performed a series of reactions at different ADP substrate (40–260 μM) and different AMP product (0 and 10 mM) concentrations and used the malachite green colorimetric assay to quantify P_i produced over time. Figs 1*a*–*b* summarize the results from fitting the Apy ADPase data using basic initial velocity analysis as well as to the three independent models (Table 1). Using both AIC and BIC to assess the goodness of fit relative to the complexity of the model, the apyrase data were best modeled to the FR mechanism (Equation 3) given the lower AIC and BIC values compared to the PI model (Δ AIC $=-17$ and ΔBIC $=-14$). However, asymmetric confidence analysis (Table 2; Fig S1*a*) indicated that not all the parameters were well constrained. Specifically, we observed large upper bounds for K_P and k_{cat}^{R} and no lower bound for k_{cat}^{R}. These results suggest that AMP product weakly binds the enzyme yet will proceed in the reverse direction to reform ADP substrate (i.e. the equilibrium constant for the chemistry step was 7 at the best fit values).

Fig 1. Global fitting of Apy and Eg5 progress curves. (*a*) P_i product formation as a function of time for Apy (closed symbols) shown at different initial substrate (ADP) concentrations (as indicated) with no added AMP product at the start of the reactions. (*b*) Time courses of Apy product formation (open symbols) at different substrate concentrations (as indicated) plus added initial AMP product (10 mM) at the start of the reactions. (*c*) P_i product formation as a function of time for Eg5 (closed symbols) shown at different initial substrate (ATP) concentrations (as indicated) with no added ADP product at the start of the reactions. (*d*) Time courses of Eg5 product formation (open symbols) at different substrate concentrations (as indicated) plus added initial ADP product (1.5 mM) at the start of the reactions. The smooth lines in each panel represent the best fit global FR model for Apy and PI model for Eg5 using NonlinearModelFit in Mathematica with fit constants provided in Table 1.

Eg5 (EC 3.6.4.4) catalyzes the hydrolysis of ATP to ADP and P_i products, which can be monitored using the malachite green colorimetric assay [34]. We designed a matrix of microtubule-stimulated Eg5 ATPase reactions with varying ATP substrate (95–310 μM) and added ADP product (0 and 1500 μM) concentrations for our analysis. Figs 1*c–d* summarize the results from fitting the microtubule-stimulated Eg5 ATPase data using the three independent models from above with a direct comparison with previously published results (Table 1) [30, 33]. We found good agreement with the steady-state constants, especially the comparison of the k_{cat}^{F} values ($5.4 \pm 0.1\ \mathrm{s}^{-1}$ in previous work [33]; $5.7 \pm 0.4\ \mathrm{s}^{-1}$ here). The estimates for K_S were consistently 3–8-fold higher than the previously determined MM value, yet the estimate from the PI model was very similar to the $K_{d,ATP}$ determined from presteady-state acid quench experiments at 20 ± 7 μM [33]. However, increasing the complexity of the model from PI to FR led to a decrease in the goodness of fit with higher AIC and BIC values ($\Delta\mathrm{AIC} = 2$ and $\Delta\mathrm{BIC} = 5$). These results are consistent with the known chemistry of Eg5, which shows no evidence of ATP synthesis in the presence of microtubules [33]. Asymmetric confidence analysis (Table 2; Fig S1*b*) showed no lower bounds for K_S and K_P. Thus, our analysis provides progress curve fitting to extract the binding constants for substrate and product as well as the forward catalytic rate constant with minimal data sets (i.e. 6–8 time courses).

Simulated enzyme mechanisms

To demonstrate the ability of our multidimensional chemical kinetics analysis of SIM1, SIM2, and SIM3 simulations to extract the four Michaelis-Menten (MM) constants from 'noisy' data, we established 2 enzymes based on the best fit constants of Apy and Eg5 as well as 20 hypothetical enzymes that show a wide range of kinetic mechanisms (E1-E20, Table 3, Fig S2*a*). These enzyme mechanisms display 1) widely varying equilibrium constants that span biologically relevant enzymes ($K_{eq}^{*} = k_{cat}^{F}K_P/k_{cat}^{R}K_S = 1.0 - 2.4\mathrm{x}10^4$), 2) a wide variety of substrate and product binding constants (K_S and K_P, respectively) including various combinations seen in natural enzymes, 3) a combinatorial spectrum of forward and reverse catalytic constants (k_{cat}^{F} and k_{cat}^{R}, respectively) as well as 4) a range of catalytic efficiencies (k_{cat}^{F}/K_S) from extremely inefficient enzymes (Apy $= 0.13\ \mu\mathrm{M}^{-1}\mathrm{s}^{-1}$) to a 'catalytically perfect' enzyme (E1 $= 1003\ \mu\mathrm{M}^{-1}\mathrm{s}^{-1}$). To generate 'noisy' data sets that realistically represent a range of experimental observations, normally distributed random additive error was added to the quantity of product formed during the simulations using Equations 4 and 5 (Fig S3). Given that the four

constants (K_S,K_P,k_{cat}^F,k_{cat}^R) were defined in these simulations, we could assess our nonlinear least squares fitting for both accuracy and precision using a scoring algorithm (Equation 9), which is a function of both the normalized constant (defined as the fitted constant divided by the actual constant) and symmetric 95% confidence interval (Fig S4).

Data set design

For the SIM1 simulations, the concentration of substrate was varied over a range of concentrations from 0.5*K_S to 5*K_S with no added product ($[P_o]$) at the start of the simulations. For the SIM2 simulations, the substrate concentration was kept constant at 5*K_S and the concentration of initial added product ($[P_o]$) at the start of each simulation was varied from 0 to 5*K_P (Table 2). The SIM3 simulations provided a matrix (3x3) using the variation of initial substrate ($[S_o]$) and initial added product ($[P_o]$) concentrations as described for SIM1 and SIM2 simulations, respectively (see Table S1 for an example of simulation conditions). The calculated time domain was optimized through the θ ratio of final instantaneous velocity (v_f) to initial velocity (v_o) to normalize the data point density across each time domain (see Materials and Methods for details). Each simulation provided a unique experimental data set with a pseudorandom Gaussian error distribution and these data sets were fit using NonlinearModelFit in *Mathematica.* For Apy and E1-E20, all four constants (K_S,K_P,k_{cat}^F,k_{cat}^R) were floated during the analysis. Since Eg5 demonstrated a minimal PI mechanism, only three constants (K_S,K_P,k_{cat}^F) were floated during the fitting. We assessed the sensitivity of the SIM1, SIM2, and SIM3 designs against varying additive error for Apy and Eg5 (Fig 2). For Apy, SIM1 was significantly inferior to the SIM2 and SIM3 designs based on overall fitting scores (Figs 2*a–d*). However, SIM3 was a significantly better design compared to the SIM2 design (Fig 2*d*), suggesting that varying both initial substrate and product concentrations provided better overall fitting scores. For Eg5, SIM1 and SIM3 were comparable for K_S and k_{cat}^F estimates (Figs 2*e–g*), but SIM3 provided a significantly improved mean fitting score for K_P (Fig 2*h*).

Fig 2. Error sensitivity for Apy and Eg5 progress curves under different simulation designs. (*a-c*) Fitting results for Apy simulations with the fitting score for each constant (K_S, orange; K_P, blue; k_{cat}^F, red; k_{cat}^R, cyan) plotted against error using SIM1, SIM2, and SIM3 conditions, respectively. (*d*) At constant 3% error, the mean fitting score for each constant was determined under different Apy simulation designs (as indicated). n=200. Error bars represent the standard deviation of the mean. (*e-g*) Fitting results for Eg5 simulations with the fitting score for each constant plotted against error using SIM1, SIM2, and SIM3 conditions, respectively. (*h*) At constant 3% error, the mean fitting score for each constant was determined under different Eg5 simulation designs (as indicated). n=200. N/A, not applicable.

