

APPENDIX S1

I. The free energy distribution

The Hamiltonian or energy function for the interactions between ligands and receptors can be described by the collections of contact interactions between the atom pairs $E = -\sum_{ij} J_{ij} \sigma_{ij}$, where σ_{ij} is the contact variable between atoms with certain distance cutoff and the J_{ij} is the coupling strength for specific contact pair. Since there are many different types of atoms and also many different cutoff distances for the interactions, different J_{ij} have different values. This forms a distribution for the coupling J_{ij} . ~~Since the number of different J_{ij} couplings are large, the statistical distribution should have a Gaussian form~~ The coupling strengths between the atomic pairs J_{ij} is assumed to be Gaussianly

distributed: $f(J_{ij}) \sim \exp[-\frac{(J_{ij} - \bar{J})^2}{2\Delta J^2}]$ ~~from the large number theorem~~. Then

we can prove that this is equivalent to a random energy model with the interaction energy follows a Gaussian distribution:

$f(E) \sim \exp[-\frac{(E - \bar{E})^2}{2\Delta E^2}]$ [1]. ~~This~~ reflects the complexity of the underlying

interactions in contrast to the conventional models for simple systems where the coupling strengths are fixed and not distributed.

The resulting random energy model can be defined as follows:

(1). The system has M^N energy levels E_i where M is the number of the

configurations for unit while N is the number of units.

(2) These energy levels are distributed according to the probability distribution $P(E) = (N\pi J^2)^{-1/2} \exp(-E^2 / NJ^2)$. J is the coupling strength between contacts.

(3) The energy levels E_i are independent random variables.

From the random energy model, we can explore the statistical properties of associated free energy (reference 498 in the main text).

The system composes of M^N energy levels E_i . The corresponding partition function Z is then given by $Z(\{E_i\}) = \sum_{i=1}^{2^N} \exp\left\{-\frac{E_i}{T}\right\}$. Once the partition function Z is known, the free energy is known.

Notice that since the energy is distributed due to the statistical properties of the coupling strengths, we expect the corresponding partition function is distributed. Therefore, we expect the corresponding free energy is also distributed. We can calculate the average free energy by the formula: $F = -T \langle \ln Z \rangle = -T \int \prod_i [P(E_i) dE_i] \ln Z(\{E_i\})$

Furthermore, the free energy can be obtained using the formula: $F = -k_B T \ln Z$ (here, $k_B = 1$). Based on the previous studies [2], we can reach the form of the probability distribution $P(Z)$ of Z from the distribution of the energy.

After some algebra, the distribution of partition function reads:

$$P(Z)dZ = g\left(\frac{Z}{Z_0}\right)d\left(\frac{Z}{Z_0}\right) \text{ for } \frac{Z}{Z_0} \text{ finite;}$$

The distribution of partition function reads:

$$p(Z)dZ = 2^N \sqrt{\frac{T^2}{NJ^2\pi}} \exp\left(-\frac{T^2}{NJ^2} \ln Z^2\right) d \ln Z \text{ for } \frac{Z}{Z_0} \text{ exponentially large}$$

in N.

The moments $\langle Z^v \rangle$ of the partition function is:

$$\langle Z^v \rangle = \int \dots \int \prod_{i=1}^{2^N} P(E_i) dE_i \text{ and the function } g(x) \text{ is}$$

$$\int x^v g(x) dx = \begin{cases} \frac{\Gamma(1-vT_c/T)}{\Gamma(1-v)} (v < T/T_c) \\ \infty (v > T/T_c) \end{cases}.$$

When $v > T/T_c$, the formula

$$p(Z)dZ = 2^N \sqrt{\frac{T^2}{NJ^2\pi}} \exp\left(-\frac{T^2}{NJ^2} \ln Z^2\right) d \ln Z \text{ dominates the integral}$$

$$\int Z^v P(Z) dZ \text{ and the moments } \langle Z^v \rangle \sim 2^N \exp[Nv^2 J^2 / 4T^2] \text{ for } T > T_c \text{ and}$$

$v > (T/T_c)^2$ or $T < T_c$ and $v > T/T_c$. can be recovered. Therefore, by the

$$\text{formula } p(Z)dZ = 2^N \sqrt{\frac{T^2}{NJ^2\pi}} \exp\left(-\frac{T^2}{NJ^2} \ln Z^2\right) d \ln Z \text{ and } F = -T \ln Z (k_B$$

$= 1, k_b = 1$), we can easily obtain the gaussian distribution: $P(F) \sim$

$\exp\left(-\frac{F^2}{NJ^2}\right)$ as shown in equation 74 in the main text. In the paper, we use

the $f(F)$ to represent the free energy distribution. Notice that the $g(x)$

mentioned above has an asymptotic form

$$g(x) \sim x^{-T/T_c-1} \left(-\frac{1}{\Gamma(-T/T_c)}\right) \text{ for } x \rightarrow \infty. \text{ Based on the formula}$$

$P(Z)dZ = g\left(\frac{Z}{Z_0}\right)d\left(\frac{Z}{Z_0}\right)$, one can also recover the moments

$\langle Z^v \rangle = \int Z^v g\left(\frac{Z}{Z_0}\right)d\left(\frac{Z}{Z_0}\right)$ for $T < T_c$ and $v < T/T_c$. Therefore, we can obtain

the distribution $P(Z)$ of Z at the tail similar to $g(x)$ above. Since partition function Z is power law distributed asymptotically at the tail, the associated free energy is therefore exponentially distributed at the tail as shown in equation 85 in the main text. The exponential distribution of

free energy can also be obtained by the formula (3).

