Subdiffusion and reaction networks in single-molecule biophysics

Samuel Kou
Department of Statistics
Harvard University

Single-Molecule Experiments

- Experimental advances make it possible to study biochemical/biological progresses at **single-molecule level**
- Example: single-molecule enzymatic kinetics, single-molecule protein/DNA dynamics

"Everything that living things can do can be understood in terms of the jigglings and wigglings of atoms"

--- Richard Feynman

Single-Molecule Advantages

- Can measure molecular properties ***individually***, instead of inferring from ensemble population statistics
- If the reaction/kinetic time is slow, ensemble experiments become almost impossible due to the ***difficulty of synchronization***
- Single-molecule trajectory provides detailed ***dynamic information***. In particular capture transient intermediates

Stochastic Challenges

- Require new stochastic modeling

  Subdiffusion

  Enzymatic reaction

- Data are noisier, and efficient inference methodology is needed

Michaelis-Menten Mechnism

$$E + S \underset{k_{-1}}{\overset{k_1}{\rightleftarrows}} ES \xrightarrow{k_2} E^0 + P, \qquad E^0 \xrightarrow{k_3} E$$

$$\frac{d[E]}{dt} = -k_1[E][S] + k_{-1}[ES]$$

$$\frac{d[ES]}{dt} = k_1[E][S] - (k_{-1} + k_2)[ES]$$

$$\frac{d[E^0]}{dt} = \frac{d[P]}{dt} = k_2[ES]$$

⇩

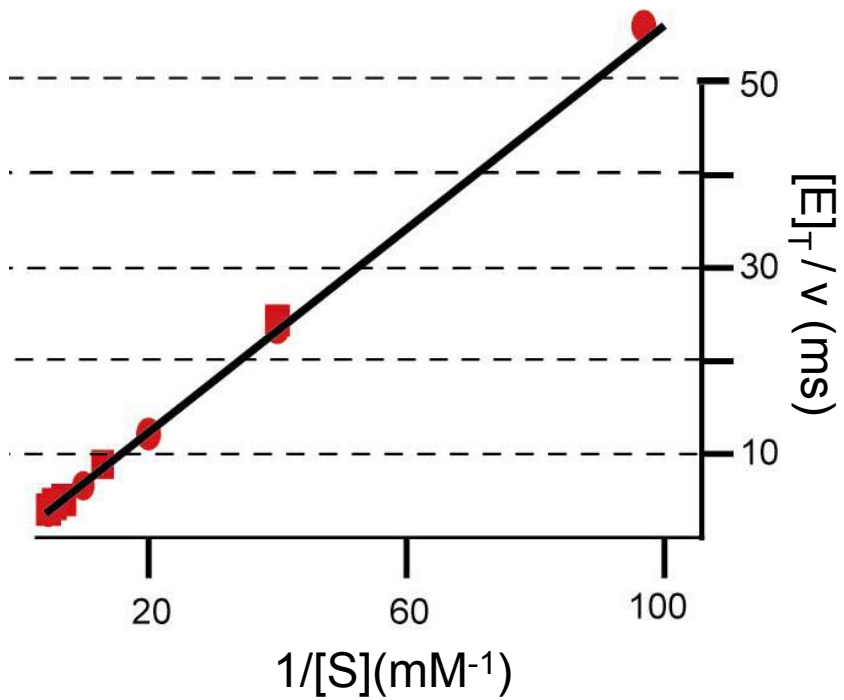
Classical Michaelis-Menten equation

$$v = \frac{v_{\max}[S]}{[S] + K_M},$$

$$v_{\max} = k_2([E] + [ES])$$

$$K_M = (k_{-1} + k_2)/k_1$$

Lineweaver-Burke plot



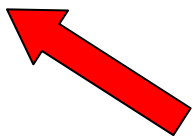
Single-molecule case

$$E + S \underset{k_{-1}}{\overset{k_1}{\rightleftarrows}} ES \xrightarrow{k_2} E^0 + P$$

$$\frac{dP_E(t)}{dt} = -k_1[S]P_E(t) + k_{-1}P_{ES}(t)$$

$$\frac{dP_{ES}(t)}{dt} = k_1[S]P_E(t) - (k_{-1} + k_2)P_{ES}(t)$$

$$\frac{dP_{E^0}(t)}{dt} = k_2 P_{ES}(t)$$



Three-state continuous-time Markov chain

Turnover time:
first passage time

$$Q = \begin{pmatrix} -k_1[S] & k_1[S] & 0 \\ k_{-1} & -(k_{-1} + k_2) & k_2 \\ k_3 & 0 & -k_3 \end{pmatrix}$$

Single-molecule case

$$E + S \underset{k_{-1}}{\overset{k_1}{\rightleftarrows}} ES \xrightarrow{k_2} E^0 + P$$

$$\frac{dP_E(t)}{dt} = -k_1[S]P_E(t) + k_{-1}P_{ES}(t)$$

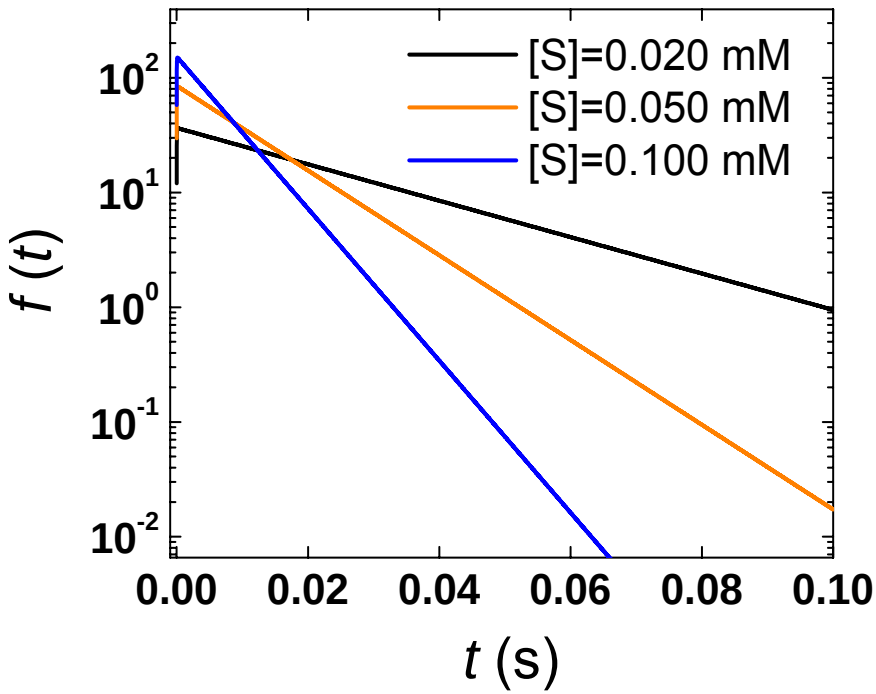
$$\frac{dP_{ES}(t)}{dt} = k_1[S]P_E(t) - (k_{-1} + k_2)P_{ES}(t)$$

$$\frac{dP_{E^0}(t)}{dt} = k_2 P_{ES}(t)$$

Turnover time distribution

$$f(t) = \frac{k_1 k_2 [S]}{2A}\left[\exp(A+B)t - \exp(B-A)t\right]$$

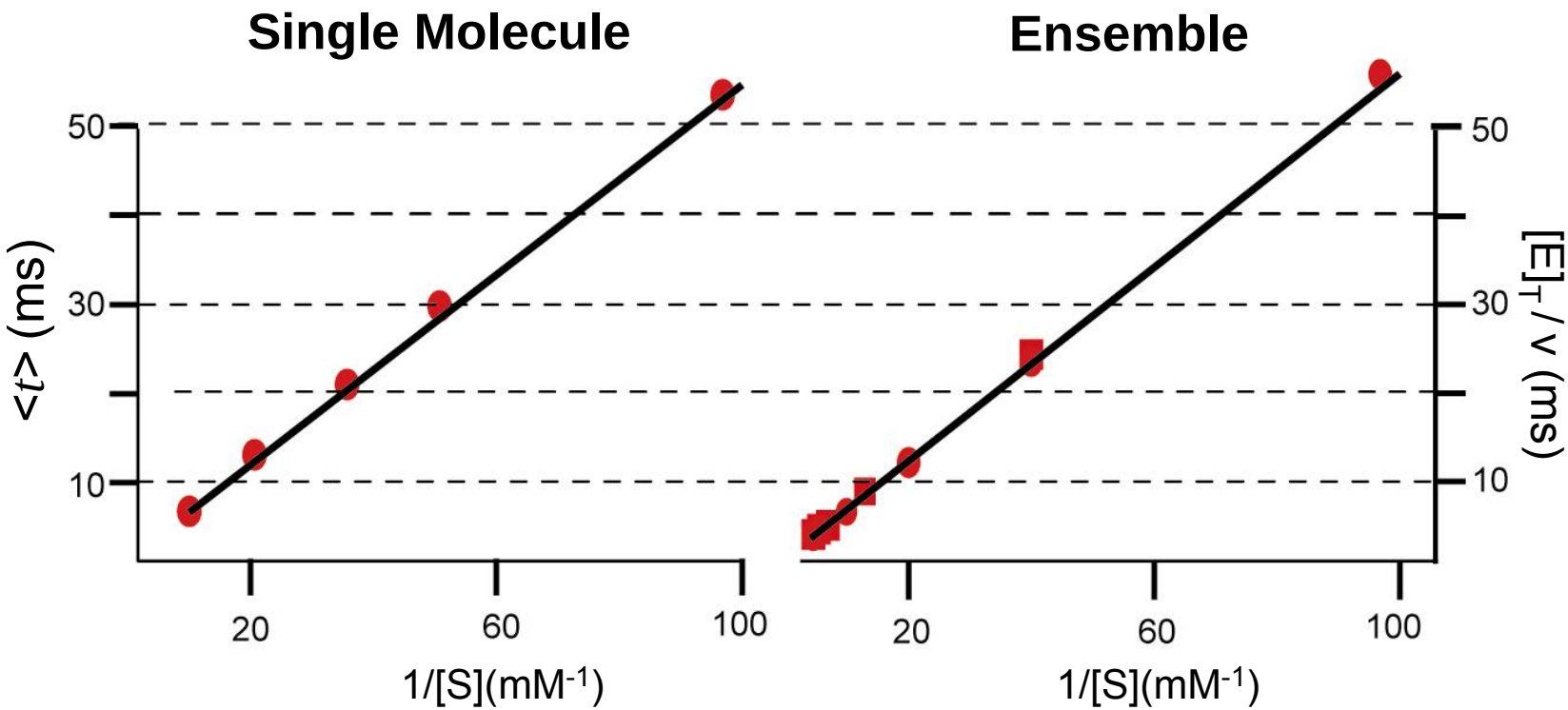
$$A = \sqrt{B^2 - k_1 k_2 [S]},\ B = -(k_1[S] + k_{-1} + k_2)/2$$