We tested the sensitivity of all three simulation designs (SIM1, SIM2, SIM3) against constant additive error for our hypothetical enzymes, E1-E20 (Fig S5). For the hypothetical enzymes with $K_S \geq K_P$ (E1-E10), SIM2 provided very low (or indistinguishable from zero) fitting scores for K_S and K_P. SIM3 provided comparable fitting scores for K_S, K_P and k_{cat}^F, but consistently higher scores for k_{cat}^R. For the hypothetical enzymes with $K_S < K_P$ (E11-E20), SIM1 provided very low (or indistinguishable from zero) fitting scores for K_P and k_{cat}^R. SIM3 provided

consistently higher scores for all four constants compared to SIM2. There was no obvious correlation to better fitting scores when comparing 1) k_{cat}^{F} relative to k_{cat}^{R}, 2) overall equilibrium constant $\left(K_{eq}^{*}\right)$, or 3) enzymatic efficiency $\left(k_{cat}^{F}/K_{S}\right)$. Overall, varying both initial substrate and added product was successful to attain higher fitting scores regardless of the enzyme mechanism.

We assessed the sensitivity of the SIM3 design against varying additive error for our hypothetical enzymes, E1-E20 (Fig S6). All enzymes showed fitting scores that dropped below 0.5 at >3% error except for E10, which had fitting scores that dropped below 0.5 around 2% error. Using asymmetric confidence estimate analysis, simulated Apy SIM3 designs were tested at various additive errors (Fig S7). Since the asymmetric confidence interval was consistently larger than the symmetric 95% confidence estimates obtained from nonlinear least squares fitting, it was difficult to directly compare quality scores (Fig S8). Given the wide range of different hypothetical enzymes tested in this analysis, the SIM3 design (where both initial substrate and added product is varied) is likely to be successful at estimating the four MM parameters from fairly 'noisy' data for any enzyme found in natural systems.

Correlations between the estimated constants during the fitting

We next examined the six possible correlations between the four fitted parameters: K_S,K_P ,k_{cat}^{F},k_{cat}^{R} (Fig 3). Using the standard asymptotically unbiased sample correlation matrix, we computed the estimated correlations as we varied through E1-E20, Apy, and Eg5 with constant error and fixed starting values. We observed that the correlations remain roughly the same, but there was some variability as the parameters for the different enzymes vary. The relationship in the model accounts for most of the nonzero correlation especially considering a given fitting curve determined only three of the four parameters uniquely as discussed (Equations S18-S29). The simplicity of the symmetry between K_S and K_P accounts for the near perfect correlation between these two constants.

Fig 3. Parameter correlations for SIM3. All six parameter correlation coefficients (ρ) are shown for E1-E20, Apy, and Eg5 for the SIM3 design. Error was held constant at 3%. E1-E20 and Apy were modeled using the FR mechanism whereas Eg5 was modeled using the PI mechanism.

Time domain length effects on the parameter estimation

We investigated the sensitivity of the SIM3 design for the length of the time domain based on the ratio of instantaneous velocity from the start of the reaction till the end ($\theta = v_f/v_o$; Fig 4). We observed similar fitting scores for Apy and Eg5 over a broad range of time domains, indicating a vast insensitivity of the SIM3 design on the short time domain for data simulation (Figs 4*a-c* and S9). At θ ratios greater than 0.7, the fitting scores became markedly lower, indicating that the fitting relies on a significant quantity of product formed during the reaction (i.e. 60-85% of initial substrate converted to product during the time course). However, if the time domain was extended beyond the optimal range and thus few data were found in the initial velocity phase of the curve, the fitting scores steadily decreased for Apy and Eg5 (Figs 4*d-f*). For the hypothetical enzymes with $K_S \geq K_P$ (E1-E10), fitting scores for K_S and K_P decreased rapidly as the time domain was lengthened (Fig S10). For the hypothetical enzymes with $K_S < K_P$ (E11-E20), the overall fitting scores were less sensitive to the longer time domains (Fig S10).

Fig 4. SIM3 fitting scores as a function of time domain. Apy and Eg5 simulations were performed by varying the time domain [q as the ratio of final instantaneous velocity (v_f) to initial velocity (v_o)]. (*a*) Representative Apy simulations with constant error (3%) are shown with increasing q ratios (as indicated). (*b*-*c*) Fitting scores as a function of the q ratio for Apy and Eg5, respectively. (*d*) Representative plots of Apy simulations at decreasing q ratios. (*e*-*f*) Fitting scores as a function of decreasing q values with constant error at 3% for Apy and Eg5, respectively.

Data point density effects on parameter estimation

The total degrees of freedom correspond to the length of each progress curve (i.e. 25 data points per curve) multiplied by the number of progress curves. The number of data points per reaction was easy to modify since the total time domain was divided evenly from 2–1024 data points (i.e. changing the time constant for the calculation of $[P]_t$). Figs 5*a*-*b* highlights the SIM3 results for Apy and Eg5 by varying these conditions. As expected, the fitting scores increased proportionally with the number of data points per time course. These results were not specific to Apy and Eg5, but similar observations were made for all hypothetical enzymes tested (Fig S11). We observed a similar increase of the fitting scores by holding the number of data points per curve constant at 25 and varying the number of progress curves (data not shown). However, using methods to solve the asymmetric confidence estimates, we observed nearly constant fitting score as a function of data points per curve (Fig 5*c*). Therefore, the asymmetric confidence estimates solved from fitting two constants while fixing the values of the remaining two constants provided a robust estimate of the confidence interval independent of the number of total data points in the data set.

Fig 5. SIM3 results as a function of data point density. (*a*) Apy and (*b*) Eg5 simulations were performed by increasing the number of data points per curve (2-1024 points). Fitting scores were plotted as a function of number of data points per curve. (*c*) Asymmetric confidence estimates for K_S in Apy simulations at 0.8 (triangles) and 0.9 (circles) contours. Data plotted with log scale x-axis and were fit to a linear function. Error was held constant at 3%.

Discussion

Different enzyme mechanisms

We have tested the sensitivity of three different experimental designs centered on the closed form (Scheme 1) of the expanded Michaelis-Menten (MM) mechanism for enzyme activity. To simultaneously fit the raw data from individual progress curves to attain the four MM constants (K_S,K_P,k_{cat}^F,k_{cat}^R), we fit real experimental data for two enzymes (Apy and Eg5) as well as simulated data for 20 different enzymes (E1–E20). These hypothetical enzymes have a wide range of equilibrium constants, substrate and product binding affinities, catalytic constants, and enzymatic efficiencies (Fig S2*a*, Table 3). Despite all the divergent mechanisms tested in this study, the SIM3 design was robust in accurately and precisely estimating all four MM constants from an experimentally feasible data set.