$f(f_a) = \rho \cdot \exp \rho(f_a - f_c) \cdot \theta(f_c - f_a)$ [3]. ρ is a function of the temperature, f_c

is a cut-off free energy needed at an intermediate stage. f_a are

independent random variables representing the free energy. The authors

had succeeded in proving the formula (3) by computing the inclusive distributions of probabilities $f^{(k)}$.

References:

1. J. Wang*, G. M. Verkhivker. (2003) Energy landscape theory, funnels, specificity and optimal criterion of biomolecular binding. Phys. Rev. Lett. 90, 188101.

2. Gardner E, Derrida B (1989) The probability distribution of the partition function of the random energy model. Journal of Physics A: Mathematical and General 22: 1975.

3. Mezard M, Parisi G, Virasoro M (1985) Random free energies in spin glasses. Journal de Physique Lettres 46: 217–222.

II. The kinetics of binding

Bryngelson and P. G. Wolynes had studied the folding time of protein since the 1980's, based on a simple folding model, which has numerous metastable states [1]. The continuous time random walk (CTRW) approach was applied to determine the time scale of folding. By a simple generalization of transition-state theory and some approximations, a simple expression of the folding time was gained. This study has become a fundamental work for studying the folding time. Furthermore, the protein folding kinetics had been studied according to the statistical energy landscape theory, with a defined expression for the mean first-passage time (MFPT) to characterize the folding kinetics.

The calculations regarding the MFPT in this study are mainly based on the analytical results in the previous studies [1-2]. The procedure to derive MFPT or time scale is rather long. We shall assume that the reader has at least a basic understanding of the ideas and results of these earlier works. Herein, we recall some basic results provided in the study [1]. First, the folding problem can be expressed by a continuous time random walk, and the waiting time distributions are correlated with the distributions of escape rates from each site. Then, the continuous time random walk can be represented by the generalized master equation. Next, the results show that the continuum limit of the generalized master

equation is a generalized Fokker-Planck equation. After that, the mean first passage time for the generalized Fokker-Planck equation is calculated, and the mean first passage time is correlated with the first inverse moment of the distribution of the rate of escape from the sites. Finally, a simple approximate expression for the mean first passage time is derived.

Due to high similarity between folding and binding, the binding kinetics can be realized in the same way. Then the kinetic binding process is also approximated by the Metropolis dynamics

$$R(E_1 \rightarrow E_2) = \begin{cases} R_0 \exp\left[-\frac{(E_2 - E_1)}{T}\right] & \text{for } E_2 > E_1 \\ R_0 & \text{for } E_2 < E_1 \end{cases} \quad (1)$$

Where $R(E_1 \rightarrow E_2)$ is the transition rate for a receptor-ligand system from binding state 1 to 2, with total energies of system E_1 and E_2 , respectively. R_0 is an overall constant characterizing the inverse time scale for the transition process between binding states (R_0 is generally of on the order of inverse nanoseconds). Thus, the transition rate from one binding state to a neighboring binding state can be determined by the energy gap between these two states. This problem can be addressed through the aforementioned CTRW. One can further reduce the multidimensional random walk problem to the simple one-dimensional CTRW, and therefore a generalized master equation. In summary, one can first stratify the binding energy landscape in the order parameter ρ space, then a

relevant energy distribution function $P(E, \underline{\rho})$ is gained. The corresponding transition rate distribution function $P(R, \underline{\rho})$ can be calculated through the Metropolis dynamics given above. By specifying the jumping rate R for a binding complex at a binding state with the specific order parameter $\underline{\rho}$ to its neighboring states adjacent in order parameter space, one can obtain the associated waiting-time distribution $\psi(\tau, \underline{\rho})$ for the complex to reside in a binding state for time (i.e. residence time) before it leaves. Finally, a CTRW can be solved with the waiting-time distribution for the binding complex system and the associated jumping probabilities between successive $\underline{\rho}$'s. These probabilities are approximated to be time-independent, and the jumping process is equivalent to the quasi equilibrium assumption. Under this assumption, these probabilities can be gained by utilizing the asymptotic distribution

$$\lim_{\tau \rightarrow \infty} G(\underline{\rho}, \tau) \propto e^{-\beta F(\underline{\rho})} \quad (2)$$

Where $G(\underline{\rho}, \tau)$ represents the function of probability distribution for the binding complex at time τ , and $\beta=1/ k_B T$ (here, $k_B=1$). Then one can obtain a generalized kinetic master equation using the CTRW approximation, it can be expressed in the Laplace-transformed space as

$$sG(\underline{\rho}, s) - n_i(\underline{\rho}) = \hat{\kappa}(\underline{\rho}, s)G(\underline{\rho}, s) \quad (3)$$

Where $n_i(\underline{\rho})$ represents the initial distribution of $G(\underline{\rho}, \tau)$. While $\hat{\kappa}(\underline{\rho}, s)$ represents a linear operator related to the jumping probabilities and the waiting time distribution. In the local connectivity

case, the generalized master equation given above is reduced to a generalized Fokker-Planck equation in the Laplace-transformed space

$$s\mathcal{G}(\rho, s) - n_i(\rho) = \frac{\partial}{\partial \rho} \left\{ D(\rho, s) \left[\mathcal{G}(\rho, s) \frac{\partial}{\partial \rho} U(\rho, s) + \frac{\partial}{\partial \rho} \mathcal{G}(\rho, s) \right] \right\} \quad (4)$$

Where
$$U(\rho, s) \equiv \frac{F(\rho)}{T} + \log \frac{D(\rho, s)}{D(\rho, 0)} \quad (5)$$

s represents the Laplace transform variable over time τ , while $D(\rho, s)$ represents the frequency-dependent diffusion parameter. $F(\rho)$ means the average of free energy from the random energy model.

