Front. Phys., 2011, 6(1): 133–140
DOI 10.1007/s11467-010-0153-0

RESEARCH ARTICLE

Protein folding as a quantum transition between conformational states

Liao-fu LUO (罗辽复)

Laboratory of Theoretical Biophysics, Faculty of Physical Science and Technology, Inner Mongolia University, Hohhot 010021, China

E-mail: lolfcm@mail.imu.edu.cn

Received September 15, 2010; accepted November 29, 2010

Assuming that the main variables in the life processes at the molecular level are the conformation of biological macromolecules and their frontier electrons a formalism of quantum theory on conformation-electron system is proposed. Based on the quantum theory of conformation-electron system, the protein folding is regarded as a quantum transition between torsion states on polypeptide chain, and the folding rate is calculated by nonadiabatic operator method. The rate calculation is generalized to the case of frequency variation in folding. An analytical form of protein folding rate formula is obtained, which can be served as a useful tool for further studying protein folding. The application of the rate theory to explain the protein folding experiments is briefly summarized. It includes the inertial moment dependence of folding rate, the unified description of two-state and multistate protein folding, the relationship of folding and unfolding rates versus denaturant concentration, the distinction between exergonic and endergonic foldings, the ultrafast and the downhill folding viewed from quantum folding theory, and, finally, the temperature dependence of folding rate and the interpretation of its non-Arrhenius behaviors. All these studies support the view that the protein folding is essentially a quantum transition between conformational states.

Keywords protein folding rate, quantum transition, torsion states, non-Arrhenius temperature dependence, exergonic and endergonic folding, ultrafast folding

PACS numbers 87.15.He, 87.15.Cc, 03.65.Xp

1 Locality of conformation states and quantum number characterizing conformation-electronic system

Assume that the main variables in the life processes at the molecular level are the conformation of biological macromolecules and their frontier electrons. Suppose that the conformation is described by a set of torsion angles $\{\theta\} = (\theta_1, \dots, \theta_N)$. The conformation potential U, as a function of θ, may generally have some minima that correspond to stable conformations. To be definite, consider the case of two minima. Set

$$U(\theta) = \begin{cases} \dfrac{k}{2}\Big(\theta + \dfrac{\pi}{2}\Big)^2, & [-\pi,\ 0] \\[2ex] \dfrac{k}{2}\Big[(1-\lambda)\Big(\theta - \dfrac{\pi}{2}\Big)^2 + \dfrac{\pi^2}{4}\lambda\Big], & [0,\ \pi] \end{cases} \tag{1}$$

Here, λ is an asymmetrical parameter. We shall discuss the dependence of steady-state solutions on the asymmetrical parameters and demonstrate that the conclusions drawn will not depend on the particular choice of the potential when we use any two-minima-asymmetrical potential [1]. If $\lambda = 0$, $U(\theta)$ is C_2 symmetrical, then the conformational wave functions are nonlocal, and no definite local conformation can be related with the molecule. However, if $\lambda \neq 0$, the C_2-symmetry is broken; by numerical solution of Schrodinger equation for the potential given by Eq. (1), one can show that for $Ik/\hbar^2$=10 (I is the inertia moment of the molecule with respect to coordinate θ), the ground state should be localized to larger than 90% as $\lambda \geqslant 10^{-3}$, and for $Ik/\hbar^2 = 40$, the ground state localized to larger than 99% as $\lambda \geqslant 10^{-6}$. Define the locality J_i of the ith level as:

$$J_i = \frac{\max(P_i, Q_i)}{\min(P_i, Q_i)} \tag{2}$$

where P_i is the probability of ith level located in the "left" conformation ($-\pi \leqslant \theta \leqslant 0$) and Q_i in the "right" conformation ($0 \leqslant \theta \leqslant \pi$). By numerical calculation, one can show that nearly all bound states are localized with $J_i > 2$ as $\lambda \geqslant 0.1(0.01)$ for the case of $Ik/\hbar^2 = 10(40)$ [1]. The above discussion shows that the small asymmetry in potential would cause the high localization of wave functions. Although the results are deduced from a particular potential (1), the generality of the conclusions can be demonstrated by the following arguments. Consider the interaction between a pair of levels localized mainly in left potential valley A (the corresponding wave function denoted as ψ_A) or in right potential valley B (the corresponding wave function ψ_{B}), respectively. The overlap integral of wave functions $\int \psi_A^* \psi_B \mathrm{d}\theta$ in barrier region is a measure of the potential asymmetry and the wave function localization. By using WKB quasi-classical approximation to solve Schrodinger equation, we obtain the overlap integral proportional to $\exp\{-\int_\alpha^\beta K(\theta')\mathrm{d}\theta'\}$, where α and β are turning points at the left and right sides of the barrier, respectively, $U(a) = U(\beta) = E$ and the wave vector equals

$$K(\theta) = \frac{1}{\hbar}\{2I[U(\theta) - E]\}^{1/2} \tag{3}$$

For $U - E \sim 2.5$ kJ/mol (typical torsion energy) and $I = 10^{-37}\mathrm{g{\cdot}cm^2}$ (torsion inertial moment of typical atomic group), one has $K(\theta) \sim 10^2$. Therefore, the overlap integral is a small quantity, and the localization does occur. Consider that the conformational potential usually has certain asymmetry, we deduce that the levels below the top of barrier are always localized apart from those near the top where $K(\theta) \sim 0$. As for the problem whether the ground state is localized in the left or right valley, it is determined by the comparison of the depths of two valleys and the values of $\partial^2 U/\partial\theta^2$ in two minima of the potential. As Anderson studied the excitations on a disordered lattice in 1958, he proposed a noted theorem: the eigenfunctions are localized if the strength of the disorder exceeds some definite value [2]. Now, the theorem on the localization of the excitation of a lattice has been generalized to the system of conformations.

Suppose the dynamical variables of the conformation-electronic system are $(\theta,\ x)$, where x is the coordinate of frontier electrons, and θ is the torsion angle of molecule. Based on the above analyses on the conformational locality, the wave function of the system can be described by $M_{kn\alpha}(\theta, x)$, where α denotes the electron-state, and $(k,\ n)$ refer to the conformation- and vibration-state, respectively. The conformational quantum number $k = 1, 2, \ldots$, refers to the local conformational state, indicating which minimum of the torsion potential the system is localized around. The vibration quantum number n indicates the harmonic energy level in a given conformational state.

As a multi-atom system, the conformation of a protein is fully determined by bond lengths r_{ij}, bond angles χ_{i-j-k}, and torsion angles $(\theta_1, \theta_2, ..., \theta_N)$. Since torsion angles are most easily changed even at room temperature, we assume that the conformation is mainly described by a set of torsion coordinates, that is, the conformational potential U is only a function of $(\theta_1, \theta_2, ..., \theta_N)$. For a real problem of protein folding, sometimes $r_{ij}$and χ_{i-j-k} may not be fixed. In these cases, we shall generalize the theoretic approach as follows. We assume that the potential can be averaged over bond lengths and bond angles, and the averaged energy landscape is described by a set of torsion coordinates only. The average is reasonable since the folding does not involve a single microscopic pathway but rather a statistical energy landscape [3]. Through the statistical average of energy landscape and the corresponding redefinition of the potential, the number of torsion-accessible states will be largely increased. For example, for a pair of conformational states that are originally inaccessible, it is impossible to reach each other by changing torsion coordinate (or torsion quantum number in quantum theory) only. However, this is possible after the statistical average of energy landscape and redefinition of the form of torsion potential. Therefore, the statistical energy approach effectively expands the set of torsion-accessible states. In the following, we shall study protein folding as a changing process between a series of torsion-accessible states by use of quantum theory of conformation-electronic system.

