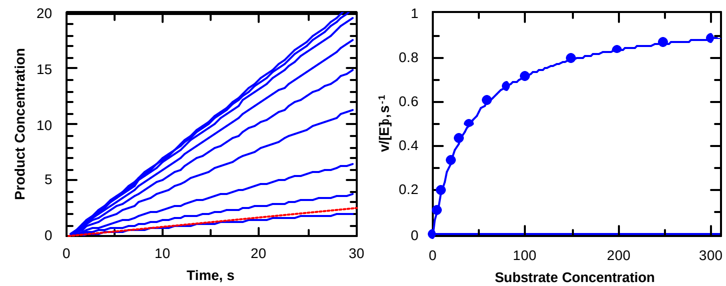
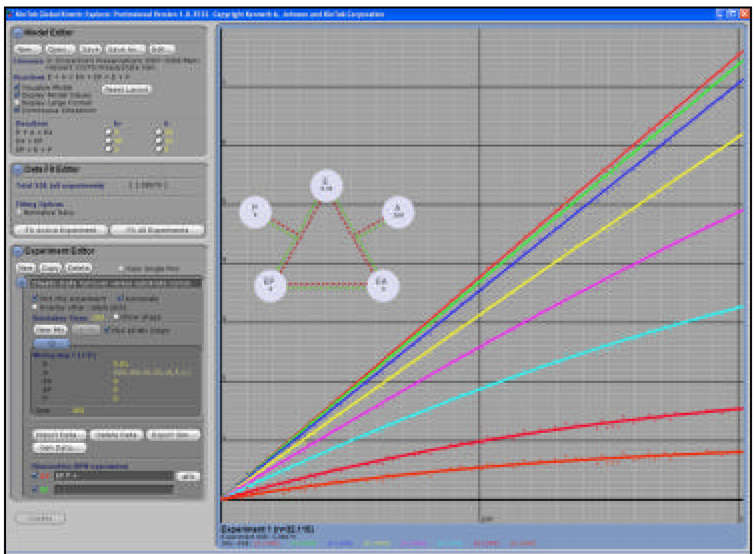
Computer Simulation and Global Data Fitting

Kenneth A Johnson
University of Texas at Austin

Conventional Steady-State Kinetics



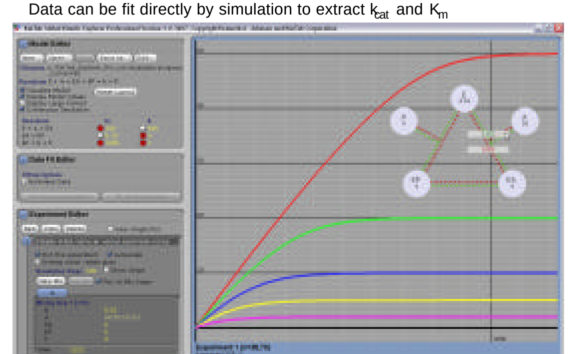
1. Measure initial rate
   a. Restrict data collection to first 10% of reaction
   b. If there is curvature, fit to polynomial to get initial rate
2. Plot rate versus concentration
3. Fit secondary plot to extract k_{cat} and K_m



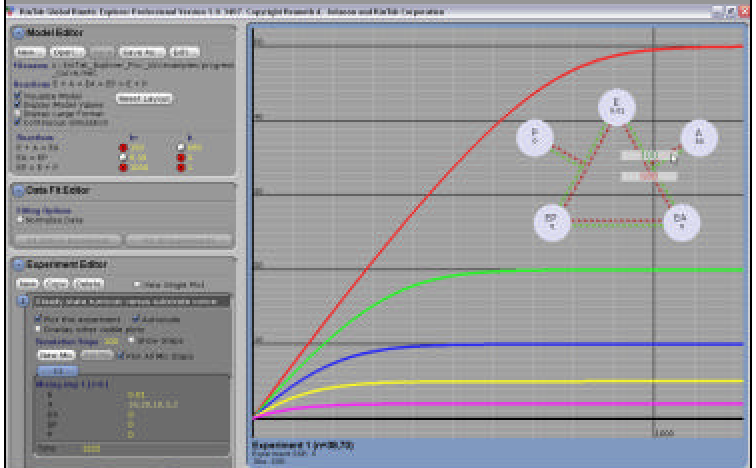
Full Time-course (Progress Curve) kinetics

$$E + S \underset{k_{-1}}{\overset{k_1}{\rightleftarrows}} ES \underset{k_{-2}}{\overset{k_2}{\rightleftarrows}} EP \underset{k_{-3}}{\overset{k_3}{\rightleftarrows}} E + P$$

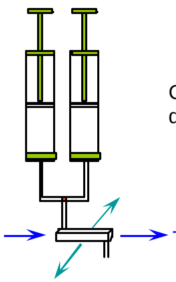
Follow reaction until reaction approaches equilibrium
Decreasing rate of turnover is due to decreasing [S] and increasing [P]
Data can be fit directly by simulation to extract k_{cat} and K_m