Scoring function provides standard assessment

Given the complex nature of the data analysis from this study, we developed a simple exponential scoring function (Equation 9; Fig S2*b*) that yielded a score from 0 to 1 based on nonlinear least squares fitting accuracy (defined as the fit estimate divided by the actual constant) and fitting precision (defined as the symmetric 95% confidence interval for each estimate). In general, more meaningful fitted constants correlate with higher accuracy and lower confidence intervals. The weighting parameters that govern the scoring function were empirically determined to provide a fitting score of 0.5–1 when accuracy and precision were less than ±50%. The same weighting constants were used for the entire analysis, and depending on the nature of the questions being addressed, these weighting factors can be scaled accordingly to provide more stringent or more flexible scoring outcomes. The weighting factors employed in this study provide a reasonably conservative approach to evaluating the fitting results.

Best experimental design of steady-state experiments

Our analysis took advantage of three experimental designs called SIM1 (increasing $[\mathrm{S_o}]$ with no added $[\mathrm{P_o}]$ at $t = 0$), SIM2 (constant $[\mathrm{S_o}]$ with increasing added $[\mathrm{P_o}]$ at $t = 0$), and SIM3 (matrix of increasing $[\mathrm{S_o}]$ with increasing added $[\mathrm{P_o}]$ at $t = 0$). Typically, steady-state experiments are designed by only varying the substrate concentration as in SIM1. However, we found that adding initial product in the beginning of the reaction will lead to better parameter fitting results from both experimental and simulated kinetic data. Indeed, the best fitting regime happened with varying both substrate and initial product concentrations in the 3x3 matrix (Fig 2). We found that >50% of initial substrate conversion to product was necessary for accurately defining all four parameters (Fig 4). If experimental error remains constant, having greater than 24 data points per curve (assuming a 3x3 matrix) provided sufficiently accurate parameter estimates, but more data narrowed the symmetric confidence limits (Fig 5). We also observed that most hypothetical enzyme mechanisms tolerated a wide range of initial substrate and initial added product concentrations for achieving accurate parameter estimates (data not shown). Given the capabilities of *Mathematica*, future investigations aim to extend this analysis to mechanisms involving two substrates and products during the reaction (i.e. bi-bi mechanisms).

Variable conditions in the design of steady-state experiments

When studying a new uncharacterized enzyme, the decisions made by the enzymologist for initial substrate concentrations, total enzyme concentration, length of time domains, the number of data points per time domain, and the number of data sets to collect are difficult. Interestingly, global fitting under different simulation designs tolerated a wide spectrum of initial substrate and/or added product concentrations (~2-orders of magnitude around the K_S and K_P values; data not shown). The total enzyme concentration and the length of the time domain are directly related, though limits are often encountered based on the solubility limit of the enzyme, the dissociation of higher order oligomers of the enzyme, the minimum time constant for data sampling, and/or the stability of the enzyme under the experimental conditions. Therefore, these caveats make it difficult to achieve the optimal conditions to meet MM steady-state assumptions. Nevertheless, we found that our SIM3 design converged on all four parameters for all enzymes tested despite the length of time domain used.

As expected, we observed higher fitting scores when data sampling frequency increased over the course of the progress curve (Figs 5*a*-*b* and S11), yet, surprisingly, the scores did not linearly improve with the added number of data sets in the matrix (data not shown). These results

indicate that an enzymologist can attain accurate and precise information about the kinetic mechanism with ~10-fold less time and materials. We also found the matrix of initial substrate and added product concentrations need not be symmetrical for convergence on the four parameters (data not shown). Even a minimum of 2 initial substrate concentrations and 2 added product concentrations are sufficient to attain a rough sketch of the behavior of the enzyme, which will rapidly guide the researcher to subsequent experimental conditions for the next round of experiments. This approach will provide rapid feedback to appropriately hone on the experimental conditions that provide the lowest range of confidence intervals and (hopefully) more accurate estimates of the steady-state parameters.

Conclusions

Results presented here will impact how steady-state experiments will be designed and how steady-state kinetic data will be analyzed in the future. In the past, MM steady-state experiments focused on capturing data during the initial velocity time regime with variant substrate concentrations. These data were subsequently fit to a linear function and the slopes that define the initial reaction velocities were plotted against the substrate concentration to be fit to the hyperbolic MM equation to extract the k_{cat} and K_S parameters. Due to restrictions that were directly related to the experimental methodologies (e.g. linear range for product detection) or to the nature of the enzyme under investigation (e.g. an enzyme that binds its substrate very tightly but has a very slow maximal rate of substrate turnover), sometimes these experiments did not accurately define the intrinsic rate and equilibrium constants. Also, when these experiments did work as intended, the researcher would only yield two of the multiple parameters that govern the enzymatic cycle. Therefore, to minimize the time and materials used in the laboratory as well as maximize mechanistic insight through the multidimensional chemical kinetic analysis, the development of a modified experimental methodology is required. Combining the variation of both initial substrate and added product concentrations to establish a matrix of reaction conditions provides a simple strategy to generate experimental data that can be fit to extract all four MM parameters from a relatively simple data set. Since results from these experiments are typically the first insight into the mechanistic character of an unknown enzyme, having a more thorough analysis gives the enzymologist a better understanding for designing more complicated kinetic and/or thermodynamic experiments. Although these results promote the advancement of steady-state kinetic methodologies, data from presteady-state experiments are indispensible. These advancements in steady-state methodologies can provide insights of the mechanistic details that are derived through presteady-state methodologies and *vice versa*. Given the ability of *Mathematica* to fit data with pseudo-infinite dimensionality, these strategies can be applied to more complex mechanisms with multiple different substrates and products, activators, and small molecule inhibitors. Also, the progress curve fitting allows for enzyme concentration being a variable parameter, thus eliminating the need to normalize all reactions to a common enzyme concentration for comparison purposes. This study sets a foundation for modeling more complex mechanisms and will provide a new strategy for obtaining accurate and precise measurements of basic MM parameters.

Acknowledgments

We thank Jeffrey Ewer for preliminary data collection as well as Kayla Bell and Benjamin Walker for insightful comments on the manuscript. We acknowledge financial support from the Simons Foundation (grant #210442 to C.G.C.), the IU Faculty Research Support Program-External Resubmission (J.C.C.) and the National Science Foundation (MCB 1614514 to J.C.C.).

References

1. Michaelis L, Menten ML, Johnson KA, Goody RS. The original Michaelis constant: translation of the 1913 Michaelis-Menten paper. Biochemistry. 2011;50(39):8264-9. Epub 2011/09/06. doi: 10.1021/bi201284u. PubMed PMID: 21888353; PubMed Central PMCID: PMC3381512.

2. Briggs GE, Haldane JB. A Note on the Kinetics of Enzyme Action. Biochem J. 1925;19(2):338-9. Epub 1925/01/01. PubMed PMID: 16743508; PubMed Central PMCID: PMC1259181.

3. Loveridge EJ, Tey LH, Allemann RK. Solvent effects on catalysis by Escherichia coli dihydrofolate reductase. J Am Chem Soc. 2010;132(3):1137-43. doi: 10.1021/ja909353c. PubMed PMID: 20047317.

4. Spies MA, Woodward JJ, Watnik MR, Toney MD. Alanine racemase free energy profiles from global analyses of progress curves. J Am Chem Soc. 2004;126(24):7464-75. doi: 10.1021/ja049579h. PubMed PMID: 15198593.

5. Toney MD, Castro JN, Addington TA. Heavy-enzyme kinetic isotope effects on proton transfer in alanine racemase. J Am Chem Soc. 2013;135(7):2509-11. doi: 10.1021/ja3101243. PubMed PMID: 23373756; PubMed Central PMCID: PMCPMC3579662.

6. Toney MD. Common enzymological experiments allow free energy profile determination. Biochemistry. 2013;52(34):5952-65. doi: 10.1021/bi400696j. PubMed PMID: 23906433; PubMed Central PMCID: PMCPMC4286415.