$\mathcal{G}(\rho, s)$ corresponds to the Laplace transform of $G(\rho, \tau)$, the latter represents the probability density function. That is to say, $G(\rho, \tau) d\rho$ is the probability for a binding complex to stay at $d\rho$ space at time τ .

$n_i(\rho)$ represents the initial condition for $G(\rho, \tau)$. One can set the boundary conditions for the equation as a reflecting one at $\rho=0$,

$$\left[\mathcal{G}(\rho, s) \frac{\partial}{\partial \rho} U(\rho, s) + \frac{\partial}{\partial \rho} \mathcal{G}(\rho, s) \right] \Big|_{\rho=0} = 0 \quad \text{where the binding complex}$$

corresponds to the non-native binding states, while an absorbing one at

$$\rho = \rho_f, \quad \mathcal{G}(\rho_f, s) = 0, \quad \text{where the binding complex is in the native binding}$$

state. Here, to facilitate the calculation of the first passage time distribution, an absorbing boundary condition at $\rho = \rho_f$ is given.

Then we denote the distribution of first passage time by $F_p(t)$ and define \sum_t to be the probability that a molecule has not made a first passage by time t. The distribution $F_p(t)$ and \sum_t are related by:

$$F_p(t) = -\frac{d}{dt} \sum (6)$$

The mean first passage time is expressed as:

$$\bar{t} \equiv \int_0^{\infty} dt t F_p(t) (7)$$

From eq.6 and eq.7 above, we can obtain a useful expression for the mean first passage time:

$$\bar{t} = \int_0^{\infty} dt \sum t (8)$$

Suppose that we make the “site” ρ_f a perfect absorber, so a molecule that makes a first passage simply leaves the system. The probability of a molecule being in the system at time t is given by $\sum t$, so

$$\sum t = \int_0^{\rho_f} d\rho G(\rho, t) (9)$$

And combining eq. 8 and eq.9, we obtain:

$$\bar{t} = \int_0^{\rho_f} d\rho G(\rho, s=0) (10)$$

In addition, the eq.4 can be also written as:

$$sG(\rho, s) - n_i(\rho) = -\frac{\partial}{\partial \rho} j(\rho, s) (11)$$

Where $j(\rho, s) = -D(\rho, s) \exp[-U(\rho, s)] \frac{\partial}{\partial \rho} \{G(\rho, s) \exp[U(\rho, s)]\}$ (12)

The boundary conditions are a reflecting boundary condition at $\rho=0$, $j(0, s)=0$, and an absorbing boundary condition at ρ_f , $G(\rho_f, s)=0$.

Integrating eq. 11 from 0 to ρ and using the reflecting boundary condition gives

$$j(\rho, s) = -\int_0^{\rho_f} d\rho' [sG(\rho', s) - n_i(\rho')] (13)$$

And substituting the expression for $j(\rho, s)$ above into eq. 12 yields

$$\frac{\partial}{\partial \rho} \{G(\rho, s) \exp[U(\rho, s)]\} = \frac{\int_0^\rho d\rho'' [sG(\rho'', s) - n_i(\rho'')] }{D(\rho, s) \exp[-U(\rho, s)]} \quad (14)$$

Next, integrating eq. 14 from ρ to ρ_f and using the absorbing boundary condition obtains

$$G(\rho, s) = -\int_0^{\rho_f} d\rho' \int_0^{\rho'} d\rho'' [sG(\rho'', s) - n_i(\rho'')] \frac{\exp[U(\rho', s) - U(\rho, s)]}{D(\rho', s)} \quad (15)$$

Here, eventually all of the molecules are assumed to make a first passage and therefore be absorbed, so from the relation $\lim_{s \rightarrow 0} sf(s) = \lim_{t \rightarrow \infty} f(t)$.

Then

$$\lim_{s \rightarrow 0} sG(\rho, s) = \lim_{t \rightarrow \infty} G(\rho, t) = 0 \quad (16)$$

Taking the limit of eq. 15 as s goes to zero and using eq. 19 yields

$$G(\rho, 0) = \int_0^{\rho_f} d\rho' \int_0^{\rho'} d\rho'' n_i(\rho'') \frac{\exp[U(\rho', 0) - U(\rho, 0)]}{D(\rho', 0)} \quad (17)$$

Which is the promised expression for $G(\rho, 0)$. Finally, we substitute eq. 17 into eq. 10 and use the property $U(\rho, 0) = \beta F(\rho)$ to obtain the mean first passage time expression:

$$\bar{t} = \int_0^{\rho_f} d\rho \int_\rho^{\rho_f} d\rho' \int_0^{\rho'} d\rho'' \frac{n_i(\rho'') \exp\{\beta[F(\rho') - F(\rho)]\}}{D(\rho', 0)} \quad (18)$$

Equation 18 simplifies when $n_i(\rho)$ is peaked about some ρ_i , so $n_i(\rho) \approx \delta(\rho - \rho_i)$, then

$$\bar{t} = \int_{\rho_i}^{\rho_f} d\rho \int_0^\rho d\rho' \frac{\exp[\beta F(\rho) - \beta F(\rho')]}{D(\rho, 0)} \quad (19)$$

The Equations 18 and 19 give us the final expressions for the folding time.

Here, we can also rewrite the Eq.4 in the order parameter Q space. The

binding kinetic process follows a diffusion equation in order parameter Q

[1-3] (in the Laplace space s):

$$s\hat{P}(Q,s) - P_i(Q) = \frac{\partial}{\partial Q} \left[D(Q,s) \frac{\partial \hat{P}(Q,s)}{\partial Q} + \hat{P}(Q,s) \frac{(\partial U(Q,s)/k_B T)}{\partial Q} \right] \quad (20)$$

Where D(Q,s) represents the frequency dependent diffusion coefficient.