Full Progress Curve Kinetics



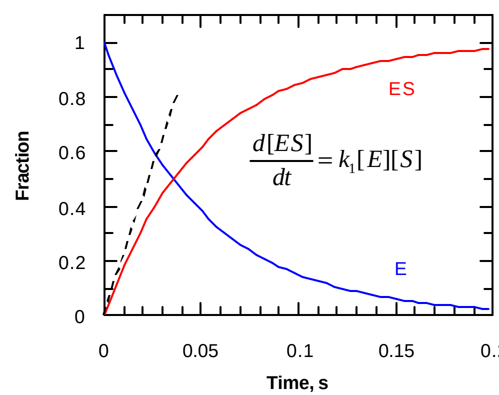
KinTek Stopped-Flow



- 30 µl per shot
- 1 msec dead time
- 10 µl observation cell

Kinetics of irreversible substrate binding

$$E + S \xrightarrow{k_1} E \cdot S$$



$$\frac{d[ES]}{dt} = k_1[E][S]$$

Kinetics of substrate binding: *irreversible binding*

$$E + S \xrightarrow{k_1} E \cdot S$$

$$d[E]/dt = -k_1[E][S]$$

$$d[E]/[E] = -k_1[S]dt$$

$$\int_{E_0}^{E} d[E]/[E] = -\int_0^t k_1[S]dt = -k_1[S]\int_0^t dt$$

$$\ln([E]/[E]_0) = -k_1[S](t - t_0)$$

$$[E]/[E]_0 = e^{-k_1[S]t}$$

$$[ES]/[E]_0 = 1 - e^{-k_1[S]t}$$

k_1 is a second order rate constant, units $M^{-1}s^{-1}$

Assume *[S]* is constant

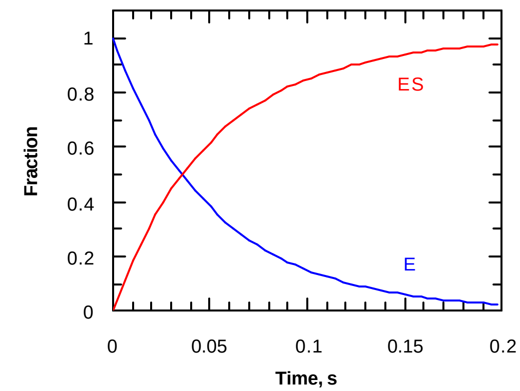
$k_1[S]$ is the pseudo-first order rate constant, unit s^{-1}

E decays and ES appears by an exponential function with rate *$k_1[S]$*

We often use the units of $\mu M^{-1}s^{-1} = 10^6 M^{-1}s^{-1}$.