2 Protein folding as a quantum transition calculated by nonadiabatic operator method

Protein folding is commonly studied in classical mechanics and classical statistical mechanics. However, this approach is too limited for the understanding of the fundamental physics underlying folding. It is well known that the fluorescence and phosphorescence are phenomena closely related to protein folding. Since no one doubts that fluorescence and phosphorescence could only be understood in terms of the quantum transition between molecules, why should the protein folding study be divorced from the framework of quantum theory? In the following we shall investigate protein folding from the quantum theory of conformation-electronic system.

Protein folding is nonradiative transition. The transition rate can be calculated by nonadiabatic operator method [4]. In adiabatic approximation, the wave function of the conformation-electronic system $M_{kn\alpha}(\theta, x)$ can be expressed as the product of conformational wave function and electronic wave function, $M(\theta,\ x) = \psi(\theta) \cdot \varphi(x,\ \theta)$. However, the form is not a rigorous eigenstate of the total Hamiltonian of the system, and there exist transitions between adiabatic states that result from the off-diagonal elements. The transition matrix element is given by off-diagonal element as follows:

$$\langle k'n'\alpha' \mid H' \mid kn\alpha\rangle = \int \psi^{+}_{k'n'\alpha'}(\theta)\Big[-\frac{\hbar^2}{2I}\int \varphi^{+}_{\alpha'} \cdot \Big(\frac{\partial^2\varphi_\alpha}{\partial\theta^2} + 2\frac{\partial\varphi_\alpha}{\partial\theta}\,\frac{\partial}{\partial\theta}\Big)\mathrm{d}^3x\Big]\psi_{kn\alpha}(\theta)\mathrm{d}\theta \tag{4}$$

Consider that the electronic state does not change in transition processes, $a' = a$. This is the generally occurring case for most protein foldings. For the case of radiation transition, the electronic transitions exist, accompanying conformational changes. We will discuss these phenomena from the point of conformation-electronic transition in another article. Since the electronic wave function φ_α is generally real, one has

$$\int \varphi_\alpha(x,\ \theta)\frac{\partial\varphi_\alpha(x,\ \theta)}{\partial\theta}\mathrm{d}^3x = 0 \tag{5}$$

from the normalization condition $\int \varphi_\alpha(x,\ \theta)\varphi_\alpha(x,\ \theta)\mathrm{d}^3x = 1$. Therefore, only the first term in right-hand-side of Eq. (4) is retained, namely

$$\langle k'n'\alpha \mid H' \mid kn\alpha\rangle = \int \psi^{+}_{k'n'\alpha}(\theta)\Big(-\frac{\hbar^2}{2I}\int \varphi^{+}_{\alpha}\frac{\partial^2\varphi_\alpha}{\partial\theta^2}\mathrm{d}^3x\Big)\psi_{kn\alpha}(\theta)\mathrm{d}\theta \tag{6}$$

For a multi-torsion-angle problem, the torsion potential $U_{\mathrm{tor}}(\theta_1,\theta_2,...,\theta_N)$ is generally argument-inseparable. Only when its conformational wave function can be approximately expressed by the product of single conformational functions, then the quantum transition problem can easily be solved. However, from the widely accepted statistical energy landscape theory [3], the protein folding involves a statistical average of a lot of microscopic pathways. One may average $U_{\mathrm{tor}}(\theta_1,\theta_2,...,\theta_N)$ over $N-1$ coordinates with only one angular coordinate θ_j left as the variable. The resulting potential is denoted as $U^{(j)}_{\mathrm{tor}}(\theta_j)$

$$U^{(j)}_{\mathrm{tor}}(\theta_j) = \langle U_{\mathrm{tor}}(\theta_1,\theta_2,...,\theta_N)\rangle_{\theta_1,...,\theta_{j-1},\theta_{j+1},...,\theta_N}$$