While, $U(Q,s) = \beta F(Q) + \log(D(Q,s)/D(Q,s=0))$ represents the 'effective free energy' (note that, F(Q) means the free energy of the binding complex system at Q). β is $1/k_B T$, $P_i(Q)$ represents the initial condition for P(Q;

$t=0$), and $\hat{P}(Q,s)$ is the Laplace transform of P(Q,t):

$$\hat{P}(Q,s) = \int_0^{\infty} \exp[-st] P(Q,t) dt, \text{ which is the probability of the binding}$$

complex at Q and time t. Here, the diffusion coefficient D(Q, s) relies on

frequency s and order parameter Q. It can be written as:

$$D(Q,s) \left\langle \frac{R}{R+s} \right\rangle_R(Q) / \left\langle \frac{1}{R+s} \right\rangle_R(Q) \quad (21)$$

where s represents the Laplace transform variable over time t. The

average $\langle \rangle_R$ is taken over f(R,Q), which is the probability distribution

function derived from the transition rate R from a binding state with the specific order parameter Q to its neighboring binding states [1-3].

As we know, the binding dynamics depends on thermodynamics driving force and the kinetic diffusion. The equation of the latter can be integrated to give the mean first passage time :

$$\bar{\tau} = \int_{Q_b}^{Q_b} dQ \int_0^Q dQ' D^{-1}(Q) \exp[1/k_B T (F(Q) - F(Q'))] \quad (22)$$

$Q_b \sim 1$ means that the native binding state is reached. One can set the

boundary conditions for the equation as a reflecting one at $Q \sim 0$,

$$\underline{[\Psi(Q, s) \frac{\partial}{\partial Q} U(Q, s) + \frac{\partial}{\partial Q} \Psi(Q, s)]|_{Q=0} = 0}$$

where the binding complex

corresponds to the non-native binding states while an absorbing one at $Q = Q_b$, $\Psi(Q_b, s) = 0$

where the binding complex is in the native binding

states. The first passage time to reach Q_b (that is, the time for the random

walker visiting order parameter Q_b for the first time) will be used as a

representative and typical time scale for binding.

For the binding complex, the RMSD (Root Mean Square Deviation)

can be also viewed as order parameter. Thus, the RMSD similar to the Q

depicts the whole binding process towards the global native binding state

along the binding energy landscape. Similarly, we can start with a general

kinetic master equation, by assuming the local connectivity, then derive a

diffusion equation [1-7]:

$$\underline{\frac{\partial}{\partial t} P(\text{RMSD}, t) = \frac{\partial}{\partial \text{RMSD}} [D(\text{RMSD}) \frac{\partial P(\text{RMSD}, t)}{\partial \text{RMSD}} + P \frac{\partial (F(\text{RMSD}) / \kappa_B T)}{\partial \text{RMSD}}]} \quad (23)$$

where $P(\text{RMSD}, t)$ represents the probability of the binding complex with

specific RMSD at time t , $D(\text{RMSD})$ represents the diffusion coefficient

and $F(\text{RMSD})$ represents the corresponding free energy of the complex

system at RMSD. Actually, the diffusion coefficient can be essentially

viewed as the average of time leaving the local minimum site. The

problem about the kinetics of the binding can be converted into one

dimensional diffusion for RMSD. The diffusion equation can be also

integrated to give the MFPT:

$$\bar{\tau} = \int_{RMSD_i}^{RMSD_f} dRMSD \int_{RMSD_i}^{RMSD} dRMSD' \frac{\exp\left[\frac{F(RMSD) - F(RMSD')}{\kappa_B T}\right]}{D(RMSD)} \quad (24)$$

RMSD_i ~ 1 A means that where the native binding state is reached. One can set the boundary conditions for the equation as a reflecting one at

$$\underline{RMSD_i}, \left[P(RMSD, t) \frac{\partial}{\partial RMSD} F(RMSD) + \frac{\partial}{\partial RMSD} P(RMSD, t) \right] \Big|_{RMSD=RMSD_i} = 0$$

where the binding system is in native state, while an absorbing one at RMSD=RMSD_f, P(RMSD_f,t)=0, where the binding system is in the non-native states. The absorbing boundary condition similar to the order parameter Q is chosen.

Reference:

1. J. D. Bryngelson and P. G. Wolynes (1989) Intermediates and Barrier Crossing in a Random Energy Model (with Applications to Protein Folding) J. Phys. Chem. 93: 6902-6915.
2. Lee CL, Stell G, Wang J (2003) First-passage time distribution and non-markovian diffusion dynamics of protein folding. The Journal of chemical physics 118: 959-968.
3. Wang J, Huang W, Lu H, Wang E (2004) Downhill kinetics of biomolecular interface binding: globally connected scenario. Biophys. J 87: 2187-2194.
4. Wang J (2006) Diffusion and single molecule dynamics on

biomolecular interface binding energy landscape. Chem.Phys.Lett. 418: 544-548.

5. Lee CL, Lin CT, Stell G, Wang J (2003) Diffusion dynamics, moments, and distribution of first-passage time on the protein-folding energy landscape, with applications to single molecules. Phys. Rev. E. 67: 41905-41909.

6. Wang J (2003) Statistics, pathways and dynamics of single molecule protein folding. J ChemPhys. 118: 952-958.

7. Wang J, Xu L, Xue K, Wang E (2008) Exploring the origin of power law distribution in single-molecule conformation dynamics: Energy landscape perspectives. Chem. Phys. Lett. 463: 405-409.