Diffusion limit is approximately $10^8 M^{-1}s^{-1} = 100\ \mu M^{-1}s^{-1}$

Kinetics of irreversible substrate binding



$k_1 = 2\ \mu M^{-1}s^{-1}$

$[S] = 10\ \mu M$

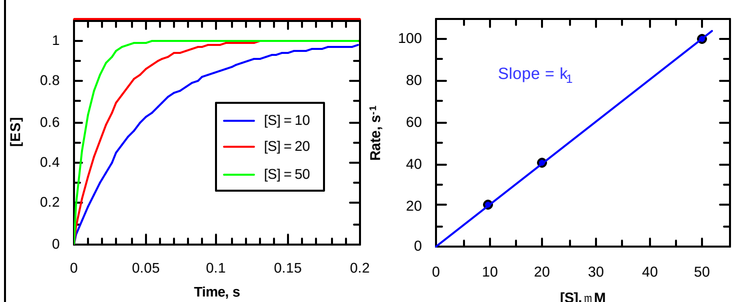
$k_1[S] = 20\ s^{-1}$

$t_{1/2} = (\ln 2)/k$

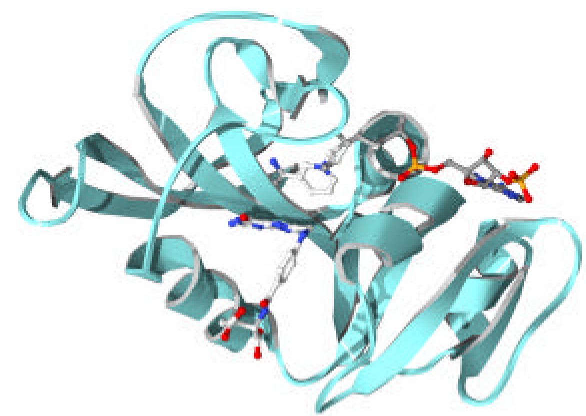
$= 34.6$ msec

Concentration dependence of binding rate

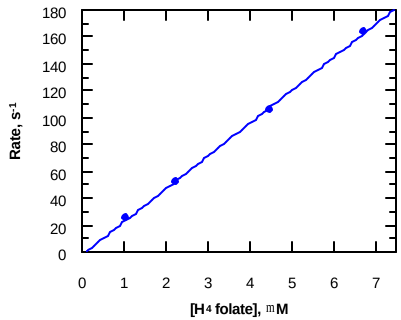
$$k_{obs} = k_1[S]$$



Dihydrofolate reductase (DHFR) with Folate and NADPH bound



Dihydrofolate reductase: Ligand binding kinetics



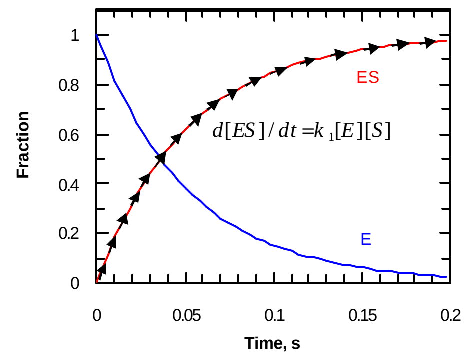
$$k_{obs} = k_1[S] + k_{-1}$$

Parameter	Value	Std. Error
a (intercept)	-2.0645	2.1998
b (gradient)	24.4393	0.5209

Numerical Integration

$$E + S \xrightarrow{k_1} E \cdot S$$

$$d[E]/dt = -k_1[E][S]$$
$$[E]_0 = [E] + [ES]$$
$$[S]_0 = [S] + [ES]$$



Analytical Solution:

$$[E]/[E]_0 = e^{-k_1[S]t}$$
$$[ES]/[E]_0 = 1 - e^{-k_1[S]t}$$

Kinetics of substrate binding: *Reversible binding*

$$E + S \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} E \cdot S$$

$$d[E]/dt = -k_1[E][S] + k_{-1}[ES]$$
$$d[ES]/dt = k_1[E][S] - k_{-1}[ES]$$
$$[E]_0 = [E] + [ES] \quad \text{mass balance}$$
$$d[E]/dt = -k_1[E][S] + k_{-1}([E]_0 - [E])$$
$$d[E]/dt = -(k_1[S] + k_{-1})[E] + k_{-1}[E]_0$$

$$[E]/[E]_0 = Ae^{-k_{obs}t}$$
$$[ES]/[E]_0 = A(1 - e^{-k_{obs}t})$$

$$k_{obs} = k_1[S] + k_{-1}$$

$k_{obs} = k_1[S] + k_{-1}$ is the sum of the forward and back rates.

$$A = K_1[S]/(K_1[S]+1)$$
$$= k_1[S]/(k_1[S]+k_{-1})$$

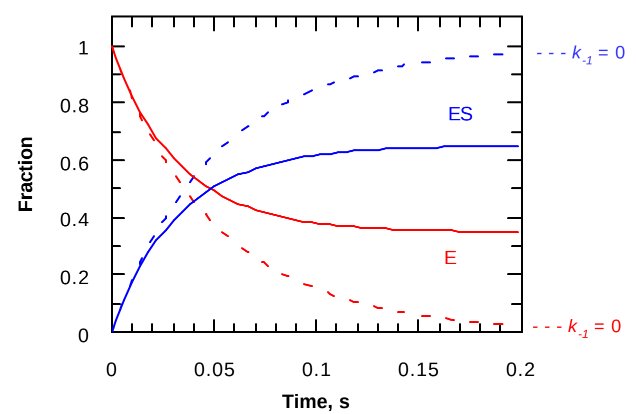
Amplitude, *A*, is a function of the equilibrium constant for forming the ES complex

General equation for data fitting: $Y = A \cdot e^{-k_{obs}t} + C$

Kinetics of reversible binding

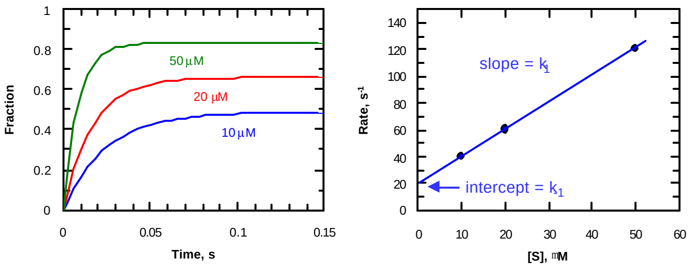
$$k_{obs} = k_1[S] + k_{-1}$$

Observed rate is the sum of forward and reverse rates.



Concentration dependence of binding rate

$$k_{obs} = k_1[S] + k_{-1}$$



NOTE: increase in amplitude <u>and</u> rate as a function of increasing [S]

One experiment can serve to define k_1, k_{-1} and K for S binding.