7. Gadda G, Fitzpatrick PF. Solvent isotope and viscosity effects on the steady-state kinetics of the flavoprotein nitroalkane oxidase. FEBS Lett. 2013;587(17):2785-9. doi: 10.1016/j.febslet.2013.04.021. PubMed PMID: 23660407; PubMed Central PMCID: PMCPMC3916838.

8. Waldauer SA, Stucki-Buchli B, Frey L, Hamm P. Effect of viscogens on the kinetic response of a photoperturbed allosteric protein. J Chem Phys. 2014;141(22):22D514. doi: 10.1063/1.4897975. PubMed PMID: 25494785.

9. Herdendorf TJ, Nelson SW. Catalytic mechanism of bacteriophage T4 Rad50 ATP hydrolysis. Biochemistry. 2014;53(35):5647-60. doi: 10.1021/bi500558d. PubMed PMID: 25137526.

10. Ankrett RJ, Walmsley AR, Bagshaw CR. Kinetic analysis of regulated myosin ATPase activity using single and limited turnover assays. J Cell Sci Suppl. 1991;14:1-5. PubMed PMID: 1832160.

11. Johnson KA. Transient-state kinetic analysis of enzyme reaction pathways. The Enzymes. 20. New York: Academic Press; 1992. p. 1-61.

12. Gilbert SP, Mackey AT. Kinetics: a tool to study molecular motors. Methods. 2000;22(4):337-54. Epub 2001/01/03. doi: 10.1006/meth.2000.1086. PubMed PMID: 11133240.

13. Francklyn CS, First EA, Perona JJ, Hou YM. Methods for kinetic and thermodynamic analysis of aminoacyl-tRNA synthetases. Methods. 2008;44(2):100-18. doi: 10.1016/j.ymeth.2007.09.007. PubMed PMID: 18241792; PubMed Central PMCID: PMCPMC2288706.

14. Johnson KA. A century of enzyme kinetic analysis, 1913 to 2013. FEBS Lett. 2013;587(17):2753-66. doi: 10.1016/j.febslet.2013.07.012. PubMed PMID: 23850893.

15. Maxwell BA, Suo Z. Recent insight into the kinetic mechanisms and conformational dynamics of Y-Family DNA polymerases. Biochemistry. 2014;53(17):2804-14. doi: 10.1021/bi5000405. PubMed PMID: 24716482; PubMed Central PMCID: PMCPMC4018064.

16. Barshop BA, Wrenn RF, Frieden C. Analysis of numerical methods for computer simulation of kinetic processes: development of KINSIM--a flexible, portable system. Anal Biochem. 1983;130(1):134-45. PubMed PMID: 6688159.

17. Zimmerle CT, Frieden C. Analysis of progress curves by simulations generated by numerical integration. Biochem J. 1989;258(2):381-7. PubMed PMID: 2705989.

18. Kuzmic P. Program DYNAFIT for the analysis of enzyme kinetic data: application to HIV proteinase. Anal Biochem. 1996;237(2):260-73. PubMed PMID: 8660575.

19. Hoops S, Sahle S, Gauges R, Lee C, Pahle J, Simus N, et al. COPASI--a COmplex PAthway SImulator. Bioinformatics. 2006;22(24):3067-74. doi: 10.1093/bioinformatics/btl485. PubMed PMID: 17032683.

20. Johnson KA. Fitting enzyme kinetic data with KinTek Global Kinetic Explorer. Methods Enzymol. 2009;467:601-26. doi: 10.1016/S0076-6879(09)67023-3. PubMed PMID: 19897109.

21. Johnson KA, Simpson ZB, Blom T. FitSpace explorer: an algorithm to evaluate multidimensional parameter space in fitting kinetic data. Anal Biochem. 2009;387(1):30-41. doi: 10.1016/j.ab.2008.12.025. PubMed PMID: 19168024.

22. Michaelis L, Menten ML. Die kinetik der invertinwirkung. Biochem Z. 1913;49:333-69.

23. Bates DJ, Frieden C. Treatment of enzyme kinetic data. 3. The use of the full time course of a reaction, as examined by computer simulation, in defining enzyme mechanisms. J Biol Chem. 1973;248(22):7878-84. Epub 1973/11/25. PubMed PMID: 4796073.

24. Duggleby RG. Quantitative analysis of the time courses of enzyme-catalyzed reactions. Methods. 2001;24(2):168-74. Epub 2001/06/01. doi: 10.1006/meth.2001.1177. PubMed PMID: 11384191.

25. Philo RD, Selwyn MJ. Use of progress curves to investigate product inhibition in enzyme-catalysed reactions. Application to the soluble mitochondrial adenosine triphosphatase. Biochem J. 1973;135(3):525-30. Epub 1973/11/01. PubMed PMID: 4272499; PubMed Central PMCID: PMC1165854.

26. Orsi BA, Tipton KF. Kinetic analysis of progress curves. Methods Enzymol. 1979;63:159-83. Epub 1979/01/01. PubMed PMID: 502859.

27. Cao W, De La Cruz EM. Quantitative full time course analysis of nonlinear enzyme cycling kinetics. Sci Rep. 2013;3:2658. doi: 10.1038/srep02658. PubMed PMID: 24029878; PubMed Central PMCID: PMC3772379.

28. Gutierrez OA, Danielson UH. Sensitivity analysis and error structure of progress curves. Anal Biochem. 2006;358(1):1-10. Epub 2006/09/19. doi: 10.1016/j.ab.2006.07.007. PubMed PMID: 16979133.

29. Maliga Z, Kapoor TM, Mitchison TJ. Evidence that monastrol is an allosteric inhibitor of the mitotic kinesin Eg5. Chem Biol. 2002;9(9):989-96. Epub 2002/09/27. PubMed PMID: 12323373.

30. Cochran JC, Sontag CA, Maliga Z, Kapoor TM, Correia JJ, Gilbert SP. Mechanistic analysis of the mitotic kinesin Eg5. J Biol Chem. 2004;279(37):38861-70. Epub 2004/07/13. doi: 10.1074/jbc.M404203200. PubMed PMID: 15247293; PubMed Central PMCID: PMC1356567.

31. Cochran JC, Gilbert SP. ATPase mechanism of Eg5 in the absence of microtubules: insight into microtubule activation and allosteric inhibition by monastrol. Biochemistry. 2005;44(50):16633-48. Epub 2005/12/14. doi: 10.1021/bi051724w. PubMed PMID: 16342954; PubMed Central PMCID: PMC2270472.

32. Zhao YC, Kull FJ, Cochran JC. Modulation of the kinesin ATPase cycle by neck linker docking and microtubule binding. J Biol Chem. 2010;285(33):25213-20. Epub 2010/06/19. doi: 10.1074/jbc.M110.123067. PubMed PMID: 20558732; PubMed Central PMCID: PMC2919084.

33. Cochran JC, Krzysiak TC, Gilbert SP. Pathway of ATP hydrolysis by monomeric kinesin Eg5. Biochemistry. 2006;45(40):12334-44. Epub 2006/10/04. doi: 10.1021/bi0608562. PubMed PMID: 17014086; PubMed Central PMCID: PMC2288585.

34. Leonard M, Song BD, Ramachandran R, Schmid SL. Robust colorimetric assays for dynamin's basal and stimulated GTPase activities. Methods Enzymol. 2005;404:490-503. PubMed PMID: 16413294.