III. The distributions of equilibrium constant and kinetics

According to the probability theory and statistics, a log-normal distribution is a continuous probability distribution of a random variable whose logarithm is normally distributed. Thus, if the random variable X is log-normally distributed, then the distribution of Y ($Y = \log(X)$) is a normal distribution. In other words, if Y is normally distributed, then X ($X = \exp(Y)$) has a log-normal distribution. Therefore, one can directly derive the distribution of equilibrium constant K ($\log K = (F_n - F_{un})/RT$) according to the free energy distribution. Since the logarithm of the equilibrium constant is directly related to the free energy difference

between the native and non-native states (the non-native states are normal distributed while the native state is relatively narrowly distributed). Furthermore, the aforementioned distribution function of free energy has been determined as Gaussian in this study, one can easily obtain the distribution function of the equilibrium constant K as log-normal. Meanwhile, as mentioned above, the time scale τ can be expressed as $\log(\tau / \tau_0) = F^\# - F_{un} / T$, we can also obtain the corresponding distribution of binding kinetics based on the distribution of free energy.

As we discussed previously in the section II, we have obtained a simple approximate expression for MPFT. Next, we will give the detailed procedure to get the FPT (time scale or kinetics) distribution. First, one has the following relation for the FPT distribution function $P_{FPT}(\tau)$:

$$\underline{P_{FPT}(\tau) = \frac{d}{d\tau}(1 - \sum \tau) = -\frac{d\sum \tau}{d\tau} \quad (1)}$$

Where

$$\underline{\sum \tau \equiv \int_0^{\rho_f} d\rho G(\rho, \tau) \quad (2)}$$

The moments of the FPT distribution function are calculated from the following relation:

$$\begin{aligned} \underline{\langle \tau^n \rangle} &\equiv \int_0^\infty d\tau \tau^n P_{FPT}(\tau) = -\int_0^\infty d\tau \tau^n \frac{d\sum \tau}{d\tau} = n \int_0^\infty d\tau \tau^{n-1} \sum \tau \\ &= n \int_0^{\rho_f} d\rho \int_0^\infty d\tau \tau^{n-1} G(\rho, \tau) \\ &= \left[n(-1)^{n-1} \int_0^{\rho_f} d\rho \left(\frac{\partial}{\partial s} \right)^{n-1} \mathcal{G}(\rho, s) \right] \Big|_{s=0} \quad (3) \end{aligned}$$

If we make a series expansion of $\mathcal{G}(\rho, s)$ and $1/D(\rho, s)$

$$\underline{\mathcal{G}(\rho, s) = \mathcal{G}_0(\rho) + s\mathcal{G}_1(\rho) + s^2\mathcal{G}_2(\rho) + \dots \quad (4)}$$

and

$$\underline{1/D(\rho, s) = a_0(\rho) + sa_1(\rho) + s^2a_2(\rho) + \dots \quad (5)}$$

Then we can obtain

$$\underline{\langle \tau^n \rangle = n!(-1)^{n-1} \int_0^{\rho_f} d\rho \mathcal{G}_{n-1}(\rho) \quad (6)}$$

and the generalized Fokker-Planck equation

$$\underline{s\mathcal{G}(\rho, s) - n_i(\rho) = \frac{\partial}{\partial \rho} \left\{ D(\rho, s) \left[\mathcal{G}(\rho, s) \frac{\partial}{\partial \rho} U(\rho, s) + \frac{\partial}{\partial \rho} \mathcal{G}(\rho, s) \right] \right\} \quad \text{mentioned}}$$

above in the section II (Eq.4) will become

$$\underline{\mathcal{G}_0(\rho) = \int_0^{\rho_f} d\rho' \int_0^{\rho''} d\rho'' n_0(\rho'') a_0(\rho) K(\rho, \rho') \quad (7)}$$

and for $n \geq 0$

$$\underline{\mathcal{G}_{n+1}(\rho) = - \int_0^{\rho_f} d\rho' \int_0^{\rho''} d\rho'' \left[\sum_{j=0}^n \mathcal{G}_{n-j}(\rho'') a_j(\rho) - n_0(\rho'') a_{n+1}(\rho) \right] \left[\sum_{j=0}^n \mathcal{G}_{n-j}(\rho'') a_j(\rho) - n_0(\rho'') a_{n+1}(\rho) \right] K(\rho, \rho')}$$

(8)

by matching each coefficient of s^n in Eq.4 and Eq.5. Therefore, one can

calculate $\mathcal{G}_n(\rho)$ recursively. Meanwhile, one can also solve the intergral

Eq. 15 in the section II directly for $\mathcal{G}(\rho, s)$, and by the observation that

$$\underline{\tilde{P}_{FPT}^0(s) = 1 - s \tilde{\Sigma}(s) \quad (9)}$$

Where $\tilde{P}_{FPT}^0(s)$ and $\tilde{\Sigma}(s)$ are Laplace transforms of $P_{FPT}(\tau)$ and $\Sigma \tau$,

respectively, we can investigate $P_{FPT}(\tau)$ by studying the behavior of

$\tilde{P}_{FPT}^0(s)$. To solve Eq. 15 in the section II numerically, one first replaces

the integrations by discrete summations. Because the equation is linear in

$\mathcal{G}(\rho, s)$, one can solve $\mathcal{G}(\rho, s)$ in the discrete ρ space by a matrix

inversion technique. In the discrete version this equation (Eq. 15 in the section II) becomes[1]

$$\underline{\hat{G}}_i = -\sum_{j,k} (\Delta\rho)^2 \underline{\hat{K}}_{ii}^{-1} \underline{\hat{D}}_{ii}^{-1} \underline{\hat{I}}_{2ij} \underline{\hat{K}}_{jj} \underline{\hat{I}}_{1jk} (s\underline{\hat{G}}_k - \underline{\hat{n}}_{0k}) \quad (10)$$