Kinetics of substrate binding: *Two-steps, four rates*

C. Complete solution $\quad E + S \underset{k_{-1}}{\overset{k_1}{\rightleftarrows}} E \cdot S \underset{k_{-2}}{\overset{k_2}{\rightleftarrows}} E \cdot X$

Each species follows a double exponential

$$[E]_i/[E]_0 = A_1 e^{-l_1 t} + A_2 e^{-l_2 t} + C$$

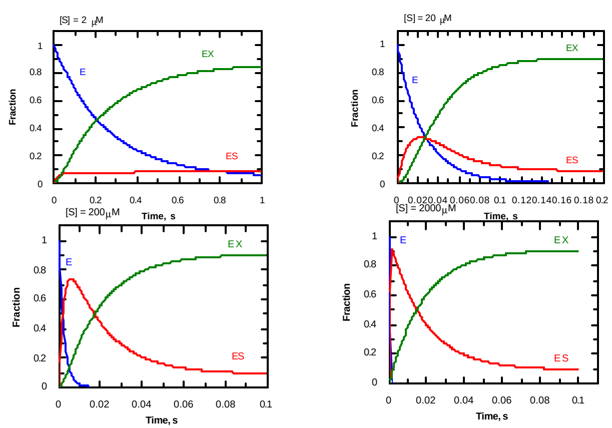
with rates of: $\quad l_1 \approx k_1[S] + k_{-1} + k_2 + k_{-2}$

$$l_2 \approx \frac{k_1[S](k_2 + k_{-2}) + k_{-1}k_{-2}}{k_1[S] + k_{-1} + k_2 + k_{-2}}$$

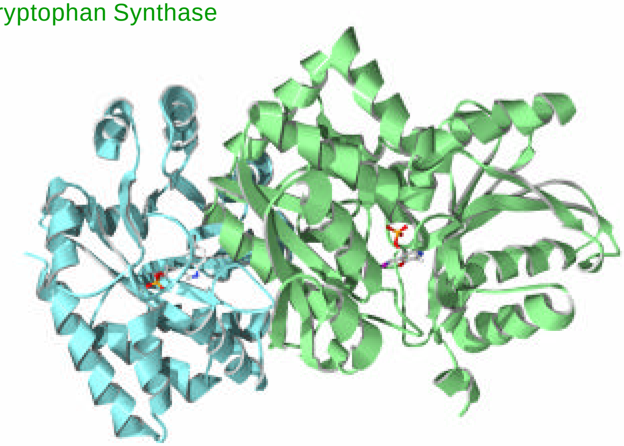
$$E + S \underset{k_{-1}}{\overset{k_1}{\rightleftarrows}} E \cdot S \underset{k_{-2}}{\overset{k_2}{\rightleftarrows}} E \cdot X$$

$k_1 = 2\ \mu M^{-1}s^{-1}$ $\quad k_1 = 2\ s^{-1}$

$k_2 = 50\ s^{-1}$ $\quad k_2 = 5\ s^{-1}$



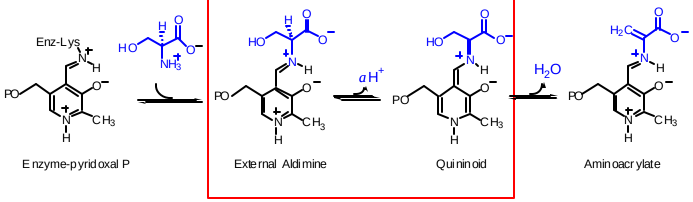
Tryptophan Synthase



α-IGP

β-Ser-PLP (aminoacrylate)

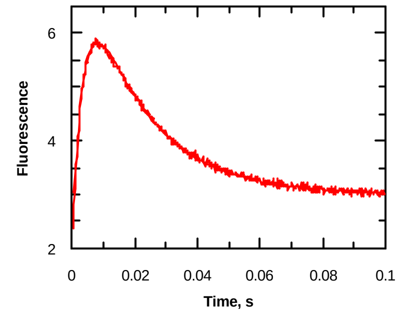
Reaction with serine with pyridoxal phosphate



$$E + Ser \underset{20s^{-1}}{\overset{0.135mM^{-1}s^{-1}}{\rightleftarrows}} E{\sim}Ser \underset{10s^{-1}}{\overset{45s^{-1}}{\rightleftarrows}} E{\sim}X$$

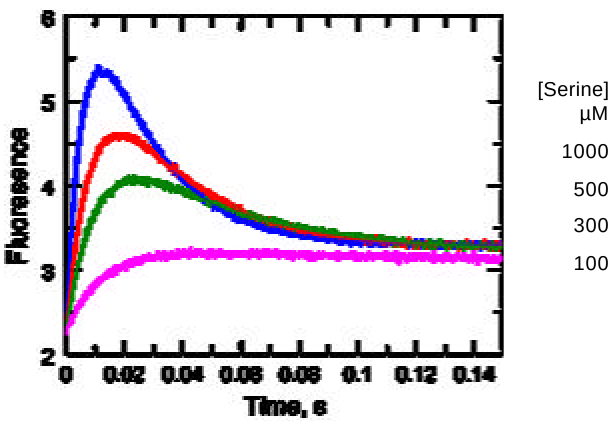
Tryptophan Synthase

$$E + Ser \underset{20s^{-1}}{\overset{0.135mM^{-1}s^{-1}}{\rightleftarrows}} E{\sim}Ser \underset{10s^{-1}}{\overset{45s^{-1}}{\rightleftarrows}} E{\sim}X$$