35. Lambert JH. Observations variae in Mathesin Puram. Acta Helvitica, physico-mathematico-anatomico-botanico-medica. 1758;3:128-68.

36. Euler L. De serie Lambertina Plurimisque eius insignibus proprietatibus. Acta Acad Scient Petropol. 1783;2:29-51.

37. Wright EM. Solution of the equation ze^z=a. Bull Amer Math Soc. 1959;65:89-93.

38. Fritsch FN, Shafer RE, Crowley WP. Algorithm 443: Solution of the transcendental equation we^w=x. Comm ACM. 1973;16(2):123-4.

39. Corless RM, Gonnet GH, Hare DEG, Jeffrey DJ, Knuth DE. On the Lambert W Function. Adv Comput Math. 1996;5:329-59.

40. Schnell S, Mendoza C. Closed form solution for time-dependent enzyme kinetics. J Theor Biol. 1997;187(2):207-12. doi: Doi 10.1006/Jtbi.1997.0425. PubMed PMID: ISI:A1997XL78400007.

41. Goudar CT, Sonnad JR, Duggleby RG. Parameter estimation using a direct solution of the integrated Michaelis-Menten equation. Biochim Biophys Acta. 1999;1429(2):377-83. Epub 1999/02/16. PubMed PMID: 9989222.

42. Goudar CT, Harris SK, McInerney MJ, Suflita JM. Progress curve analysis for enzyme and microbial kinetic reactions using explicit solutions based on the Lambert W function. J Microbiol Methods. 2004;59(3):317-26. Epub 2004/10/19. doi: 10.1016/j.mimet.2004.06.013. PubMed PMID: 15488275.

43. Golicnik M. Explicit analytic approximations for time-dependent solutions of the generalized integrated Michaelis-Menten equation. Anal Biochem. 2011;411(2):303-5. doi: Doi 10.1016/J.Ab.2011.01.016. PubMed PMID: ISI:000288519300020.

44. Schnell S, Mendoza C. Time-dependent closed form solutions for fully competitive enzyme reactions. Bull Math Biol. 2000;62(2):321-36. Epub 2000/05/29. doi: 10.1006/bulm.1999.0156. PubMed PMID: 10824432.

45. Kettlun AM, Uribe L, Calvo V, Silva S, Rivera J, Mancilla M, et al. Properties of two apyrases from Solanum tuberosum. Phytochemistry. 1982;21(3):551-8. doi: http://dx.doi.org/10.1016/0031-9422(82)83139-7.

Table 1. Parameter estimates from experimental data for Apy and Eg5.

Apy	**Initial Velocity**	**MM Model**		**PI Model**		**FR Model**	
Constant		Estimate ± SE	CI (95%)	Estimate ± SE	CI (95%)	Estimate ± SE	CI (95%)
K_S (μM)	75 ± 12	153 ± 22	110–197	218 ± 24	170–266	176 ± 20	137–214
K_P (μM)	ND [k]	NA	NA	25300 ± 2400	20600–30000	38500 ± 7000	24600–52300
k_{cat}^{F} (s^{-1})	18.7 ± 0.9	18.3 ± 1.5	15.3–21.2	22.7 ± 1.6	19.4–25.9	19.8 ± 1.3	17.2–22.4
k_{cat}^{R} (s^{-1})	NA	NA	NA	NA	NA	2.9 ± 1.1	0.7–5.1
AIC [i]	NA	NA		1620		1603	
BIC [j]	NA	NA		1634		1620	
Eg5	**Initial Velocity [a]**	**MM Model [b]**		**PI Model [c]**		**FR Model [d]**	
Constant		Estimate ± SE[e]	CI (95%)[f]	Estimate ± SE	CI (95%)	Estimate ± SE	CI (95%)
K_S (μM)	7.0 ± 0.4 [g]	58.2 ± 4.6	49–67	23.2 ± 10.3	2.9–43.5	23.1 ± 10.9	2–45
K_P (μM)	117 ± 86 [g]	NA	NA	124 ± 49	28–220	124 ± 49	27–221
k_{cat}^{F} (s^{-1})	5.4 ± 0.1 [g]	6.0 ± 0.1	5.7–6.3	5.6 ± 0.4	4.9–6.3	5.6 ± 0.4	4.8–6.4
k_{cat}^{R} (s^{-1})	NA [h]	NA	NA	NA	NA	1×10^{-13}	0–0.2
AIC	NA	NA		1119		1121	
BIC	NA	NA		1131		1136	

[a] Constants provided by the typical MM analysis of initial velocities. [b] Irreversible MM mechanism defined as in Equation 1. [c] Product inhibition (PI) mechanism defined as in Equation 2. [d] Fully reversible (FR) mechanism as in Equation 3. [e] Nonlinear least-squares fitting estimates with standard error for each fit parameter. [f] CI, confidence interval. [g] Steady-state and equilibrium binding constants obtained under similar experimental conditions [30, 33]. [h] NA, not applicable. [i] AIC, Akaike information criterion. [j] BIC, Bayesian information criterion. [k] ND, not determined.

Table 2. Asymmetric confidence limits for Apy and Eg5 experimental data.

Apy Constant	**FR Model** Estimate ± SE	**Confidence** $\mathrm{MSE}_{min}/\mathrm{MSE}_{x,y} = 0.9$	**Confidence** $\mathrm{MSE}_{min}/\mathrm{MSE}_{x,y} = 0.8$
K_S (μM)	176 ± 20	105 – 324	83 – 483
K_P (μM)	38,500 ± 7,000	19,410 – 238,360	15,105 – 345,170
k_{cat}^{F} (s^{-1})	19.8 ± 1.3	16.3 – 32.3	14.6 – 43.8
k_{cat}^{R} (s^{-1})	2.9 ± 1.1	0 – 50	0 – 74
Eg5 Constant	**PI Model** Estimate ± SE	**Confidence** $\mathrm{MSE}_{min}/\mathrm{MSE}_{x,y} = 0.9$	**Confidence** $\mathrm{MSE}_{min}/\mathrm{MSE}_{x,y} = 0.8$
K_S (μM)	23.2 ± 10.3	0 – 80	0 – 134
K_P (μM)	124 ± 49	0 – 334	0 – 457
k_{cat}^{F} (s^{-1})	5.6 ± 0.4	4.6 – 7.5	4.3 – 9.3

Table 3. Simulated enzyme constants.