Where the integral operators become matrices. $\underline{\hat{I}}_{1ij} = 1$ if $i > j$, and $\underline{\hat{I}}_{1ij} = 0$ if $i < j$. $\underline{\hat{I}}_{2ij} = 1$ if $i > j$, and $\underline{\hat{I}}_{2ij} = 0$ if $i < j$. i, j , and k are discrete labels of the order parameter. $\underline{\hat{K}}$ and $\underline{\hat{D}}$ are diagonal matrices with non-zero elements $K(\rho_i)$ and $D(\rho_i, s)$, respectively. $\underline{\hat{G}}$ and $\underline{\hat{n}}_0$ are vectors of elements $\mathcal{G}(\rho_i, s)$ and $n_0(\rho_i)$. With these notations one can easily get $\underline{\hat{G}} = (\underline{\hat{D}} \underline{\hat{K}} + s\underline{\hat{I}}_2 \underline{\hat{K}} \underline{\hat{I}}_1)^{-1} \underline{\hat{I}}_2 \underline{\hat{K}} \underline{\hat{I}}_1 \underline{\hat{n}}_0$. In our calculations we make the matrix inversion with a Gaussian elimination method with scaled-column pivoting, and the results do not change wildly with the number of grids we choose in the discrete space, so we conclude this is a stable procedure for solving $\mathcal{G}(\rho, s)$ and therefore $\underline{\hat{P}}_{FPT}(s)$.

Except for the mathematical transition and the solving procedure given above, some previous studies have characterized the FPT and its distribution based on analytical theory and simulation studies [1-5]. Bryngelson and P. G. Wolynes had discussed some properties of distribution of folding kinetics [2]. In addition, Chi-Lun Lee and Jin Wang analytically studied the fluctuations and high order moments of the FPT to infer the corresponding distribution by exploring the kinetics of protein folding. It was found that for temperatures well above T_0 , the conformation dynamic process is self-averaging and its FPT distribution

obeys a Poisson distribution. But when the temperature is lower, the fluctuations start to diverge. This means that the actual conformation dynamic process may happen on multiple time scales, and the non-self-averaging behavior emerges. In this case, the full distribution of the FPT is required in order to characterize the system. From analytical analysis the distribution of FPT turns out to be close to a Levy distribution, which has a power-law tail for long time [1].

The authors had calculated the moments $\langle \tau^n \rangle$ of the FPT distribution function [1]. Using $\langle \tau^2 \rangle$ as an example, the authors obtained a V shape curve for each setting of $\delta\varepsilon/\Delta\varepsilon$ with the minimum of the curve having a temperature close to T_0 . Similar to the behavior of $\langle \tau \rangle$, at low temperature $\langle \tau^2 \rangle$ is only dependent on $\Delta\varepsilon/T$, and in the high-temperature regime it is mainly dependent on $\delta\varepsilon/T$. The same results are also obtained for higher moments. Furthermore, by exploring the behavior of the reduced second moment, $\langle \tau^2 \rangle / \langle \tau \rangle^2$, the authors find that the reduced second moment starts to diverge at temperature around T_0 , where MFPT is at its minimum. The degree of divergence increases rapidly as temperature drops below T_0 . This indicates that the average is not a good representative of the system and a long tail in the FPT distribution is developed. The intermittency where rare events make great contribution occurs. The divergence of the second moment also shows that the dynamics is exhibiting nonself-averaging behavior.

From the study of higher moments, the authors find at high temperature the relationship $\langle \tau^n \rangle = n! \langle \tau \rangle^n$. Therefore, the FPT distribution function is Poissonian in the high-temperature regime. But when $T < T_0$, it is hard to get more information from the moments because of their diverging behavior. On the other hand, the folding dynamics can be also studied by solving the linear integral Eq. 15 in the section II directly by making the inversion of the linear operator. The authors have also investigated the behavior of the FPT distribution function in the Laplace-transformed space. The results show that for $T < T_0$ $\beta_{FPT}^0(s)$ decays slowly over decades, which suggests that the usual numerical Laplace inversion techniques cannot be applied. And, the authors further find that there is approximately a linear relation over several orders of magnitude in s (see Fig. 6 in reference 1). This indicates that for $T < T_0$ $\beta_{FPT}^0(s)$ can be approximated by a stretched exponential $\beta_{FPT}^0(s) \approx e^{-cs^\alpha}$ which is the Laplace transform of the Le'vy distribution in the time space. Then by the transform, the authors derive $P_{FPT}(\tau) \approx -\frac{1}{\pi} \sum_{n=1}^{\infty} \frac{(-c)^n \Gamma(\alpha n + 1)}{\tau^{\alpha n + 1} (n + 1)} \sin(\pi \alpha n)$ α lies between 0 and 1. From the asymptotic property of the Le'vy distribution function the authors learn that $P_{FPT}(\tau) \sim \tau^{-(1+\alpha)}$ for large τ . And the calculation results show that α decreases when the temperature decreases.