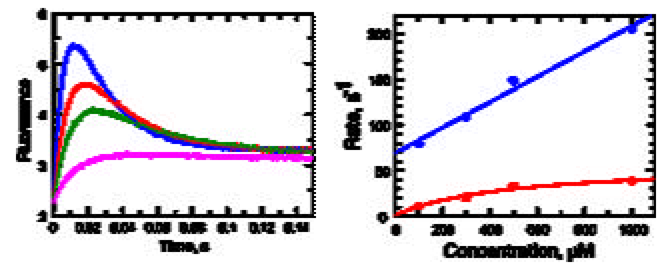
Anderson, K.A., Miles, E. W. and Johnson K. A. (1991) J. Biol. Chem 266, 8020-8033

Concentration dependence of fluorescence transient



Conventional Data Fitting – Example of simple two-step reaction

$$E + S \underset{k_{-1}}{\overset{k_1}{\rightleftarrows}} ES \underset{k_{-2}}{\overset{k_2}{\rightleftarrows}} EX$$



$$Y = A_1 \cdot e^{-l_1 t} + A_2 \cdot e^{-l_2 t} + C$$

Fit to 5 parameters for each curve. Disregard amplitudes and endpoints and plot rates versus concentration.

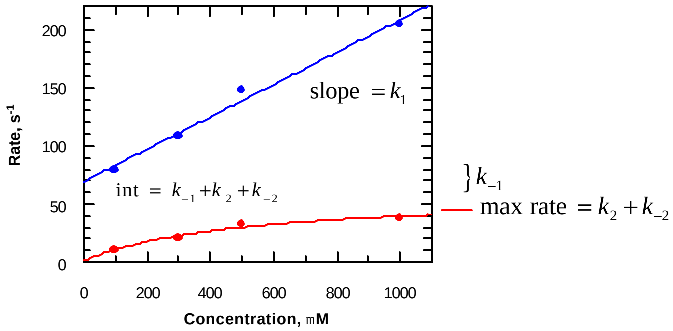
Fit to a total of 23 parameters and get only 3 rate constants!

$$l_1 \approx k_1[S] + k_{-1} + k_2 + k_{-2}$$

$$l_2 \approx \frac{k_1[S](k_2 + k_{-2}) + k_{-1}k_{-2}}{k_1[S] + k_{-1} + k_2 + k_{-2}}$$

Rates are roots of quadratic, simplified here. Get 3 of the 4 rate constants.

$$E+S \underset{20 s^{-1}}{\overset{0.135 mM^{-1}s^{-1}}{\rightleftharpoons}} E\sim S \underset{10 s^{-1}}{\overset{45 s^{-1}}{\rightleftharpoons}} E\sim X$$



$$l_1 \approx k_1[S]+k_{-1}+k_2+k_{-2} \qquad l_2 \approx \frac{k_1[S](k_2+k_{-2})+k_{-1}k_{-2}}{k_1[S]+k_{-1}+k_2+k_{-2}}$$

$$E+S \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} ES \underset{k_{-2}}{\overset{k_2}{\rightleftharpoons}} EX$$

$$l_{1,2}=\frac{(k_1[S]+k_{-1}+k_2+k_{-2})\pm\sqrt{(k_1[S]+k_{-1}+k_2+k_{-2})^2-4*(k_1[S]*(k_2+k_{-2})+k_{-1}k_2)}}{2}$$

if $(k_1[S]+k_{-1}+k_2+k_{-2})^2 \gg 4*(k_1[S]*(k_2+k_{-2})+k_{-1}k_2)$

then $l_1 \approx k_1[S]+k_{-1}+k_2+k_{-2}$

$$l_2 \approx \frac{k_1[S](k_2+k_{-2})+k_{-1}k_{-2}}{k_1[S]+k_{-1}+k_2+k_{-2}}$$

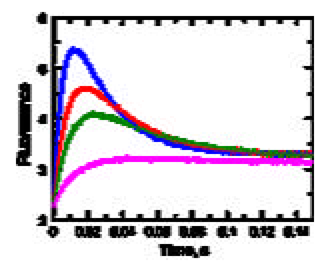
- Complex math for a two-step reaction
- Simplifications are approximations to reality

Conventional Data Fitting

1. Derive mathematical expressions for time dependence from a model
   - Almost always requires simplifying assumptions
   - Math soon gets complex (one exponential for each step)
2. Fit time dependence to mathematical expression to extract a rate
   - Fit to more independent variables than are relevant to the model
   - Loose relationships between rate and amplitude
3. Replot rate as a function of concentration
   - Observe patterns and develop model
4. Derive another mathematical expression to account for concentration dependence of the rate(s)
   - Requires more simplifying assumptions
5. Fit the concentration dependence to the mathematical expression to extract primary kinetic constants (k_{cat}, K_m, or rate constants)
   - Propagate errors through all steps of data fitting

Global Data Fitting based upon Simulation

$$E+S \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} ES \underset{k_{-2}}{\overset{k_2}{\rightleftharpoons}} EX$$