The summation of $U^{(j)}_{\mathrm{tor}}(\theta_j)$ over j gives an approximate expression of $U_{\mathrm{tor}}(\theta_1,\theta_2,...,\theta_N)$, and the argument-separable wave function is obtained. Therefore, the argument-separable torsion wave function as an approximation can be deduced from the statistical energy landscape approach. In argument-separable case, one has

$$\psi_{kn\alpha}(\theta) = \psi_{k_1,n_1,\alpha_1}(\theta_1),...,\psi_{k_N,n_N,\alpha_N}(\theta_N) \tag{7}$$

where $\psi_{k_j,n_j,\alpha_j}(\theta_j)$ is approximately expressed by a wave function of harmonic oscillator with quantum number n_j. Note that the harmonic potential has equilibrium position at $\theta_j = \theta^{(0)}_{kj}$ with the corresponding k_jth minimum of potential E_{kj} ($k_j = 1, 2, \ldots$). Inserting Eq. (7) into Eq. (6), one obtains

$$\langle k'n'\alpha \mid H' \mid kn\alpha\rangle = \int \psi^{+}_{k'n'\alpha}(\theta)\sum\Big(-\frac{\hbar^2}{2I_j}\int \varphi^{+}_{\alpha}\frac{\partial^2\varphi_\alpha}{\partial\theta_j^2}\mathrm{d}^3x\Big)\psi_{kn\alpha}(\theta)\mathrm{d}\theta \tag{8}$$

After thermal average of the initial states, the transition rate will be

$$W = \frac{2\pi}{\hbar}\sum_{\{n\}}\big|\langle k'n'\alpha \mid H' \mid kn\alpha\rangle\big|^2 B(\{n\},T) \cdot \delta\Big(\sum_j (n'_j\hbar\omega'_j - n_j\hbar\omega_j - \delta E_j)\Big) \tag{9}$$

Here, ω_j and ω'_j are frequency parameters of the jth mode in conformational state k_j and k'_j, respectively, and $\delta E_j = E_{kj} - E_{k'j}$. $B(\{n\},T)$ denotes the Boltzmann factor:

$$B(\{n\},T) = \prod_j B(n_j,T) = \prod_j \Big[\mathrm{e}^{-n_j\beta\hbar\omega_j}(1-\mathrm{e}^{-\beta\hbar\omega_j})\Big] \tag{10}$$

After summing over final states, we have

$$W = \frac{2\pi}{\hbar}\sum_{\{n\}}|\langle k'n'\alpha \mid H' \mid kn\alpha\rangle|^2 B(\{n\},T)\rho_E \tag{11}$$

Here, ρ_E means state density,

$$\begin{aligned}\rho_E &= \frac{1}{\partial E_f/\partial N_f} = \frac{1}{\hbar\bar{\omega}'}, \quad N_f = \sum_j n'_j \\ E_f &= \sum_j (n'_j\hbar\omega'_j + E_{k'j})\end{aligned} \tag{12}$$

$\overline{\omega'}$ is the average of ω'_j over j.

3 Protein folding rate in single torsion transition

If only one torsion angle participates in the conformational transition, we call it single-mode or single torsion transition; if several torsion angles participate simultaneously in one conformational transition, then we call it multimode or multitorsion transition. Consider single-mode case at first. In this case, the subscript j will be dropped. The transitional rate

$$W = \frac{2\pi}{\hbar^2\omega'}I_E I_V \tag{13}$$

$$I_E = \left|\frac{-\hbar^2}{2I}\int \varphi_\alpha\frac{\partial^2}{\partial\theta^2}\varphi_\alpha\mathrm{d}^3x\right|^2_{\theta=\theta_0} \tag{14}$$

$$I_V = \sum_n\left|\int \psi^{+}_{k'n'\alpha}(\theta)\psi_{kn\alpha}(\theta)\mathrm{d}\theta\right|^2 B(n,T) \tag{15}$$

θ_0 is the conformational coordinate value in the largest overlap region of vibration functions. Eq. (14) can be estimated by the square of rotational kinetic energy of the frontier electron during conformational change (see next section). The quantum number n' in Eq. (15) is determined by n through energy conservation. The overlap integral in Eq. (15) was first calculated by Manneback