	Kin5	**Apy**	**E1**	**E2**	**E3**	**E4**
$K_S(\mu\mathrm{M})$	30	150	0.11	0.11	1.1	10.1
$K_P(\mu\mathrm{M})$	120	25000	0.12	0.12	1.2	10.2
$k_{cat}^F(\mathrm{s}^{-1})$	6	20	110.3	5.3	5.3	10.3
$k_{cat}^R(\mathrm{s}^{-1})$	0.001	3	2.4	0.4	0.4	0.4
K_{eq}^*	24000	1111	50.1	14.5	14.5	26
$k_{cat}^F/K_S(\mu\mathrm{M}^{-1}\mathrm{s}^{-1})$	0.2	0.13	1003	48.2	4.8	1.0
	E5	**E6**	**E7**	**E8**	**E9**	**E10**
$K_S(\mu\mathrm{M})$	0.11	1.1	10.1	100.1	1.1	10.1
$K_P(\mu\mathrm{M})$	0.12	1.2	10.2	100.2	0.62	1.2
$k_{cat}^F(\mathrm{s}^{-1})$	20.3	2.3	15.3	100.3	10.3	5.3
$k_{cat}^R(\mathrm{s}^{-1})$	20.4	2.4	15.4	100.4	4.4	0.4
K_{eq}^*	1.1	1.0	1.0	1.0	1.3	1.6
$k_{cat}^F/K_S(\mu\mathrm{M}^{-1}\mathrm{s}^{-1})$	185	2.1	1.5	1.0	9.4	0.5
	E11	**E12**	**E13**	**E14**	**E15**	**E16**
$K_S(\mu\mathrm{M})$	0.1	0.1	1.1	10.1	0.1	1.1
$K_P(\mu\mathrm{M})$	1.2	1.2	10.2	100.2	1.2	10.2
$k_{cat}^F(\mathrm{s}^{-1})$	10.3	5.3	5.3	10.3	2.3	2.3
$k_{cat}^R(\mathrm{s}^{-1})$	4.4	0.4	0.4	3.4	2.4	2.4
K_{eq}^*	28.1	159	123	30	11.5	8.9
$k_{cat}^F/K_S(\mu\mathrm{M}^{-1}\mathrm{s}^{-1})$	103	53	4.8	1.0	23	2.1
	E17	**E18**	**E19**	**E20**		
$K_S(\mu\mathrm{M})$	1.1	10.1	0.1	1.1		
$K_P(\mu\mathrm{M})$	100.2	100.2	1.2	100.2		
$k_{cat}^F(\mathrm{s}^{-1})$	2.3	100.3	1.3	1.3		
$k_{cat}^R(\mathrm{s}^{-1})$	2.4	100.4	5.4	10.4		
K_{eq}^*	87	9.9	2.9	11.4		
$k_{cat}^F/K_S(\mu\mathrm{M}^{-1}\mathrm{s}^{-1})$	2.1	9.9	13	1.2		

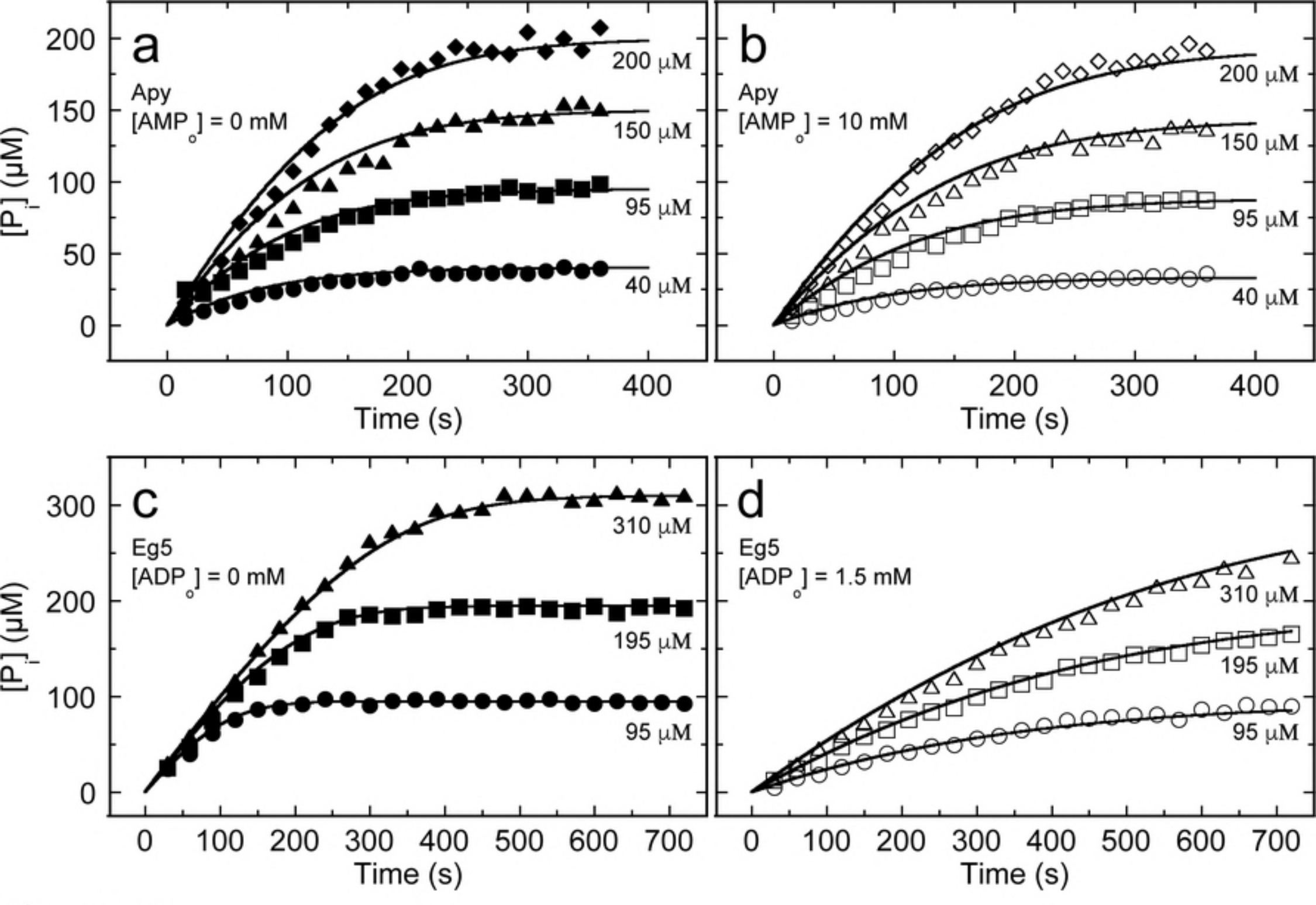
a
Apy
[AMP$_o$] = 0 mM
200 μM
150 μM
95 μM
40 μM
b
Apy
[AMP$_o$] = 10 mM
200 μM
150 μM
95 μM
40 μM
c
Eg5
[ADP$_o$] = 0 mM
310 μM
195 μM
95 μM
d
Eg5
[ADP$_o$] = 1.5 mM
310 μM
195 μM
95 μM
[P$_i$] (μM)
Time (s)