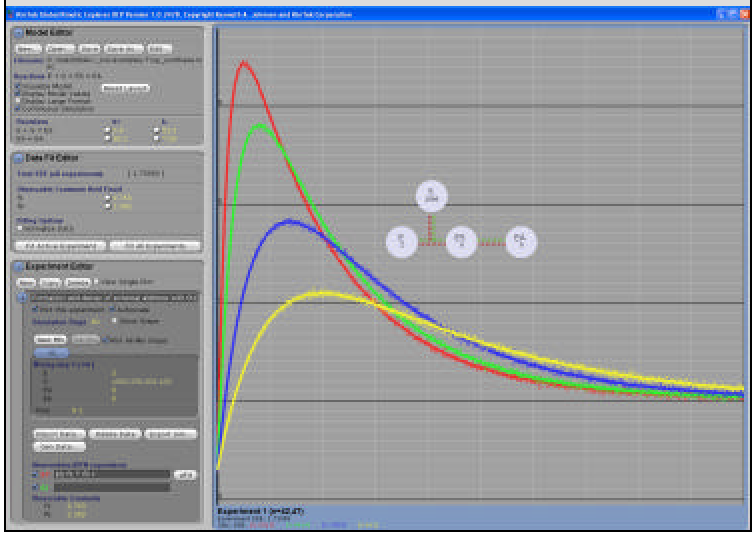


$$F = F_0 + \Delta F \cdot [ES]$$

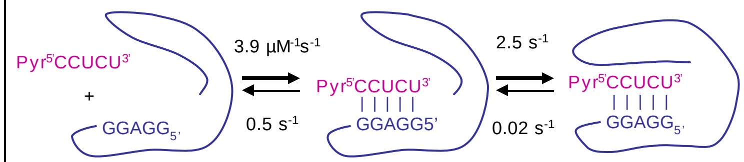
$$E+S \underset{20 s^{-1}}{\overset{0.135 mM^{-1}s^{-1}}{\rightleftharpoons}} E\sim S \underset{10 s^{-1}}{\overset{45 s^{-1}}{\rightleftharpoons}} E\sim X$$

Fit data directly to the model, get 4 rate constants and two fluorescence output factors.

Simultaneous Fitting All Four Curves by Simulation to the Model.



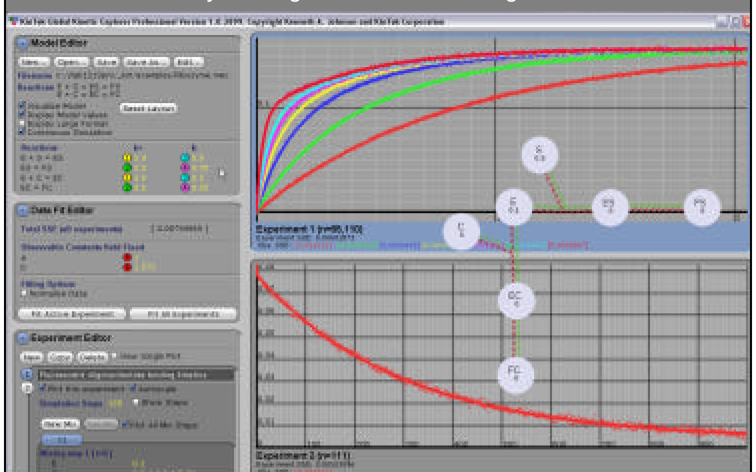
Two-step docking of oligonucleotide to hammerhead ribozyme



K_{-1} = 130 nM K_2 = 125 K_d = 1.1 nM

Bevilacqua, P. C., Kierzek, R., Johnson, K. A. and Turner, D. H. (1992). Dynamics of Ribozyme Binding of Substrate Revealed by Fluorescence Detected Stopped-Flow. *Science* 258, 1355-1358.

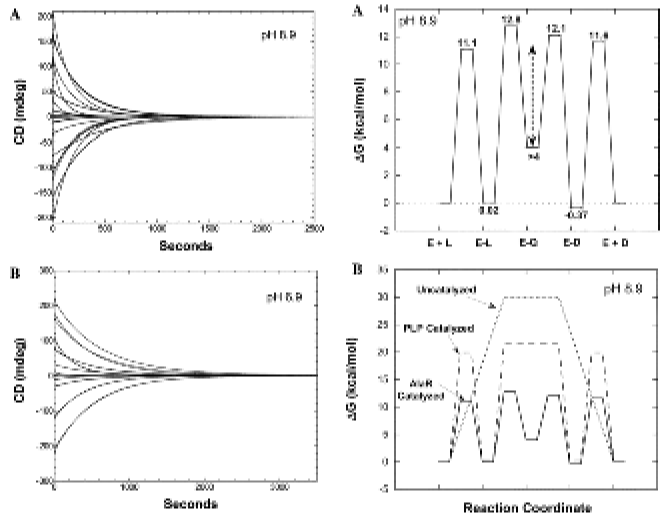
Ribozyme Oligonucleotide Binding Kinetics



Alanine Racemase: Free Energy Profiles from Global Analysis of Progress Curves.