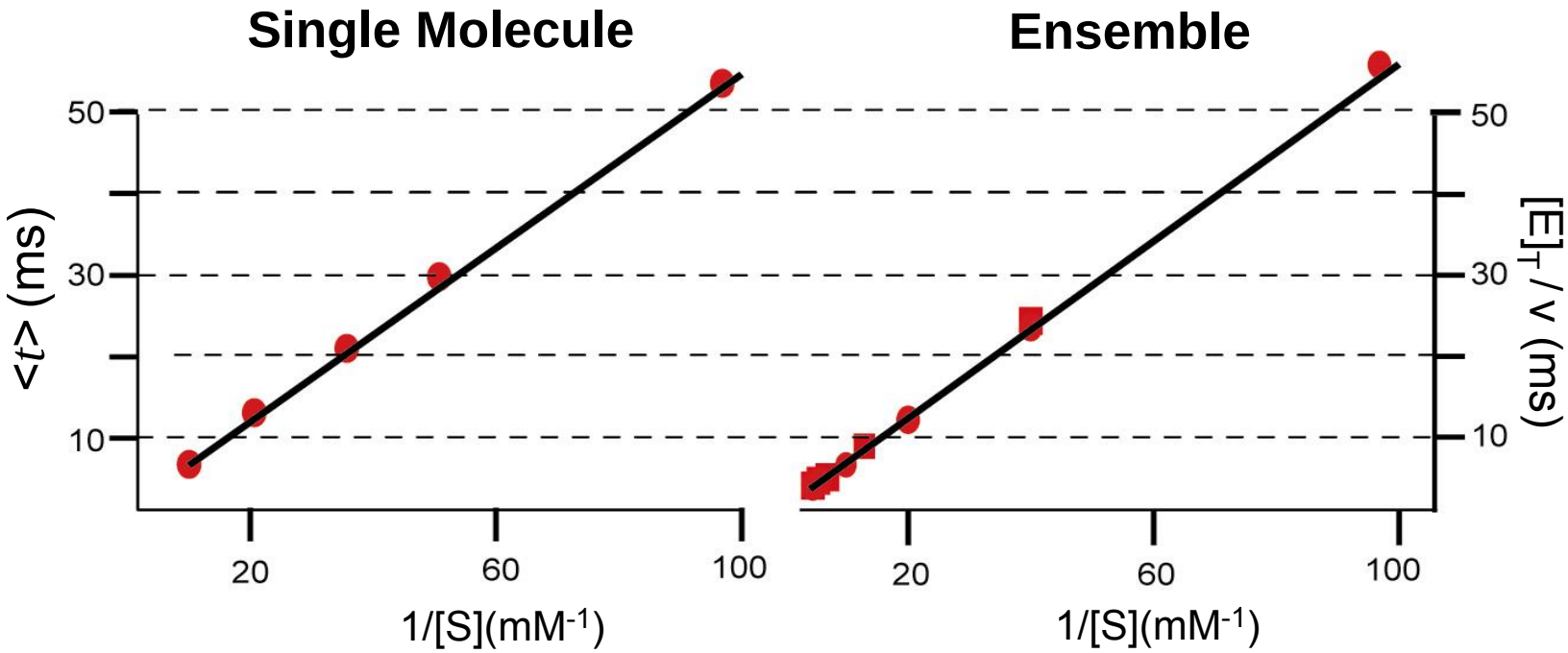
Reaction rate:

$$v = \frac{1}{\langle t \rangle} = \frac{k_2[S]}{[S] + K_M}$$

: Still obey the hyperbolic form

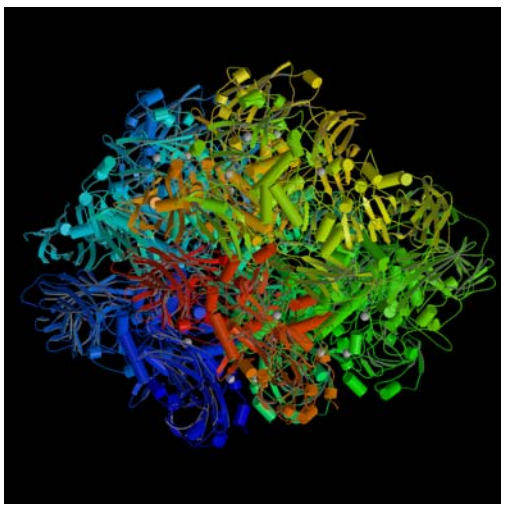


Obtained from English *et al*. (Nature Chem. Biol., 2, 87, 2006)'s single-molecule study of β-galactosidase



English *et al*., Nature Chem. Biol., 2, 87 (2006)

β-galactosidase



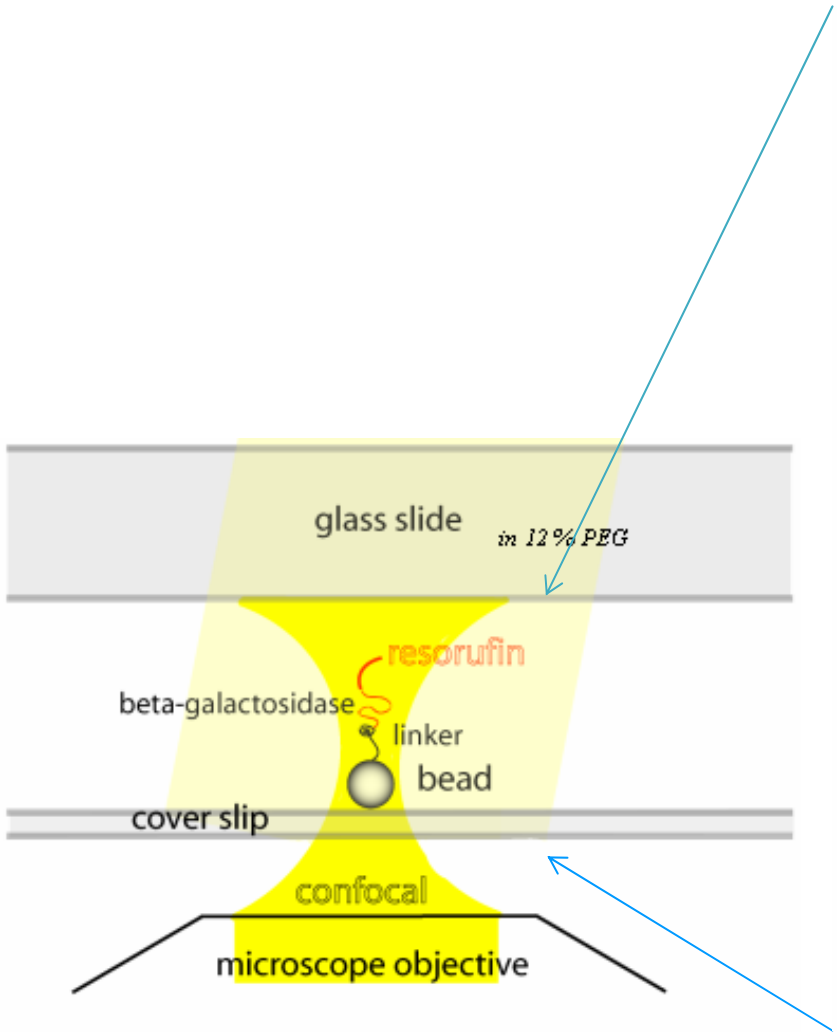
E. coli ß-gal catalyzes Hydrolysis of Lactose

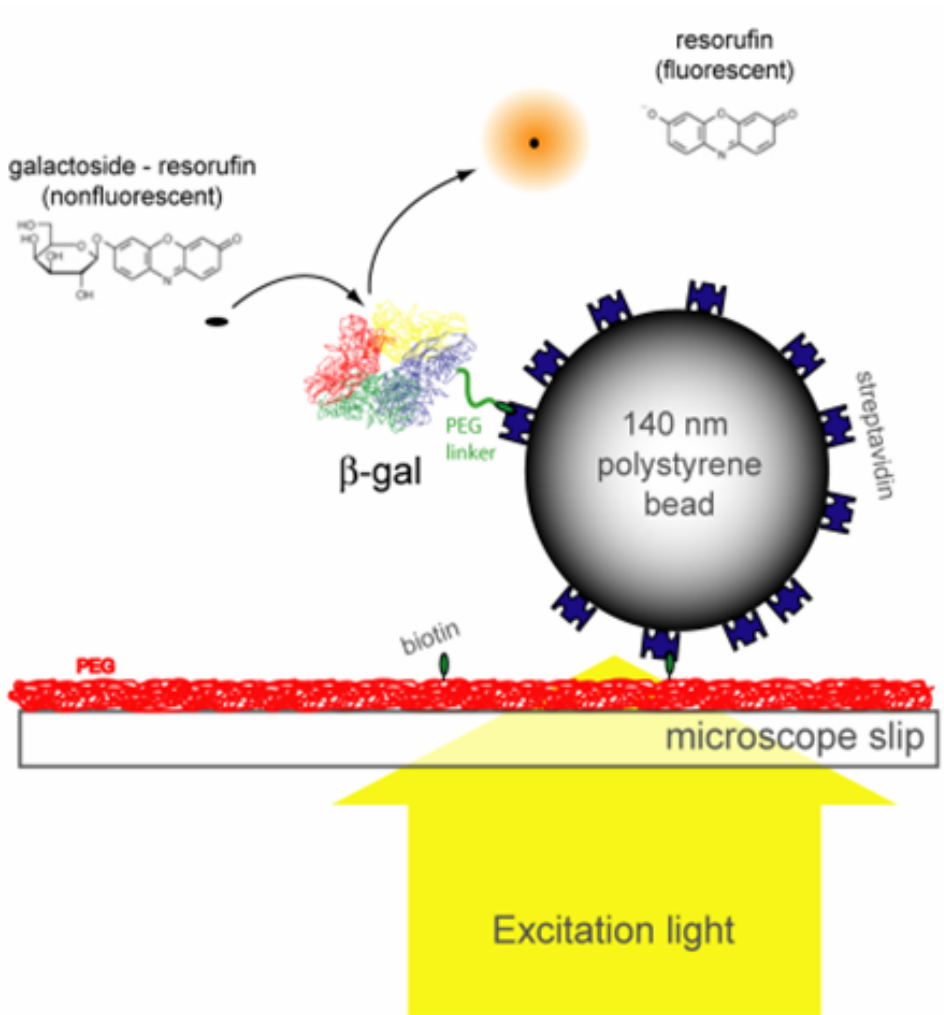
The experiment uses **photogenic** substrate

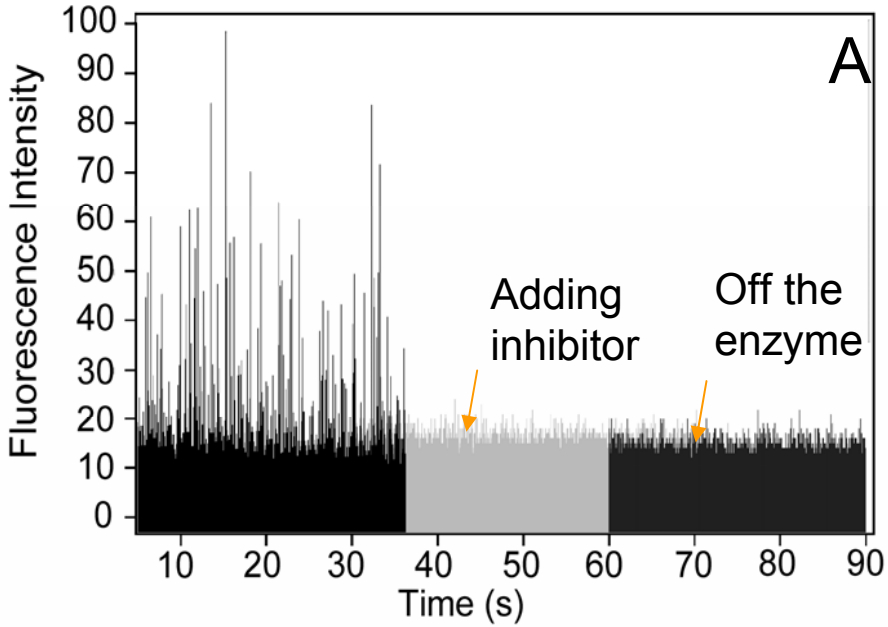
$$E + S \underset{k_{-1}}{\overset{k_1}{\rightleftarrows}} ES \xrightarrow{k_2} E^0 + P$$