Figure 1

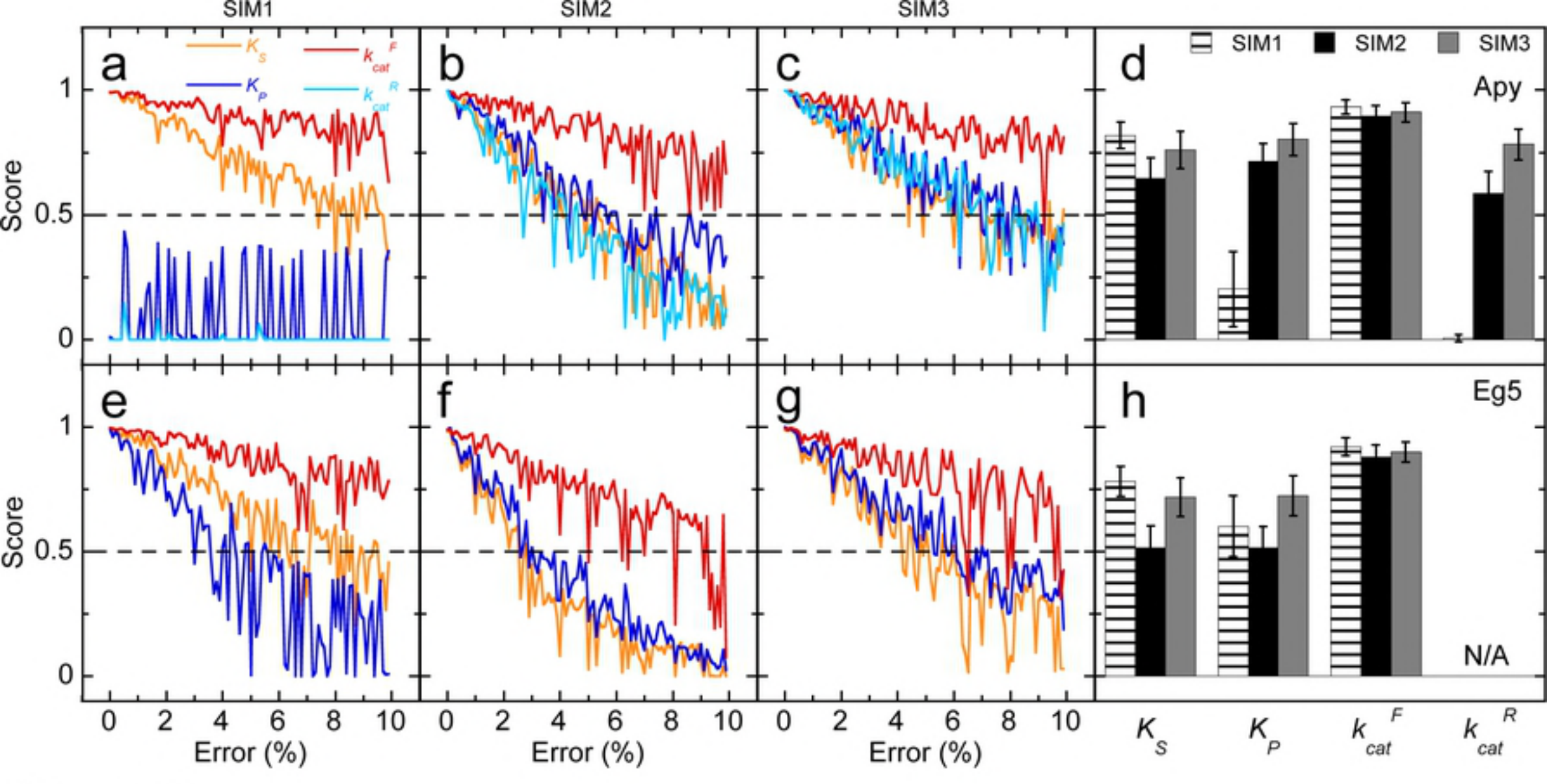

Figure 2

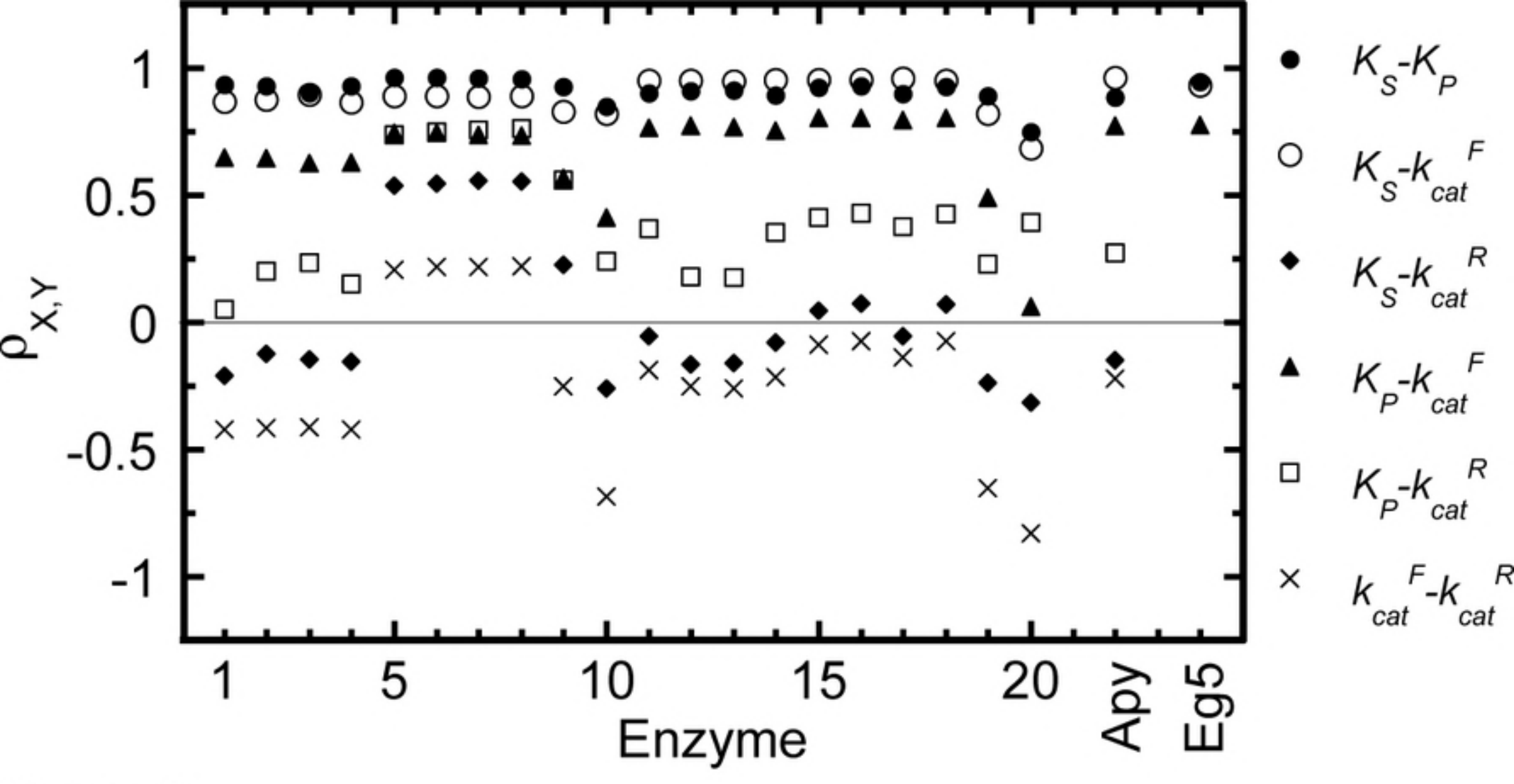
1
0.5
0
-0.5
-1
$\rho_{X,Y}$
1
5
10
15
20
Apy
Eg5
Enzyme
K_S-K_P
K_S-k_{cat}^F
K_S-k_{cat}^R
K_P-k_{cat}^F
K_P-k_{cat}^R
k_{cat}^F-k_{cat}^R