Furthermore, Zhou et.al. had performed more than 22000 folding kinetic simulations to study conformational dynamics of the second β -hairpin fragment of protein G. The authors found that the MFPT has a

U(or V)-shaped dependence on the temperature and becomes smaller as the energy bias gap increases. Above a kinetic transition temperature T_{opt}^* , the MFPT is well behaved; and the distribution of FPT tends to be log-normal, as the temperature decreases, the distribution of FPT starts to become broader and has a power-law-like tail[3]. Wang et.al. also analytically explored the origin of power law distribution observed in single molecule conformational dynamics experiments. By establishing a kinetic master equation approach to statistically study the microscopic state dynamics, the results show that the underlying landscape with exponentially distributed density of states leads to power law distribution of kinetics ($P_{FPT}(\tau) \sim \tau^{-1-T/T_c}$). The exponential density of states emerges when the system becomes glassy and landscape becomes rough with significant trapping under low temperatures. For the theoretical results the power law decay coefficient is monotonically dependent on temperature which can be tested from singular molecular experiments. This may bridge statistics from single molecule kinetic experiments and topography of conformational energy landscape [4]. Due to the high similarity between folding and binding, for the binding kinetics, Wang et.al, further analytically studied the diffusion dynamics on biomolecular interface binding energy landscape. The resulting MFPT has also U curve dependence on the temperature. The complex kinetics and the associated fluctuations give the clue of the underlying statistical distribution which

reflects the topography of the binding energy landscape upon temperature changes [5]. In short, the analytical studies in conformational dynamics, protein folding and ligand binding all indicate the kinetics follow log-normal distribution at higher temperatures and power law distribution at the low temperatures. These analytical studies formed a mathematical foundation and motivation to explore the distribution in kinetics in real ligand binding.

Reference:

1. Lee CL, Stell G, Wang J* (2003) First-passage time distribution and non-markovian diffusion dynamics of protein folding. The Journal of chemical physics 118: 959–968.
2. J. D. Bryngelson and P. G. Wolynes (1989) Intermediates and Barrier Crossing in a Random Energy Model (with Applications to Protein Folding) J. Phys. Chem. 93: 6902-6915.
3. Y. Zhou*, C. Zhang, G. Stell, J. Wang*. Temperature Dependence of the Distribution of the First Passage Time: Results from Discontinuous Molecular Dynamics Simulations of an All-Atom Model of the Second α -Hairpin Fragment of Protein G. J. Am. Chem. Soc. 125 (2003) 6300-6305.
4. Wang J*, Xu L, Xue K, Wang E* (2008) Exploring the origin of power law distribution in single-molecule conformation dynamics: energy landscape perspectives. Chem. Phys. Lett. 463:405–409.

5. Wang J* (2006) Diffusion and single molecule dynamics on biomolecular interface binding energy landscape. Chem. Phys. Lett. 418: 544–548.

IVH. The fittings for the simulation results

The aim of nonlinear fitting is to get the parameter values which best describe the simulation results. The standard way of finding the best fit is to choose the parameters that would minimize the deviations of the theoretical curve(s) from the experimental points. We call the method as chi-square minimization.

$$X^2 = \sum_{i=1}^n \left[\frac{Y_i - f(x'_i, \hat{\theta})}{\sigma_i} \right]^2 \quad (1)$$

Where x'_i is the row vector for the i th ($i = 1, 2, \dots, n$) observation. Where $X = (x_1, x_2, \dots, x_k)'$ is the independent variables and $\theta = (\theta_1, \theta_2, \dots, \theta_k)'$ is the parameters. To estimate the $\hat{\theta}$ value with the least square method, we need to solve the normal equations which are set to be zero for the partial derivatives of X^2 with respect to each $\hat{\theta}_p$.

$$\frac{\partial X^2}{\partial \hat{\theta}_p} = -2 \sum_{i=1}^n \frac{1}{\sigma_i^2} \left[Y_i - f(x'_i, \hat{\theta}) \right] \left[\frac{\partial f(x'_i, \hat{\theta})}{\partial \hat{\theta}_p} \right] = 0 \quad (2)$$

In the study, we can't find a suitable fitting function in the built-in function library of the Origin package. Then we have defined a new fitting function for use in nonlinear fitting based on our analytical model,

defined as follows:

$$y = y_0 + A_1 e^{-0.5k_1(|x-xc_1|+(xc_1-x))} + \frac{A_2}{w\sqrt{\frac{\pi}{2}}} e^{-\frac{2(x-xc_2)^2}{w^2}} + A_3 e^{-0.5k_2(|x-xc_3|+(x-xc_3))} \quad (3)$$

We employ an iterative strategy to estimate the parameter values due to the difficulty of getting the explicit solutions to the equations. This process starts with some initial values, with each iteration, a X^2 value is computed and then the parameter values are adjusted so as to reduce the X^2 . When the X^2 values computed in two successive iterations are small enough (compared with the tolerance, herein, the tolerance is set to 1E-15, the maximum number of iterations is set to 50000), the fitting procedure has converged. We have also chosen the option "reduced chi-square" during the fitting procedure. The Levenberg-Marquardt (L-M) algorithm is used to adjust the fitting parameter values in the above iterative procedure in the current study. This algorithm combines ~~combining~~ the Gauss-Newton method and the steepest descent method works for most cases. In order to get a good value for parameter initialization, the function `peak_pos` is used to estimate the peak's XY coordinate, peak's width, area, etc for the Gauss distribution. There are two methods to calculate the confidence intervals for parameters: Asymptotic-Symmetry method and Model-Comparison method in the Origin package. We applied the former to obtain asymptotic, symmetrical confidence intervals. All Curve-fittings are carried out with software Origin

(www.originlab.com). We have listed the fitting reports for global fit including the Parameters of Figure 2-4 (See also Figure S1-S24),

Figure2: (the parameters of the fitting)