$$E + L \underset{k_{-1}}{\overset{k_1}{\rightleftarrows}} EL \underset{k_{-2}}{\overset{k_2}{\rightleftarrows}} EI \underset{k_{-3}}{\overset{k_3}{\rightleftarrows}} ED \underset{k_{-4}}{\overset{k_4}{\rightleftarrows}} E + D$$

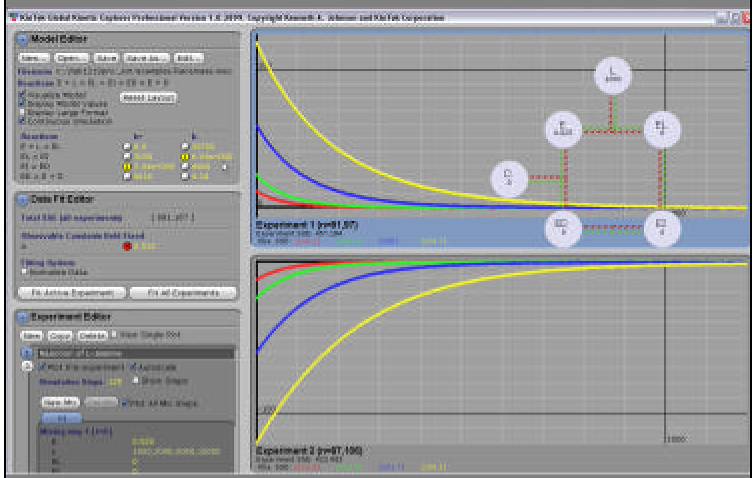
Spies, Woodward, Watnik and Toney (2004) JACS 126, 7464-7475



$$E + L \underset{39700}{\overset{8.9}{\rightleftharpoons}} EL \underset{6.64x10^{8}}{\overset{2650}{\rightleftharpoons}} EI \underset{4960}{\overset{2.44x10^{9}}{\rightleftharpoons}} ED \underset{4.18}{\overset{9610}{\rightleftharpoons}} E + D$$

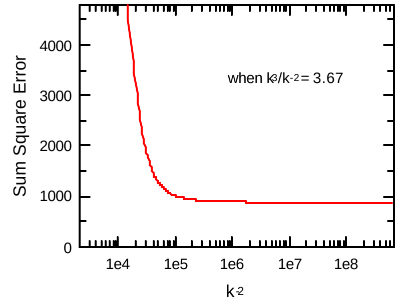
Can we get estimates for all 8 constants by global analysis of progress curves?

Alanine Racemase Progress Curves

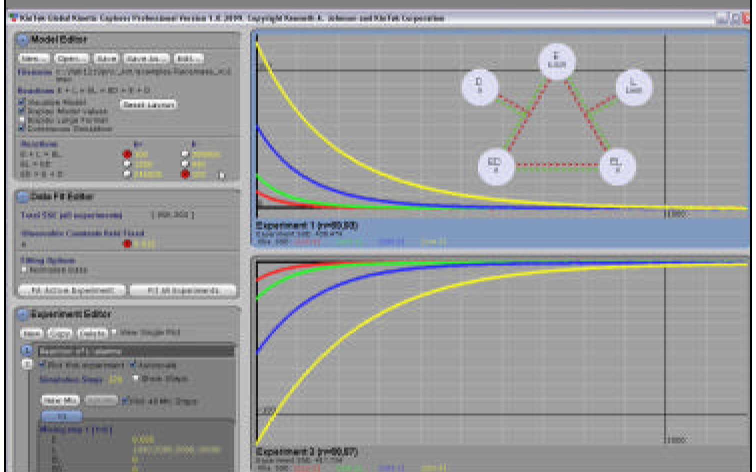


Alanine Racemase (Michael Toney rate constants)

$$E + L \underset{39700}{\overset{8.9}{\rightleftharpoons}} EL \underset{6.64x1\ 0^{8}}{\overset{2650}{\rightleftharpoons}} EI \underset{4960}{\overset{2.44x1\ 0^{9}}{\rightleftharpoons}} ED \underset{4.18}{\overset{9610}{\rightleftharpoons}} E + D$$



Alanine Racemase: Model without Intermediate



$$E + L \underset{39700}{\overset{8.9}{\rightleftharpoons}} EL \underset{6.64x10^{8}}{\overset{2650}{\rightleftharpoons}} EI \underset{4960}{\overset{2.44x10^{9}}{\rightleftharpoons}} ED \underset{4.18}{\overset{9610}{\rightleftharpoons}} E + D$$

$$E + L \underset{388000}{\overset{(100)}{\rightleftharpoons}} EL \underset{990}{\overset{1580}{\rightleftharpoons}} ED \underset{(100)}{\overset{245000}{\rightleftharpoons}} E + D$$