Figure 3

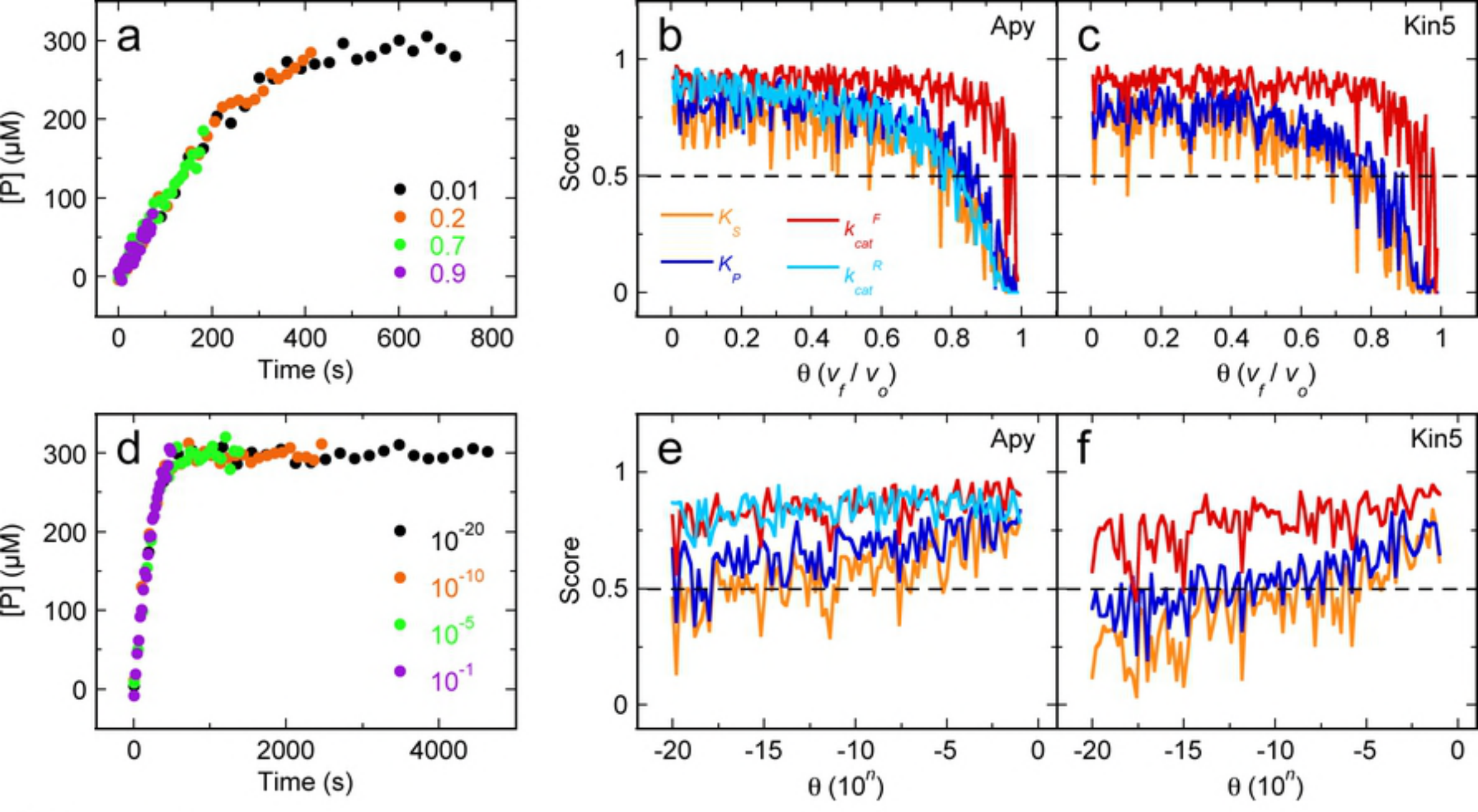
a
[P] (μM)
Time (s)
0.01
0.2
0.7
0.9
b
Apy
Score
θ (v_f / v_o)
K_S
K_P
k_cat^F
k_cat^R
c
Kin5
d
10^-20
10^-10
10^-5
10^-1
e
f
θ (10^n)

Figure 4

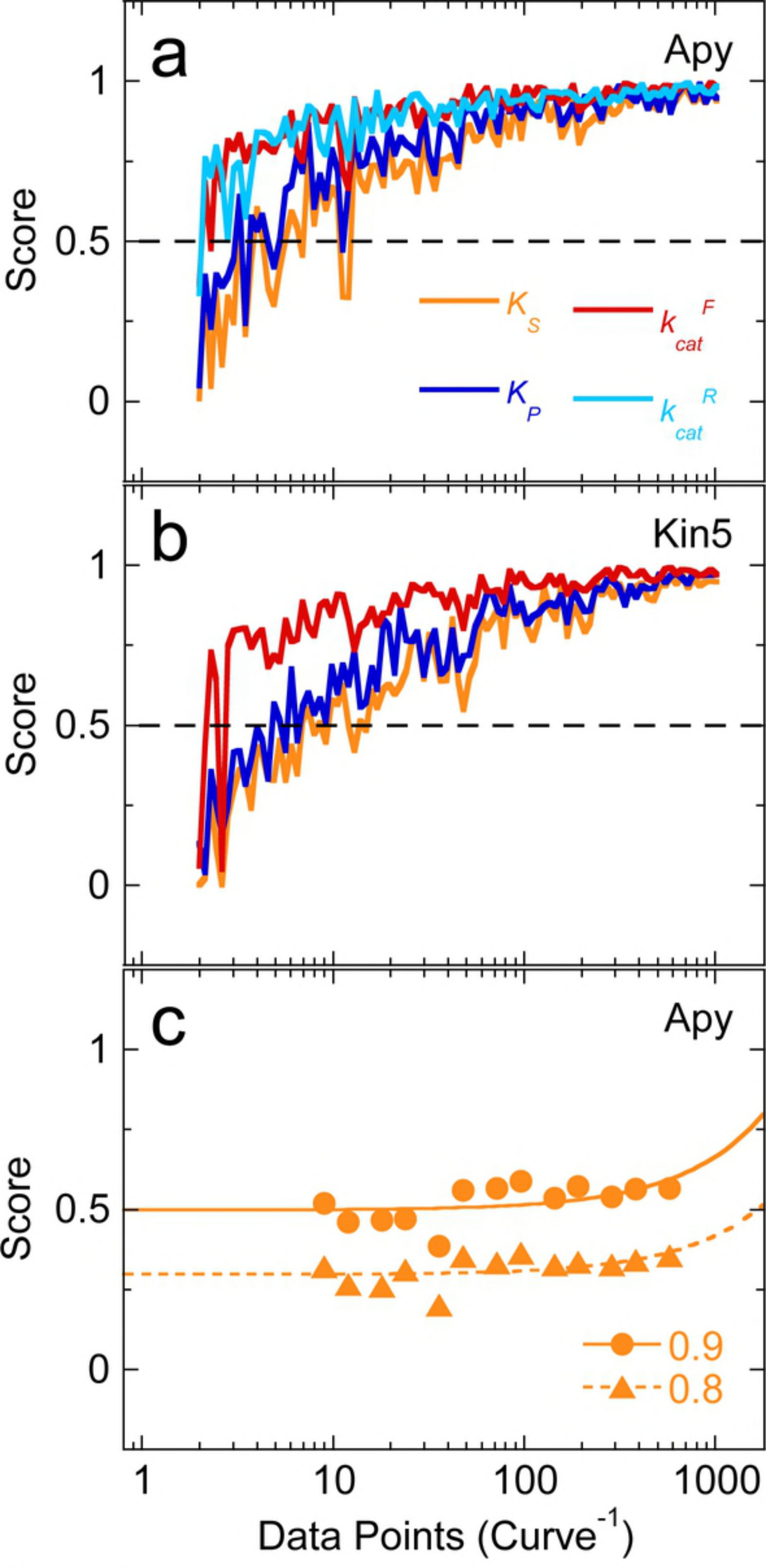
a
Apy
b
Kin5
c
Apy
Score
1
0.5
0
K_S
k_{cat}^F
K_P
k_{cat}^R
0.9
0.8
1
10
100
1000
Data Points (Curve^{-1})

Figure 5