Equation	$y = y_0 + (A_1 * \exp(-0.5 * k_1 * (\text{abs}(x - x_{c1}) + (x_{c1} - x)))) + (A_3 * \exp(-0.5 * k_2 * (\text{abs}(x - x_{c3}) + (x - x_{c3})))) + (A_2 / (w * \sqrt{\pi/2})) * \exp(-2 * ((x - x_{c2}) / w)^2)$		
Adj.R-Square	0.99568		
		Value	Standard Error
Count	A1	246.97799	154.4804
Count	k1	1.88664	0.71839
Count	xc1	-7	0.29012
Count	A3	274.33063	160.24916
Count	k2	4	6.95249
Count	xc3	-5.5	0.31448
Count	y0	-271.90891	162.56591
Count	A2	1200.11898	349.83219
Count	w	1.06054	0.14476
Count	xc2	-6.02538	0.01637

Figure3: (the parameters of the fitting)

Equation	$y = y_0 + (A_1 * \exp(-0.5 * k_1 * (\text{abs}(x - x_{c1}) + (x_{c1} - x)))) + (A_3 * \exp(-0.5 * k_2 * (\text{abs}(x - x_{c3}) + (x - x_{c3})))) + (A_2 / (w * \sqrt{\pi/2})) * \exp(-2 * ((x - x_{c2}) / w)^2)$		
Adj.R-Square	0.99968		
		Value	Standard Error
Count	A1	384.23565	114.75243
Count	k1	1.1516	0.14429
Count	xc1	-11.47884	0.0452

Count	A3	392.46673	111.6123
Count	k2	4	3.4372
Count	xc3	-9.11778	0.10152
Count	y0	-390.15399	110.93278
Count	A2	1426.54232	404.84755
Count	w	1.68243	0.19921
Count	xc2	-10.14335	0.01006

Figure4: (the parameters of the fitting)

Time_{off}:

Equation	$y = y_0 + (A_1 * \exp(-0.5 * k_1 * (\text{abs}(x - x_{c1}) + (x_{c1} - x)))) + (A_3 * \exp(-0.5 * k_2 * (\text{abs}(x - x_{c3}) + (x - x_{c3})))) + (A_2 / (w * \sqrt{\text{PI}/2})) * \exp(-2 * ((x - x_{c2}) / w)^2)$		
Adj.R-Square	0.98077		
		Value	Standard Error
Count	A1	24.76535	24.524
Count	k1	4	7.71118
Count	xc1	6	0.10519
Count	A3	22.77664	24.40812
Count	k2	4	9.13005
Count	xc3	6.504	0.09885
Count	y0	-31.26799	34.4957
Count	A2	20.08047	1.81771
Count	w	0.18409	0.01697
Count	xc2	6.10355	0.00656

Time_{on}:

Equation	$y = y_0 + (A_1 * \exp(-0.5 * k_1 * (\text{abs}(x - x_{c1}) + (x_{c1} - x)))) + (A_3 * \exp(-0.5 * k_2 * (\text{abs}(x - x_{c3}) + (x - x_{c3})))) + (A_2 / (w * \sqrt{\text{PI}/2})) * \exp(-2 * ((x - x_{c2}) / w)^2)$		
Adj.R-Square	0.94671		
		Value	Standard Error
Count	A1	36.08573	684.33881
Count	k1	1.38737	31.1384
Count	xc1	3.84055	0.27757
Count	A3	10.69652	151.35361
Count	k2	4	88.21797

Count	xc3	4.33465	0.44884
Count	y0	-40.84552	700.78675
Count	A2	18.80556	3.85953
Count	w	0.24076	0.03814
Count	xc2	3.97482	0.01146

V. The force field used by AutoDock

In this study, the Autodock 4.2 was applied to perform the docking simulations. The AutoDock 4.2 uses a semi-empirical free energy force field to evaluate conformations during docking simulations. The force field was parameterized using a large number of protein-inhibitor complexes for which both the structures and inhibition constants or K_i , are available. The force field also includes an updated charge-based desolvation term, improvements in the directionality of hydrogen bonds, and several improved models of the unbound state. The AutoDock uses a computationally (relatively) inexpensive "hybrid" force field that contains terms based on molecular mechanics as well as empirical terms. The prediction of absolute binding energies may be less accurate compared to more computationally expensive, purely force field-based methods such as molecular dynamics, but this semi-empirical approach is considered as well-suited for the relative rankings.

The AutoDock semi-empirical force field includes intra-molecular terms, a "full" desolvation model, and also considers directionality in hydrogen bonds. The conformational entropy is calculated from the sum

of the torsional degrees of freedom. Water molecules are not modeled explicitly though, but pair-wise atomic terms are used to estimate the water contribution (dispersion/repulsion, hydrogen bonding, electrostatics, and desolvation), where weights are added for calibration. The evaluation step in a nutshell: firstly, calculate the energy of ligand and protein in the unbound state. Secondly, calculate the energy of the protein-ligand complex. Then take the difference between 1 and 2.

$$\Delta G = (V_{\text{bound}}^{\text{L-L}} - V_{\text{unbound}}^{\text{L-L}}) + (V_{\text{bound}}^{\text{P-P}} - V_{\text{unbound}}^{\text{P-P}}) + (V_{\text{bound}}^{\text{P-L}} - V_{\text{unbound}}^{\text{P-L}} + \Delta S_{\text{conf}})$$

Where P refers to the protein, L refers to the ligand, V are the pair-wise evaluations mentioned above, and ΔS_{conf} denotes the loss of conformational entropy upon binding (R Huey et al., 2006 1).

Reference:

1 Huey, Ruth, Garrett M. Morris, Arthur J. Olson, and David S. Goodsell. "A Semiempirical Free Energy Force Field with Charge-Based Desolvation." Journal of Computational Chemistry 28, no. 6 (April 30, 2007): 1145–52. doi:10.1002/jcc.20634.