Single Molecule Turnover Experiment of ß-galactosidase

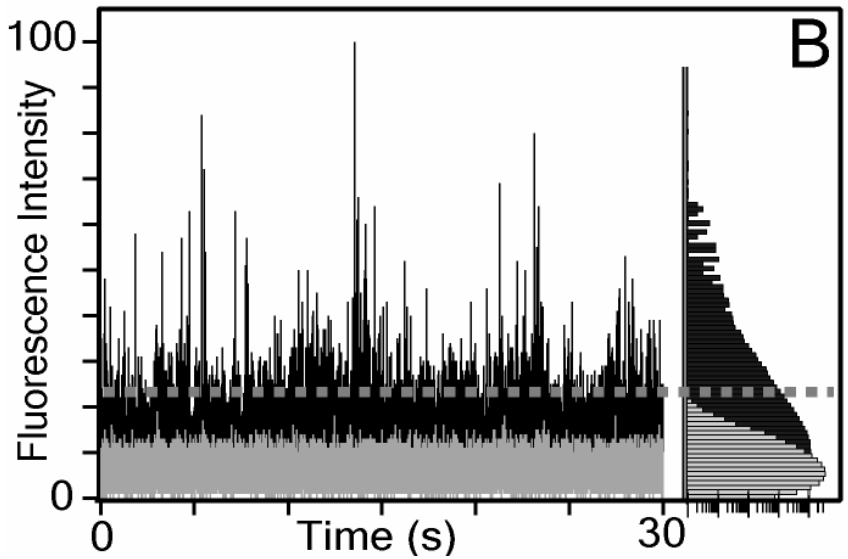
Each enzymatic turnover creates a fluorescent burst





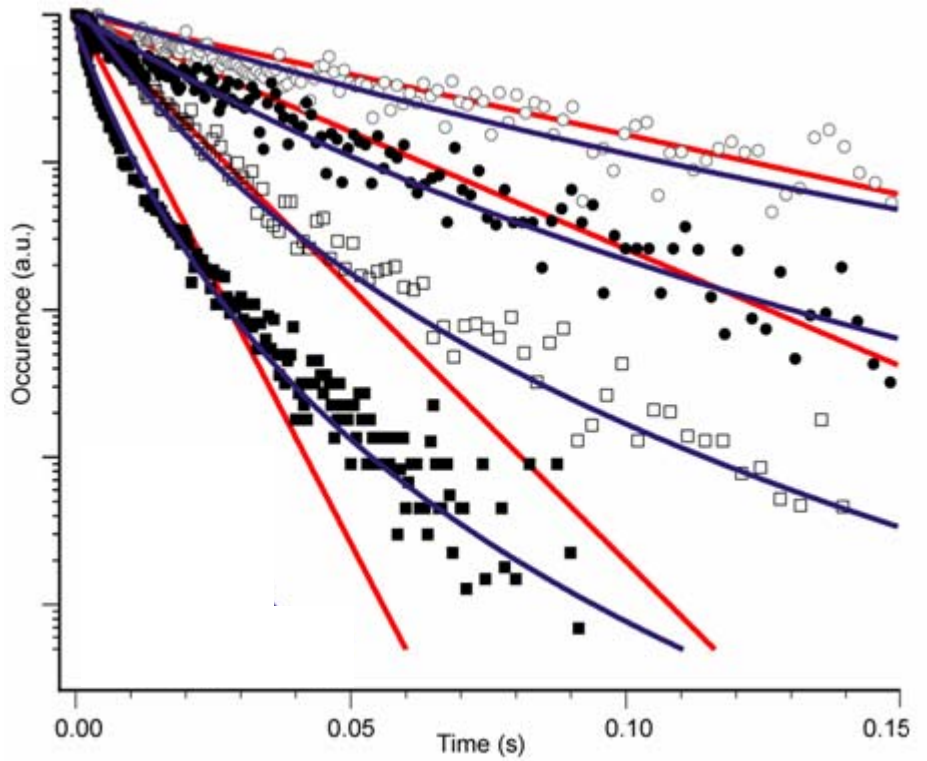


Low Substrate Concentration 20μM



High Substrate Concentration 100μM

Multi-exponential Distributions of Turnover Times



[S]=10 μM

[S]=20 μM

[S]=50 μM

[S]=100 μM

- Multi-exponential decay at high substrate concentration
- Mono-exponential decay at low substrate concentration

Memory between successive turnover times

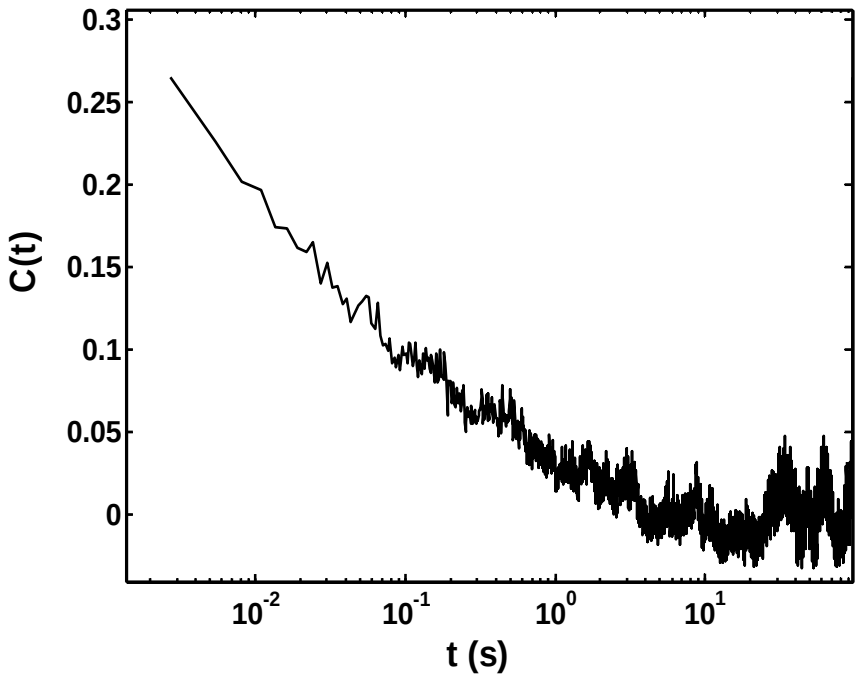
- Under Michaelis-Menten Mechnism

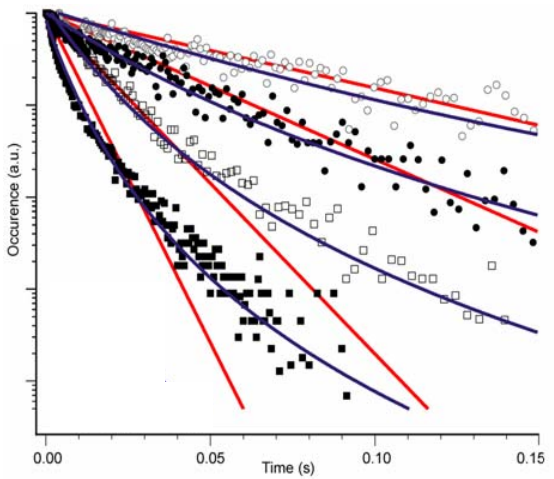
$$E + S \underset{k_{-1}}{\overset{k_1}{\rightleftarrows}} ES \xrightarrow{k_2} E^0 + P$$

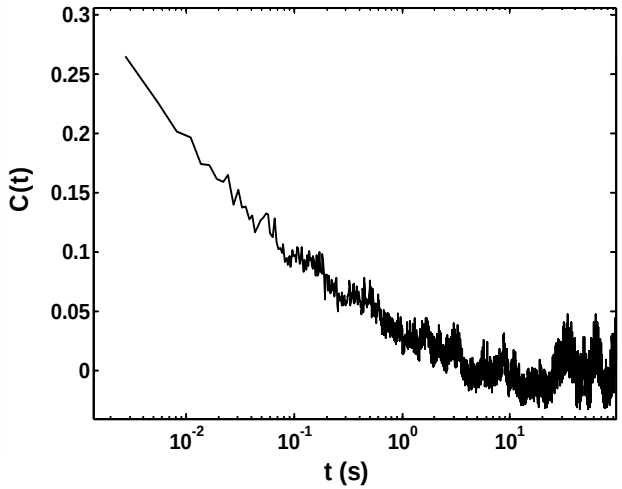
three-state continuous-time **Markov** chain

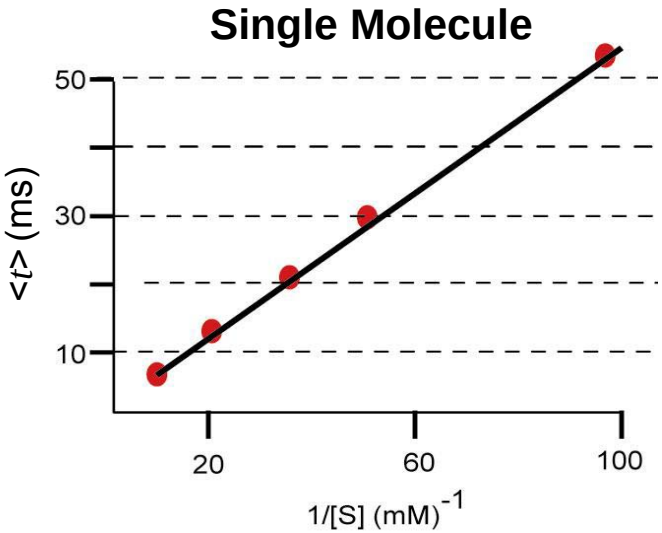
- Successive turnover times should have **NO** correlation

Experimental data




Occurence (a.u.)
0.00
0.05
Time (s)
0.10
0.15


C(t)
0.3
0.25
0.2
0.15
0.1
0.05
0
10^{-2}
10^{-1}
10^{0}
10^{1}
t (s)