Only k_{cat} and K_m (or k_{cat}/K_m) can be determined

$$E + L \underset{39700}{\overset{8.9}{\rightleftharpoons}} EL \underset{6.64x10^{8}}{\overset{2650}{\rightleftharpoons}} EI \underset{4960}{\overset{2.44x10^{9}}{\rightleftharpoons}} ED \underset{4.18}{\overset{9610}{\rightleftharpoons}} E + D$$

	k_{cat} (s^{-1})	k_{cat}/K_m ($\mu M^{-1}s^{-1}$)
Forward	1570	0.40
Reverse	980	0.40

$$E + L \underset{388000}{\overset{(100)}{\rightleftharpoons}} EL \underset{990}{\overset{1580}{\rightleftharpoons}} ED \underset{(100)}{\overset{245000}{\rightleftharpoons}} E + D$$

	k_{cat} (s^{-1})	k_{cat}/K_m ($\mu M^{-1}s^{-1}$)
Forward	1563	0.40
Reverse	983	0.40

Only k_{cat} and K_m (or k_{cat}/K_m) can be determined

$$E + L \underset{388000}{\overset{(100)}{\rightleftharpoons}} EL \underset{990}{\overset{1580}{\rightleftharpoons}} ED \underset{(100)}{\overset{245000}{\rightleftharpoons}} E + D$$

	k_{cat} (s^{-1})	k_{cat}/K_m ($\mu M^{-1}s^{-1}$)
Forward	1580	0.40
Reverse	990	0.40

$$E + L \underset{6650}{\overset{1.46}{\rightleftharpoons}} EL \underset{1790}{\overset{3590}{\rightleftharpoons}} ED \underset{1.83}{\overset{4190}{\rightleftharpoons}} E + D$$

	k_{cat} (s^{-1})	k_{cat}/K_m ($\mu M^{-1}s^{-1}$)
Forward	1572	0.40
Reverse	989	0.40

$$E + S \underset{k_{-1}}{\overset{k_1}{\rightleftarrows}} E{\cdot}S \underset{k_{-2}}{\overset{k_2}{\rightleftarrows}} E{\cdot}P \underset{k_{-3}}{\overset{k_3}{\rightleftarrows}} E + P$$

Forward

$$k_{cat,S} = \frac{k_2 k_3}{k_2 + k_{-2} + k_3}$$

$$K_{m,S} = \frac{k_2 k_3 + k_{-1}(k_{-2} + k_3)}{k_1(k_2 + k_{-2} + k_3)}$$

$$k_{cat,S} / K_{m,S} = \frac{k_1 k_2 k_3}{k_2 k_3 + k_{-1}(k_{-2} + k_3)}$$

Reverse

$$k_{cat,P} = \frac{k_{-1} k_{-2}}{k_2 + k_{-2} + k_{-1}}$$

$$K_{m,P} = \frac{k_2 k_3 + k_{-1}(k_{-2} + k_3)}{k_{-3}(k_2 + k_{-2} + k_{-1})}$$

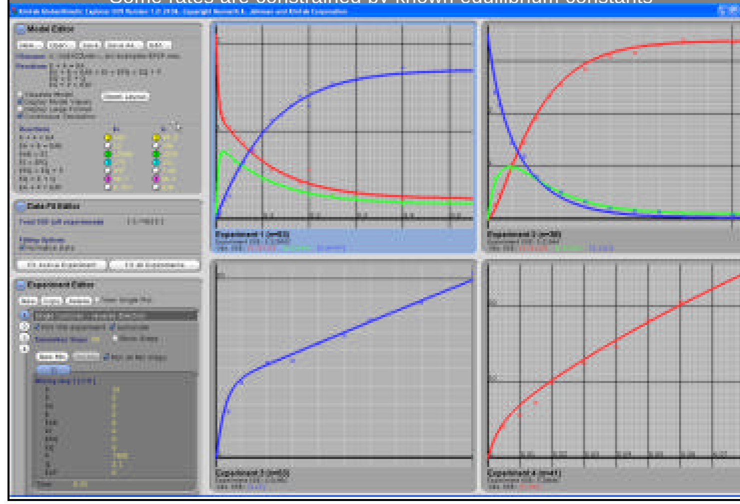
$$k_{cat,P} / K_{m,P} = \frac{k_{-1} k_{-2} k_{-3}}{k_2 k_3 + k_{-1}(k_{-2} + k_3)}$$

$$\frac{k_{cat,S} / K_{m,S}}{k_{cat,P} / K_{m,P}} = \frac{k_1 k_2 k_3}{k_{-1} k_{-2} k_{-3}} = K_{net}$$

Only 4 pieces of information: concentration dependence and maximum rate in forward and reverse direction

EPSP Synthase Example: Fit 4 experiments simultaneously

Some rates are constrained by known equilibrium constants



Guidelines for Data Fitting by Simulation

1. Conventional data fitting to defined equations can illuminate patterns in the data, help construct a model, and get initial estimates of rate constants
2. Enter complete model including all known steps
3. Constrain the fitting of constants that are not determined by the data
4. Evaluate the fit using dynamic simulation to explore the space over which data can be fitted
5. Keep the number of fitted parameters within the range of the information content of the data

Do not over-interpret the data!


Zachary Booth Simpson
www.Mine-Control.com