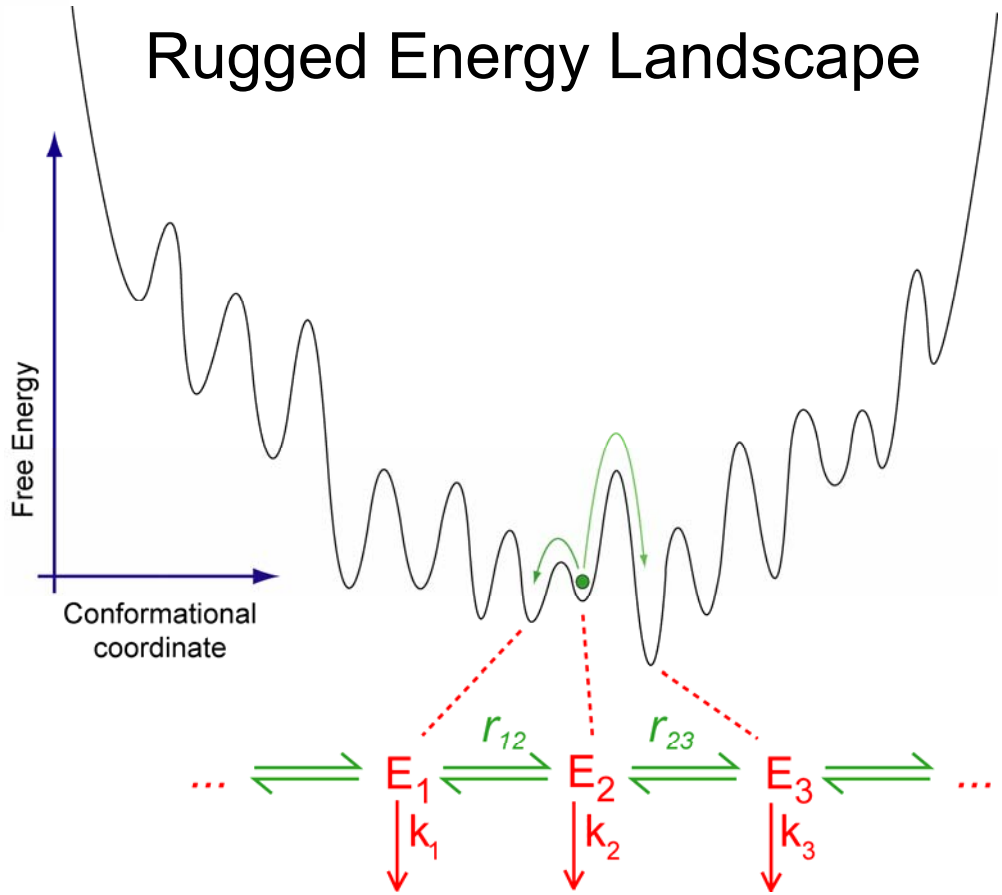

Single Molecule
50
30
10
<t> (ms)
20
60
100
1/[S] (mM)$^{-1}$

Dynamic disorder – the fluctuation of enzyme

$$
\begin{array}{ccccccc}
S + E_1 & \underset{k_{-11}}{\overset{k_{11}[S]}{\rightleftarrows}} & ES_1 & \overset{k_{21}}{\rightarrow} & P + E_1^0 & \overset{\delta_1}{\rightarrow} & E_1 \quad \cdots \\
\downarrow\uparrow & & \downarrow\uparrow & & \downarrow\uparrow & & \cdots \quad \cdots \\
S + E_2 & \underset{k_{-12}}{\overset{k_{12}[S]}{\rightleftarrows}} & ES_2 & \overset{k_{22}}{\rightarrow} & P + E_2^0 & \overset{\delta_2}{\rightarrow} & E_2 \quad \cdots \\
\cdots & & \cdots & & \cdots & & \cdots \\
\downarrow\uparrow & & \downarrow\uparrow & & \downarrow\uparrow & & \\
S + E_n & \underset{k_{-1n}}{\overset{k_{1n}[S]}{\rightleftarrows}} & ES_n & \overset{k_{2n}}{\rightarrow} & P + E_n^0 & \overset{\delta_n}{\rightarrow} & E_n \quad \cdots
\end{array}
$$

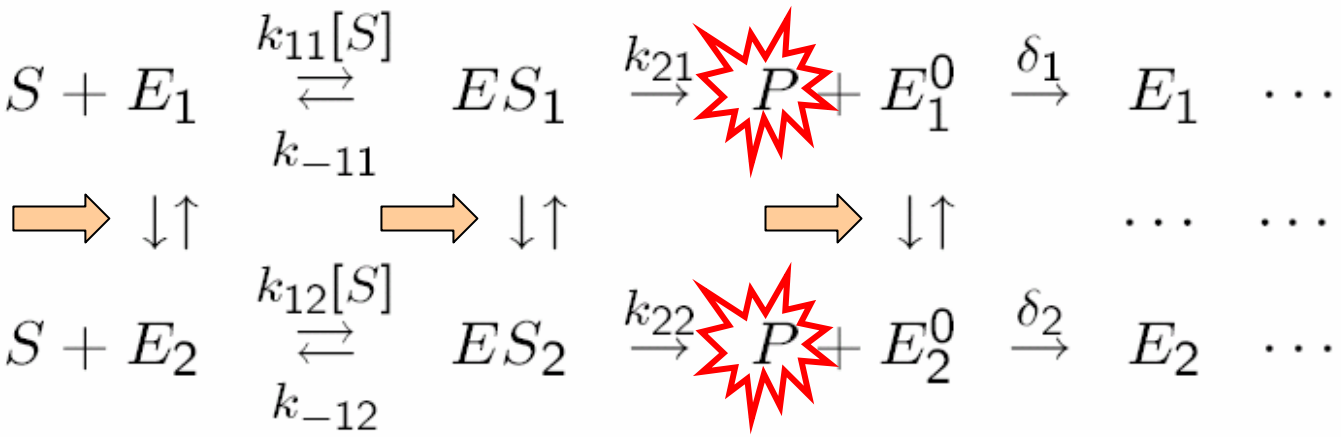
Dynamic disorder – the fluctuation of enzyme

$$
\begin{array}{ccccccc}
S+E_1 & \underset{k_{-11}}{\overset{k_{11}[S]}{\rightleftarrows}} & ES_1 & \overset{k_{21}}{\rightarrow} & P+E_1^0 & \overset{\delta_1}{\rightarrow} & E_1 \quad \cdots \\
\downarrow\uparrow & & \downarrow\uparrow & & \downarrow\uparrow & & \cdots \quad \cdots \\
S+E_2 & \underset{k_{-12}}{\overset{k_{12}[S]}{\rightleftarrows}} & ES_2 & \overset{k_{22}}{\rightarrow} & P+E_2^0 & \overset{\delta_2}{\rightarrow} & E_2 \quad \cdots \\
\cdots & & \cdots & & \cdots & & \cdots \\
\downarrow\uparrow & & \downarrow\uparrow & & \downarrow\uparrow & & \\
S+E_n & \underset{k_{-1n}}{\overset{k_{1n}[S]}{\rightleftarrows}} & ES_n & \overset{k_{2n}}{\rightarrow} & P+E_n^0 & \overset{\delta_n}{\rightarrow} & E_n \quad .
\end{array}
$$



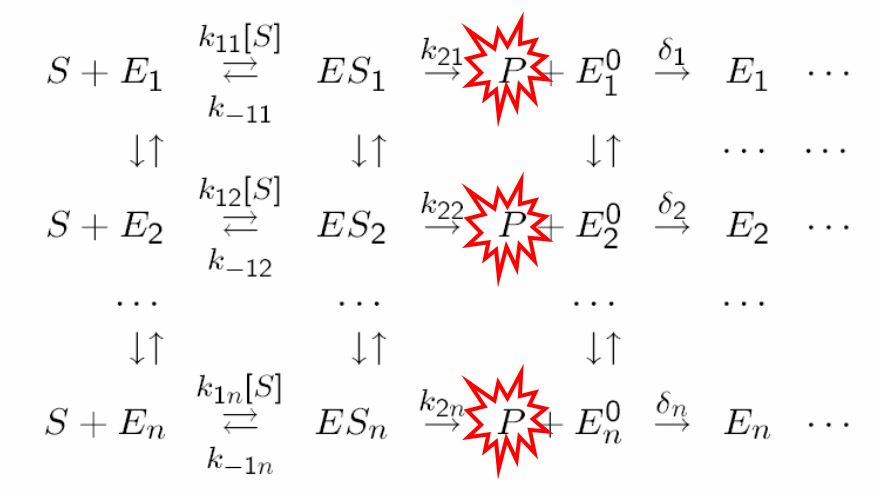
Explain the memory

If conformers intervert slowly



Transitions will stay in one channel for a quite while before going to the next

Naturally give rise to the strong correlation



$$\begin{array}{ccccccc}
S+E_1 & \underset{k_{-11}}{\overset{k_{11}[S]}{\rightleftarrows}} & ES_1 & \xrightarrow{k_{21}} & P+E_1^0 & \xrightarrow{\delta_1} & E_1 \quad \cdots \\
\downarrow\uparrow & & \downarrow\uparrow & & \downarrow\uparrow & & \cdots \quad \cdots \\
S+E_2 & \underset{k_{-12}}{\overset{k_{12}[S]}{\rightleftarrows}} & ES_2 & \xrightarrow{k_{22}} & P+E_2^0 & \xrightarrow{\delta_2} & E_2 \quad \cdots \\
\cdots & & \cdots & & \cdots & & \cdots \\
\downarrow\uparrow & & \downarrow\uparrow & & \downarrow\uparrow & & \\
S+E_n & \underset{k_{-1n}}{\overset{k_{1n}[S]}{\rightleftarrows}} & ES_n & \xrightarrow{k_{2n}} & P+E_n^0 & \xrightarrow{\delta_n} & E_n \quad \cdots
\end{array}$$

Can be solved via Laplace transform and matrix algebra

Kou *et al.*, *J. Phys. Chem.* B, 109, 19068 (2005)

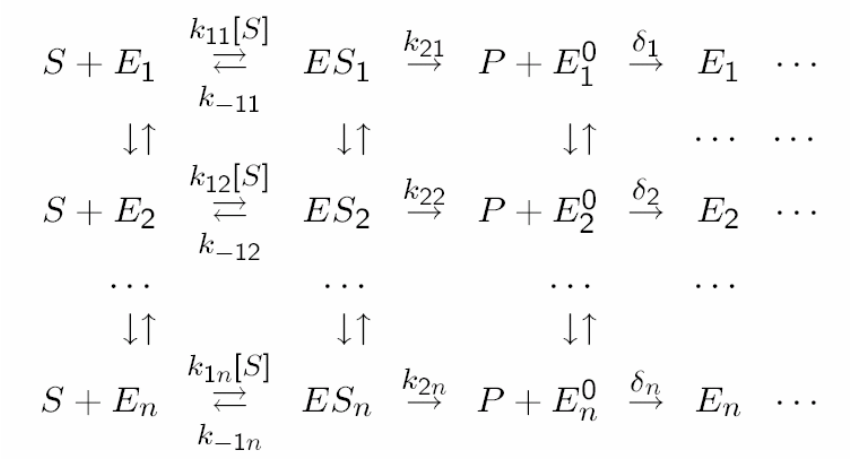
Under physically meaningful conditions

Single Molecule Michaelis-Menten Equation

$$v = \frac{1}{\langle t \rangle} = \frac{\gamma_2 [S]}{[S] + C_M}$$

$$\gamma_2 = \left[\int_0^\infty \frac{p(k_2)}{k_2} dk_2 \right]^{-1} \qquad C_M = \frac{\gamma_2 + k_{-1}}{k_1}$$

Weighted harmonic mean



Reaction rate

$$\nu = \frac{1}{\langle t \rangle} = \frac{\gamma_2 [S]}{[S] + C_M}$$

If one of the following conditions holds:

(a) slow interconversion between E_i

(b) slow interconversion between ES_i

} Quasi static

(c) fast interconversion between E_i

(d) fast interconversion between ES_i

(e) $k_{-1i} >> k_{2i}$ Quasi equilibrium

(f) k_{2i}/k_{-1i} = const Conformational equilibrium

Turnover Time distribution for fluctuating enzymes

$$
\begin{array}{ccccccc}
S+E_1 & \underset{k_{-11}}{\overset{k_{11}[S]}{\rightleftarrows}} & ES_1 & \overset{k_{21}}{\rightarrow} & P+E_1^0 & \overset{\delta_1}{\rightarrow} & E_1 \quad \cdots \\
\downarrow\uparrow & & \downarrow\uparrow & & \downarrow\uparrow & & \cdots \quad \cdots \\
S+E_2 & \underset{k_{-12}}{\overset{k_{12}[S]}{\rightleftarrows}} & ES_2 & \overset{k_{22}}{\rightarrow} & P+E_2^0 & \overset{\delta_2}{\rightarrow} & E_2 \quad \cdots \\
\cdots & & \cdots & & \cdots & & \cdots \\
\downarrow\uparrow & & \downarrow\uparrow & & \downarrow\uparrow & & \\
S+E_n & \underset{k_{-1n}}{\overset{k_{1n}[S]}{\rightleftarrows}} & ES_n & \overset{k_{2n}}{\rightarrow} & P+E_n^0 & \overset{\delta_n}{\rightarrow} & E_n \quad \cdots
\end{array}
$$

Quasi Static Limit:

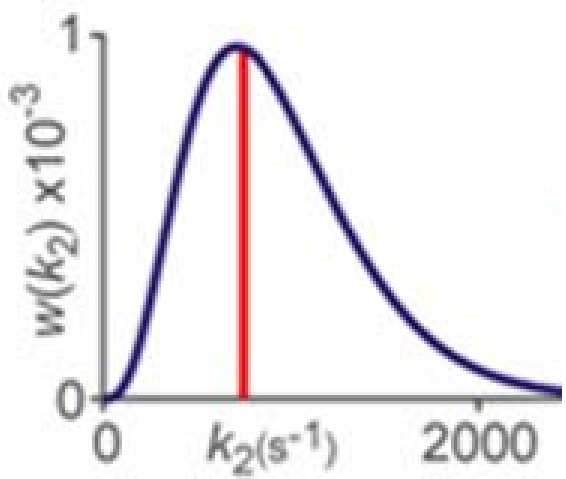
$$f(t) = \int_0^\infty dk_2\, w(k_2)\, \frac{k_1 k_2 [S]}{2A} \left[\exp(A+B)t - \exp(B-A)t\right]$$

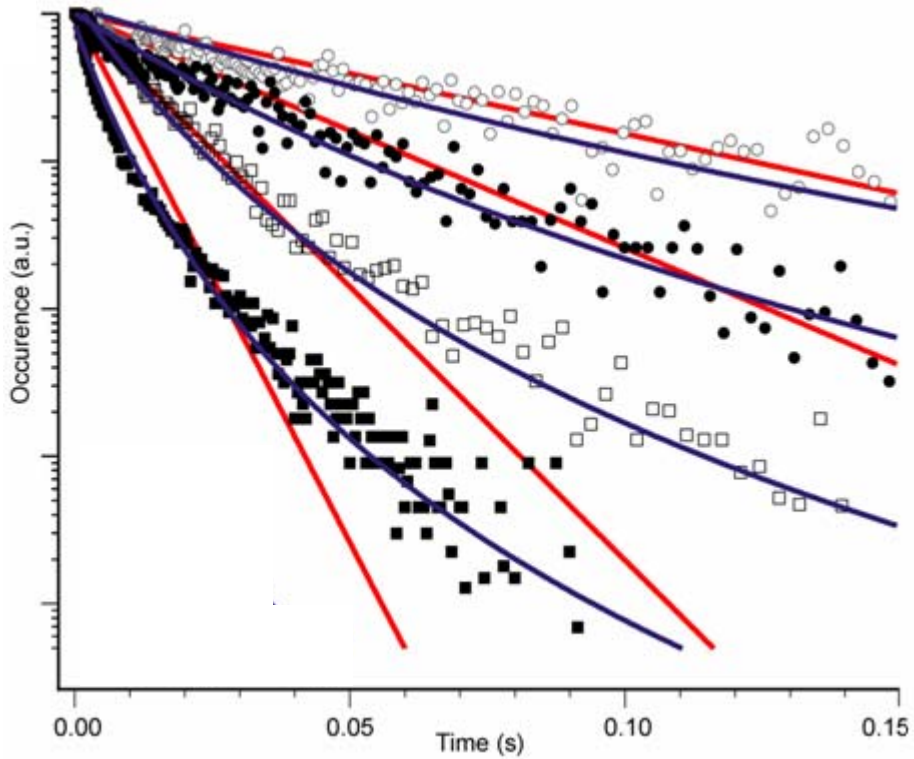
$$A = \sqrt{B^2 - k_1 k_2 [S]},\ B = -(k_1[S] + k_{-1} + k_2)/2$$

Multi-exponential Distributions of Turnover Times

Gamma Distribution

$$w(k_2) = \frac{1}{b^a \Gamma(a)} k_2^{\,a-1} \exp(-k_2 / b)$$





[S]=10 μM

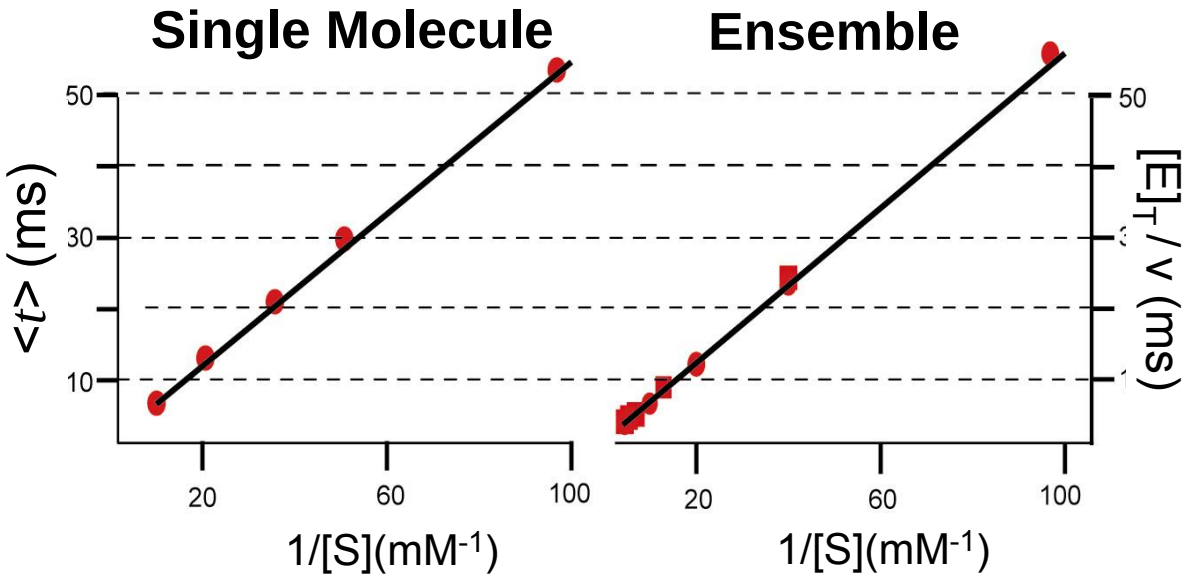
[S]=20 μM

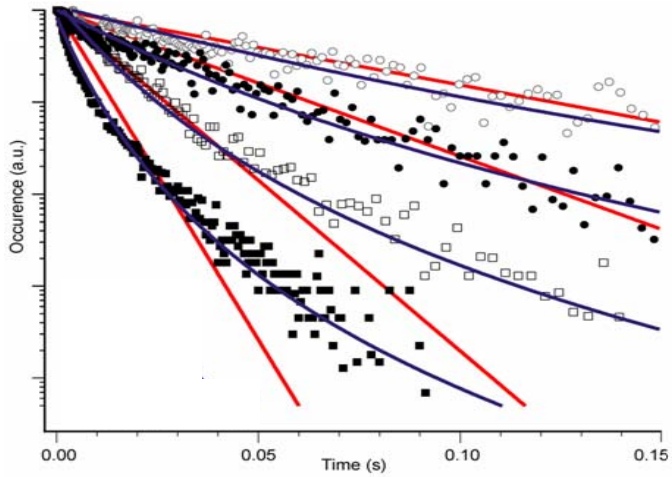
[S]=50 μM

[S]=100 μM

$k_1 = 5 \times 10^7\ \mathrm{M^{-1}s^{-1}}$ $\quad k_{-1} = 1.83 \times 10^4\ \mathrm{s^{-1}}$

$a = 4.2,\ b = 224\ \mathrm{s^{-1}}$





Consequences:

(a) Single Molecule reconciled with ensemble study

(b) Dynamic disorder might be masked in ensemble studies!

(c) The apparent k_2 and K_M of the Michaelis-Menten equation are complex functions of the k_2 and K_M of a large distribution of conformers, different from their conventional interpretations.

Brownian diffusion

- Brownian motion described by Langevin equation

$$m\frac{dv_t}{dt} = -\zeta v_t + F(t)$$

where $F(t)$ is white noise process satisfying

$$\langle F(t)F(t')\rangle = \zeta k_B T \cdot \delta(t-t')$$

by **fluctuation-dissipation** theorem.

- Solving the equation
mean square displacement $x(t) = \int_0^t v(s)ds$

$$\langle x(t)^2\rangle \sim 2\frac{k_B T}{\zeta}t, \text{ for large } t$$

--- Corner stone for statistical mechanics

Subdiffusion

- Brownian diffusion, however, cannot explain subdiffusion:

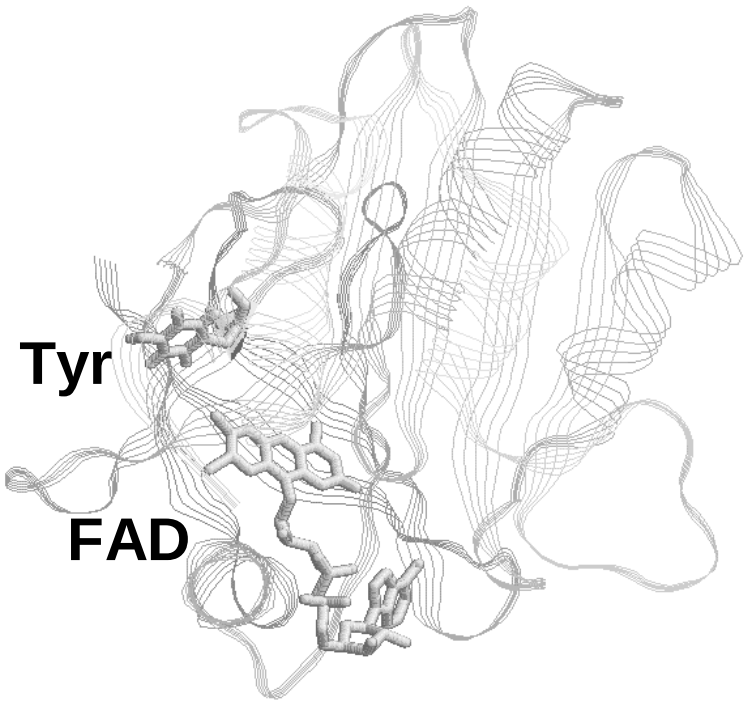
$$\langle x(t)^2 \rangle \propto t^{\alpha}, \quad 0 < \alpha < 1$$

- Distance fluctuation between electron transfer donor and acceptor within a single protein molecule

Yang *et al*., *Science*, 302, 266 (2003)

Conformational fluctuation within a protein

Yang *et al*. (*Science*, 302, 266, 2003) studied flavin-enzyme **Fre**: catalyzes the reduction of flavin, which contains a flavin adenine dinucleotide (**FAD**) and a tyrosine (**Tyr**).



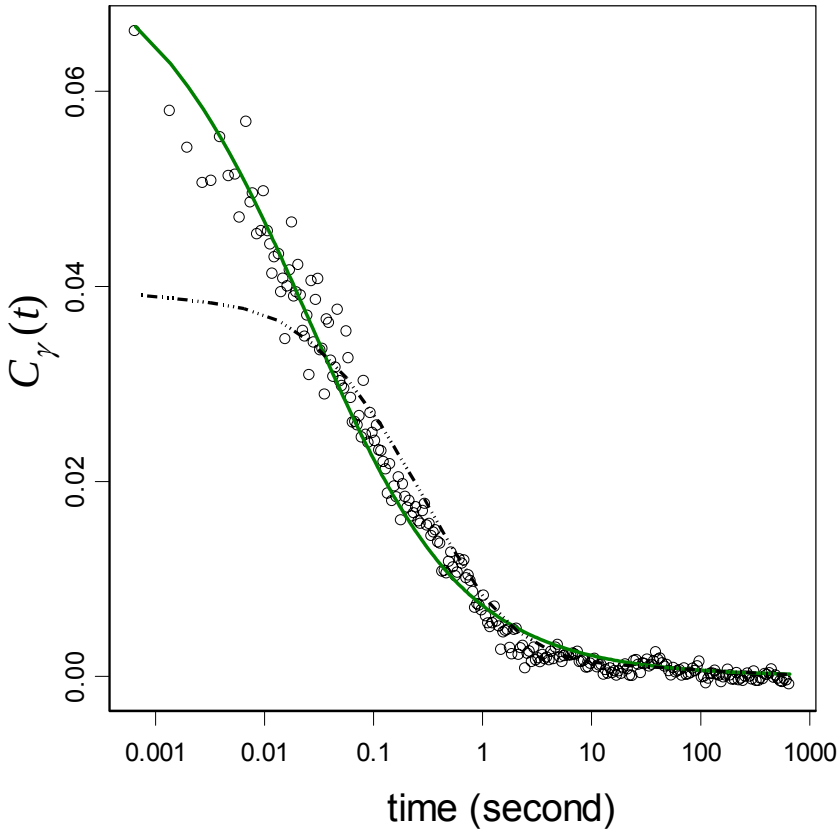
Fluorescence lifetime of FAD varies due to **electron transfer** from Tyr,

depending on the distance between ET donor and acceptor

$$\gamma^{-1}(t) = [k_0 e^{-\beta(X_{eq}+X_t)}]^{-1}$$

Autocorrelation function $\langle \Delta\gamma^{-1}(0)\,\Delta\gamma^{-1}(t)\rangle$

$$\Delta\gamma^{-1}(t) = \gamma^{-1}(t) - \langle\gamma^{-1}(t)\rangle$$



- Need tools beyond Langevin equation and BM.

- The model: Generalized Langevin equation with fractional Gaussian noise (GLE with fGn).

Generalized Langevin Equation with fGn

- Langevin equation $m\frac{dv_t}{dt} = -\zeta v_t + F(t)$
- Generalized Langevin equation $m\frac{dv_t}{dt} = -\zeta \int_{-\infty}^{t} v_u K(t-u)du + G_t,$
- **Fluctuation-dissipation theorem** links memory kernel $K(t)$ with fluctuating force $E\{G_t G_s\} = \zeta k_B T \cdot K(t-s)$

- Key question: How to introduce the noise structure?
- Understand the white noise: White noise is the derivative of the Wiener process $F_t = \frac{d}{dt} B_t$
- $B(t)$ is the unique process: (i) Gaussian, (ii) independent increment, (iii) stationary increment, (iv) self-similar

- Natural generalization: (i) Gaussian (ii) stationary increment (iii) self-similar.

- The **ONLY** candidate fBM $B^{(H)}(t)$, 0 < *H* < 1
  $E\{B_t^{(H)}\} = 0$, and covariance function
  $$E\{B_t^{(H)} B_s^{(H)}\} = (|t|^{2H} + |s|^{2H} - |t-s|^{2H})/2 \quad \text{for } t, s \geq 0.$$
  when *H* = 1/2, reduces to *B*(*t*).

- Fractional Gaussian noise $F^{(H)}(t) = \frac{dB_t^{(H)}}{dt}$:
  Gaussian & stationary.
- Memory kernel
  $$K_H(t) = E\{F^{(H)}(0)F^{(H)}(t)\}$$
  $$= H(2H-1)|t|^{2H-2}, \text{for } t \neq 0$$
- Spectral density
  $$\tilde{K}_H(\varpi) = \int_{-\infty}^{\infty} e^{it\varpi} K_H(t)dt = \Gamma(2H+1)\sin(H\pi)|\varpi|^{1-2H}$$

Toward subdiffusion

- Applying Fourier transform on

$$m\frac{dv_t}{dt} = -\zeta\int_{-\infty}^{t} v_u K_H(t-u)du + F_t^{(H)},$$

$v(t)$ Gaussian $E\{v(t)\} = 0$,

$$C(t) = E\{v(0)v(t)\} = \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{it\varpi}\tilde{C}(\varpi)d\varpi$$

$$\tilde{C}(\varpi) = \frac{k_B T\zeta\tilde{K}_H(\varpi)}{\left|\zeta\tilde{K}_H^{+}(\varpi) - i\varpi m\right|^2}$$

- For displacement $X(t) = \int_0^t v(s)ds$

$$E\{X(t)^2\} = \int_0^t\int_0^t E\{v(s)v(u)\}du\,ds$$

$$E\{X(t)^2\} \sim \frac{k_B T}{\zeta}\frac{2\sin(2H\pi)}{\pi H(2H-1)(2H-2)}t^{2-2H} \propto t^{2-2H}$$

- $H > 1/2$ leads to subdiffusion!

Harmonic potential

- For harmonic potential $U(x) = \frac{1}{2} m\omega^2 x^2$, model becomes

$$m\frac{d^2x(t)}{dt^2} = -\zeta \int_{-\infty}^{t} \frac{dx(u)}{du} K_H(t-u)du - \frac{dU(x)}{dx} + F^{(H)}(t)$$

- Overdamped limit:

$$m\omega^2 x(t) = -\zeta \int_{-\infty}^{t} \frac{dx(u)}{du} K_H(t-u)du + F^{(H)}(t)$$

- Solution: stationary Gaussian

$$\langle x(t) \rangle = 0$$

$$C_x(t) = \langle x(0)x(t) \rangle = \frac{k_B T}{m\omega^2} E_{2-2H}\left(-\frac{m\omega^2}{\zeta \Gamma(2H+1)} t^{2-2H}\right)$$

$E_a(z)$: Mittag-Leffler function $E_\alpha(z) = \sum_{k=1}^{\infty} z^k / \Gamma(\alpha k + 1)$

$$C_{xx}(t) = E\{X(0)X(t)\}$$

$$= \frac{k_B T}{m\omega^2} E_{2-2H}\left(-\frac{m\omega^2}{\zeta\Gamma(2H+1)} t^{2-2H}\right)$$

$$E_\alpha(z) = \sum_{k=0}^{\infty} z^k / \Gamma(\alpha k + 1)$$

- For all *H*,

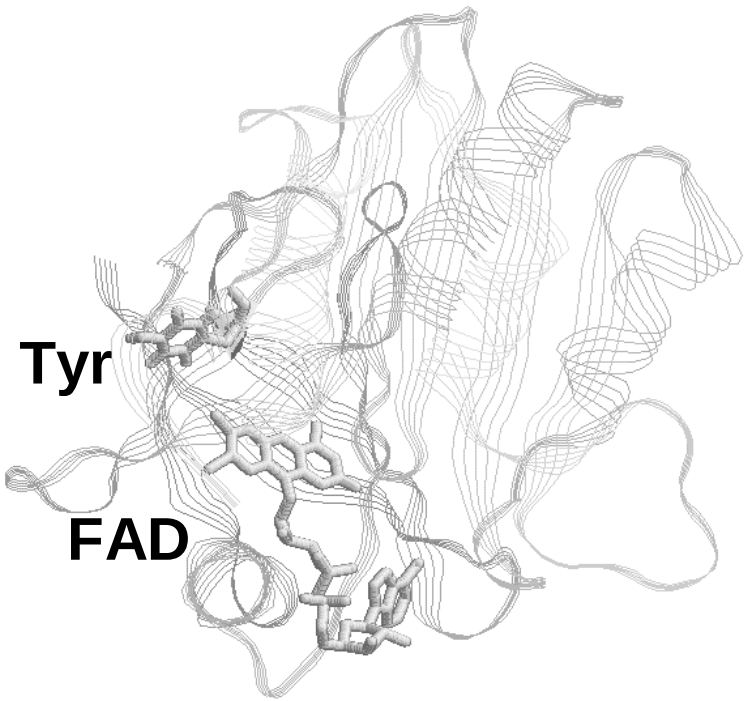
$$C_{xx}(0) = E\{X(0)X(0)\} = \frac{k_B T}{m\omega^2}$$

  the **thermal equilibrium** value.

- For *H* = 1/2, recovers the Brownian diffusion result

$$C_{xx}(t) = E\{X(0)X(t)\} = \frac{k_B T}{m\omega^2} e^{-2\frac{m\omega^2}{\zeta}t}$$

Back to experiment



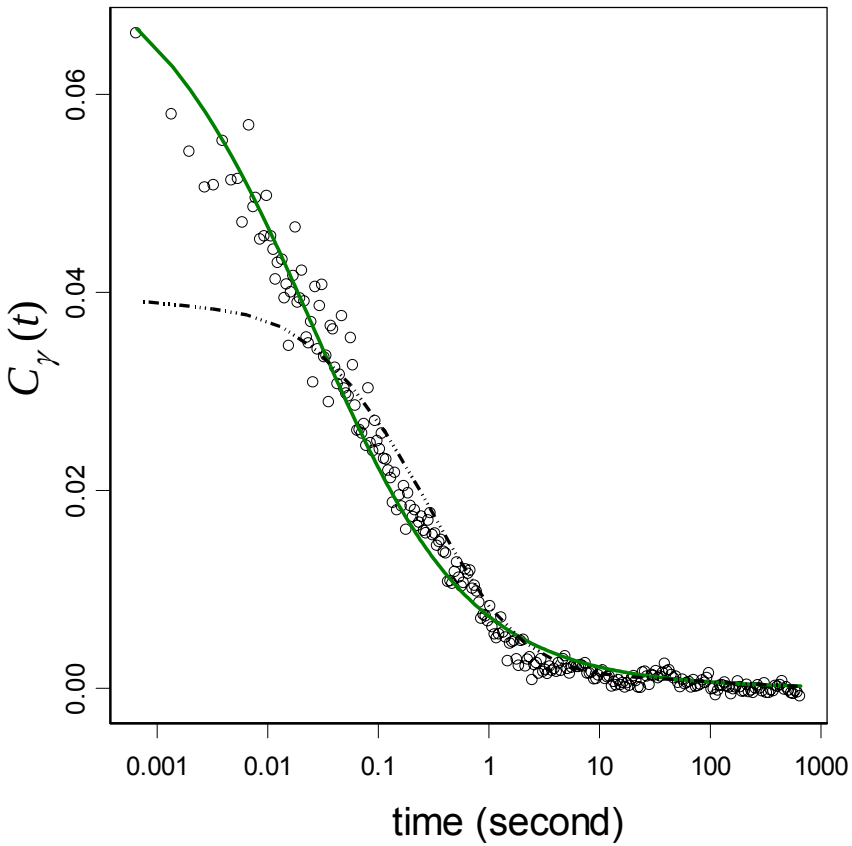
Fluorescence lifetime of FAD depends on the distance between ET donor and acceptor

$$\gamma^{-1}(t) = [k_0 e^{-\beta(X_{eq}+X_t)}]^{-1}$$

⇨ Model $X(t)$ by GLE with fGn under harmonic potential

easy calculation of lifetime autocorrelation

Fitting experimental autocorrelation

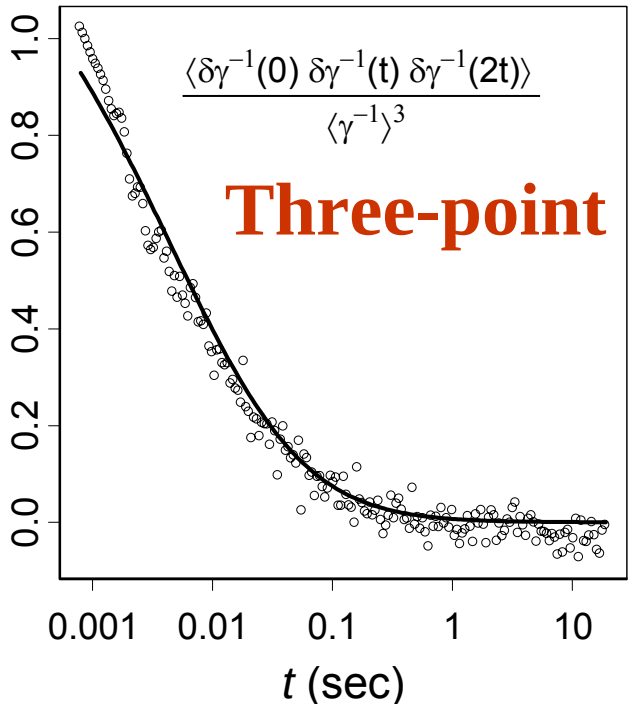


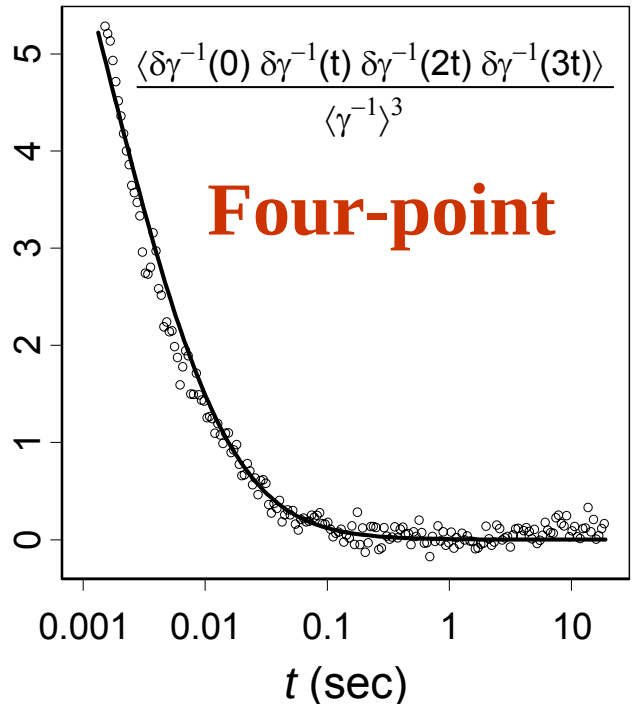
H	$\frac{\zeta}{m\omega^2}$	$\frac{k_B T}{m\omega^2}\beta^2$
0.74	0.40	0.81

Kou and Xie, *Phys. Rev. Lett.*, **93**, 180603 (2004).

Higher order autocorrelation functions

$$\delta\gamma^{-1}(t) := \gamma^{-1}(t) - \langle\gamma^{-1}(t)\rangle$$





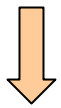
Same parameters:

H	$\frac{\zeta}{m\omega^2}$	$\frac{k_B T}{m\omega^2}\beta^2$
0.74	0.40	0.81

A prediction for three-point autocorrelation

- GLE with fGn predicts time symmetry

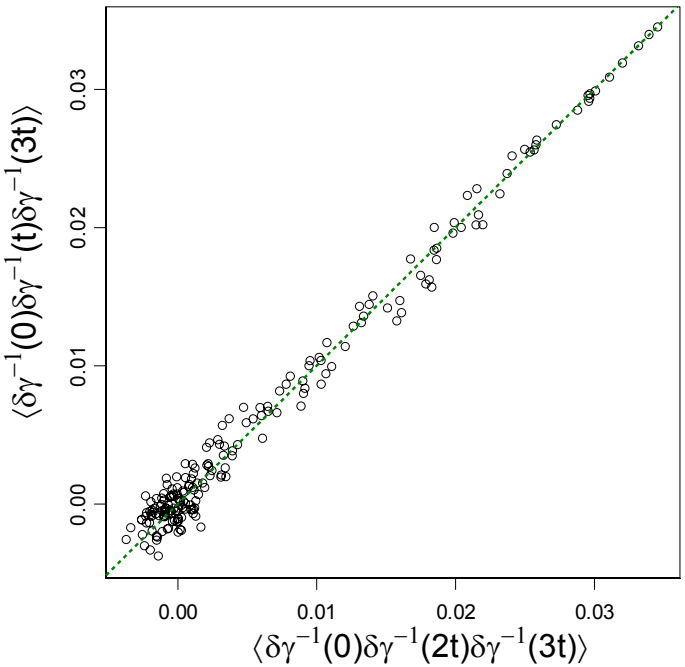
$$\langle \delta\gamma(0)^{-1}\, \delta\gamma(t_1)^{-1}\, \delta\gamma(t_1+t_2)^{-1} \rangle$$
$$= \langle \delta\gamma(0)^{-1}\, \delta\gamma(t_2)^{-1}\, \delta\gamma(t_1+t_2)^{-1} \rangle$$



- In particular

$$\langle \delta\gamma^{-1}(0)\, \delta\gamma^{-1}(t)\, \delta\gamma^{-1}(3t) \rangle$$
$$= \langle \delta\gamma^{-1}(0)\, \delta\gamma^{-1}(2t)\, \delta\gamma^{-1}(3t) \rangle \text{ for all } t$$

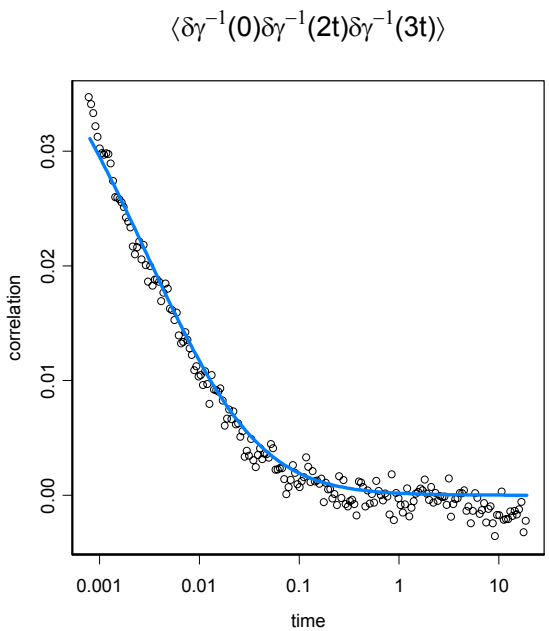
- Check with experiments:

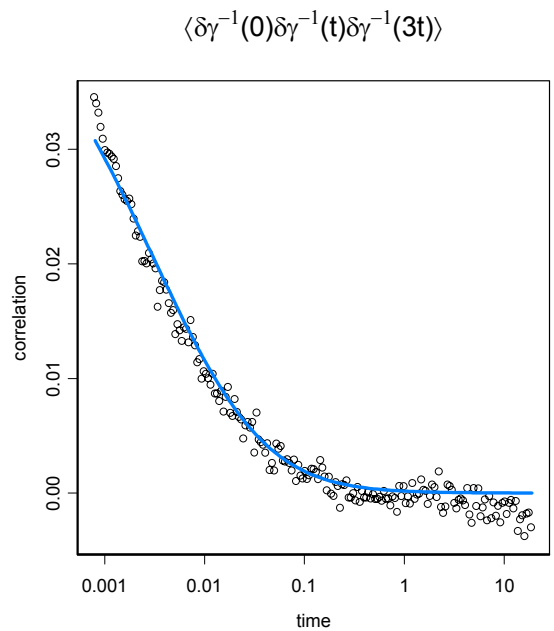


$$\langle \delta\gamma^{-1}(0)\,\delta\gamma^{-1}(t)\,\delta\gamma^{-1}(3t)\rangle$$

versus

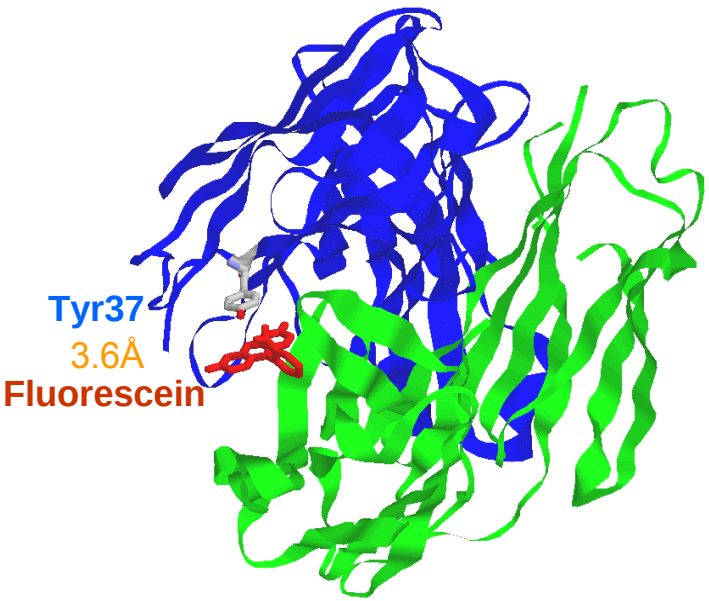
$$\langle \delta\gamma^{-1}(0)\,\delta\gamma^{-1}(2t)\,\delta\gamma^{-1}(3t)\rangle$$

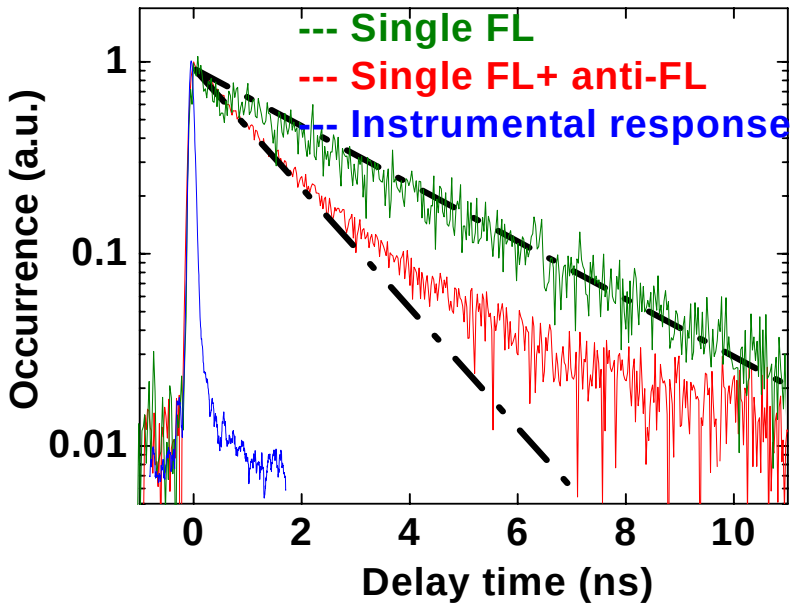




Another system

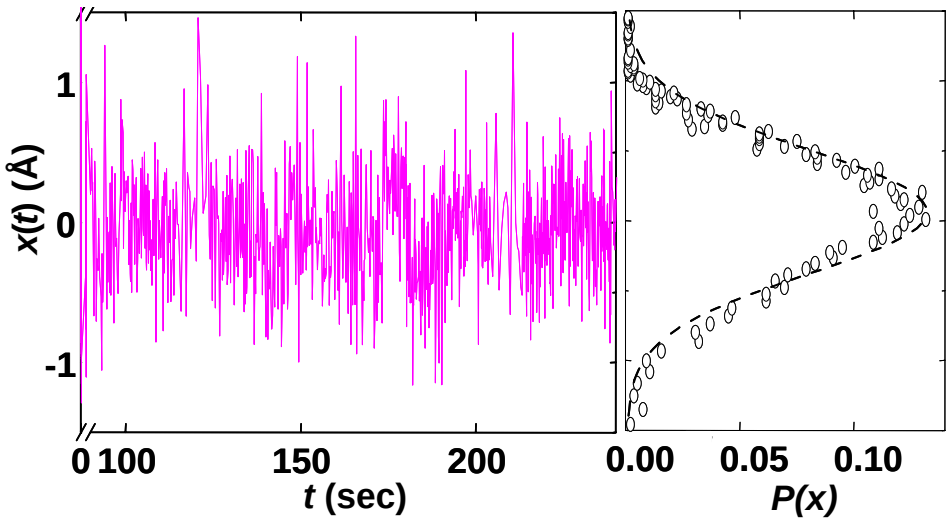
A protein complex formed by fluorescein (FL) and monoclonal anti-fluorescein (anti-FL)



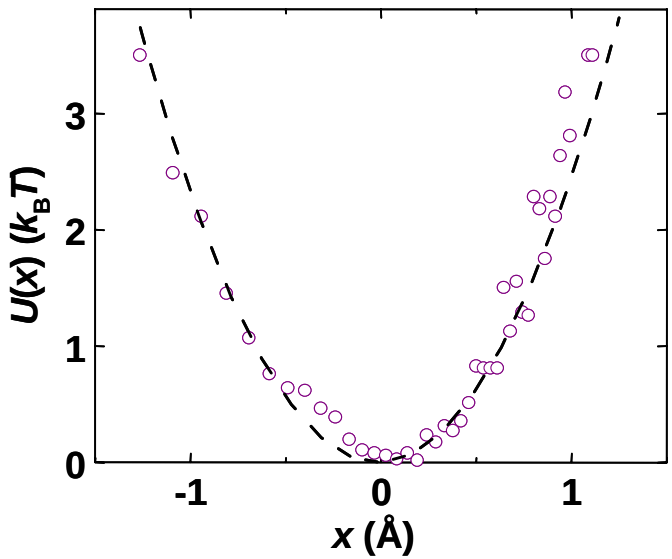


Min, *et al*. *Phys. Rev. Lett.* **94**, 198302 (2005).

Obtain distance fluctuation from $\gamma^{-1}(t) = [k_0 e^{-\beta(X_{eq}+X_t)}]^{-1}$



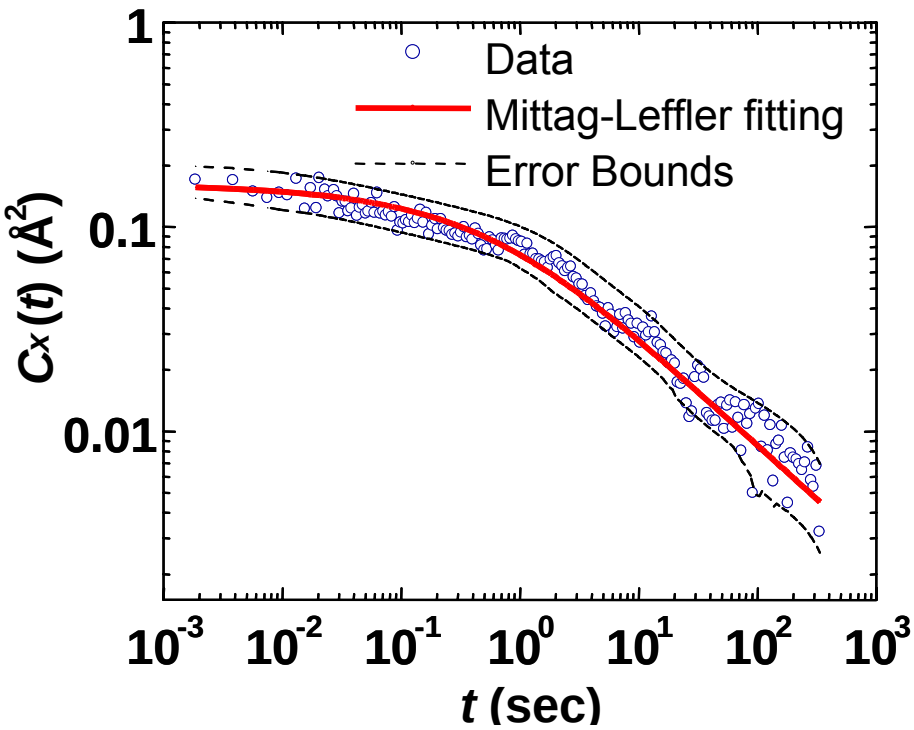
$$P(x) \propto \exp(-U(x)/k_B T)$$



Potential mean force

$$U(x) = -k_B T \ln P(x)$$

Autocorrelation of distance fluctuation



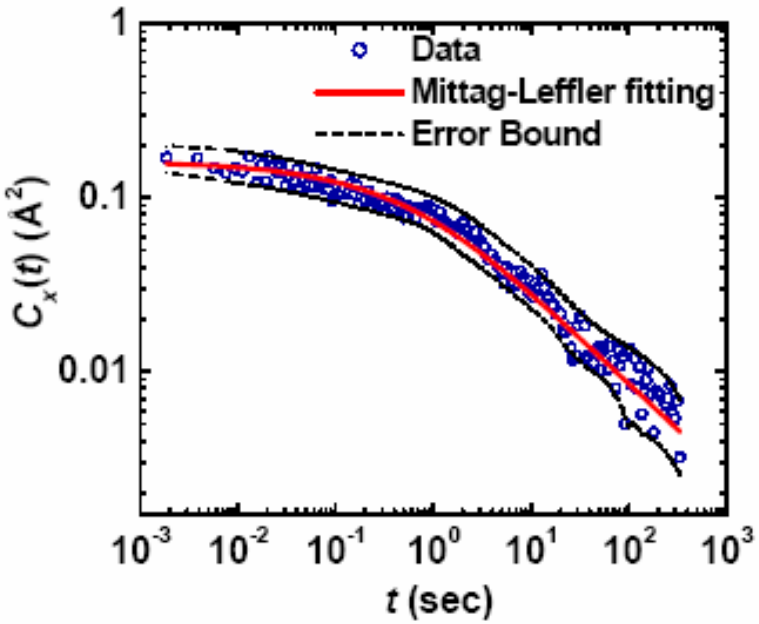
$$m\omega^2 x(t) = -\zeta \int_{-\infty}^{t} \frac{dx(u)}{du} K(t-u)du + F(t)$$

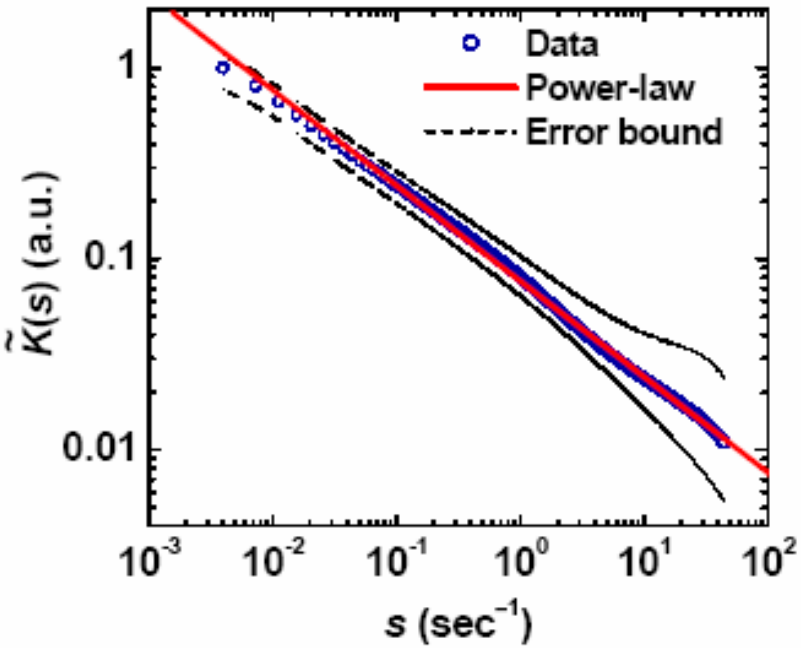
one-to-one correspondence

$$\tilde{K}(s) = \frac{m\omega^2}{\zeta} \frac{\tilde{C}_x(s)}{C_x(0) - s\tilde{C}_x(s)}$$

Memory kernel of conformational fluctuations

$$\tilde{K}(s) = \frac{m\omega^2}{\zeta} \frac{\tilde{C}_x(s)}{C_x(0) - s\tilde{C}_x(s)}$$





$$K(t) \propto t^{-\alpha-1} = t^{-0.51 \pm 0.07}$$

Summary

- Michaelis-Menten with dynamic disorder
  - Explains experimental data
  - Derive single-molecule Michaelis-Menten equation
  - Dynamic disorder might be masked in ensemble studies!
- Generalized Langevin Equation with fGn
  - Explains the observed conformational dynamics.
  - Experimentally obtain power law memory kernel

Summary (Continued)

- The connection: both highlight the conformation fluctuation:
  - An enzyme is a dynamic entity with conformational interconversion on a board range of time scales.
  - The interconverting conformers have different enzymatic reaction rate constants.

Acknowledgement

- Xie group at Harvard Chemistry & Chemical Biology
- Binny Cherayil
- Attila Szabo
- Hong Qian

- NSF grant DMS-02-04674
- NSF Career Award