(1951) [5, 6]. For simplicity, we shall assume the same frequency for initial and final states ($\omega = \omega'$) first. After thermal average, it gives

$$I_V = \Big(\frac{\overline{n}+1}{\overline{n}}\Big)^{p/2} J_p(2Q\sqrt{\bar{n}(\bar{n}+1)})\exp(-Q(2\overline{n}+1)) \tag{16}$$

in which

$$\bar{n} = (\mathrm{e}^{\beta\hbar\omega} - 1)^{-1} \tag{17}$$

$$Q = I\omega(\delta\theta)^2/(2\hbar), \quad p = \frac{\delta E}{\hbar\omega} \tag{18}$$

$\delta\theta = \theta_k^{(0)} - \theta_{k'}^{(0)}$ is the angular displacement, and $\delta E = E_k - E_{k'}$ the energy gap between initial and final states. J_P denotes the modified Bessel function [7], and here, p is related to the net change in oscillator quantum number. The modified Bessel function is introduced from the expansion:

$$\exp\Big(\frac{x}{2}\Big(y + \frac{1}{y}\Big)\Big) = \sum_{n=-1}^{\infty} y^n J_n(x) \tag{19}$$

By use of the asymptotic formula for Bessel function [8]

$$\mathrm{e}^{-z} J_p(z) = (2\pi z)^{-1/2} \exp\Big(-\frac{p^2}{2z}\Big), \quad \text{for } z \gg 1 \tag{20}$$

I_V can be further simplified. Note that the equality $\bar{n} + \frac{1}{2} = \sqrt{\bar{n}(\bar{n}+1)}$ holds approximately with error 2% as $\bar{n} = 2$ and smaller than 1% as $\bar{n} > 3$. Considering $\bar{n} > 1$ ($\hbar\omega < 0.69k_{\mathrm{B}}T$), from Eqs. (16) and (20), it leads to

$$I_V = (2\pi z)^{-1/2} \exp\Big(-\frac{p^2}{2z}\Big) \exp\Big(\frac{\delta E}{2k_{\mathrm{B}}T}\Big) \tag{21}$$

where

$$z = (\delta\theta)^2 \cdot \frac{k_{\mathrm{B}}T}{\hbar^2} I \tag{22}$$

4 Protein folding rate in multitorsion transition

For multitorsion transition, from Eqs. (11) and (12), one has

$$W = \frac{2\pi}{\hbar^2\bar{\omega}'} I_E \sum_{\{p_j\}} \prod_j I_{Vj} \tag{23}$$

$$I_E = \Big|\sum_j \frac{-\hbar^2}{2I_j} \int \varphi_\alpha \frac{\partial^2}{\partial\theta_j^2} \varphi_\alpha \mathrm{d}^3x\Big|^2_{\theta_j = \theta_{j_0}} \tag{24}$$

$$\begin{aligned} I_{Vj} &= \exp(-Q_j(2\bar{n}_j + 1))\Big(\frac{\bar{n}_j + 1}{\bar{n}_j}\Big)^{p_j/2} \\ &\quad \cdot J_{p_j}\Big(2Q_j\sqrt{\bar{n}_j(\bar{n}_j+1)}\Big) \end{aligned} \tag{25}$$

with

$$\bar{n}_j = (\mathrm{e}^{\beta\hbar\omega_j} - 1)^{-1} \tag{26}$$

$$Q_j = I_j\omega(\delta\theta_j)^2/(2\hbar) \tag{27}$$

$$p_j = \frac{\delta E_j}{\hbar\omega_j} \tag{28}$$

Here, $\delta\theta_j$ is the angular displacement, and δE_j is the energy gap between the initial and final states for the jth mode. p_j represents the net change in quantum number for oscillator mode j, which satisfies the constraint

$$\sum_j p_j = p \tag{29}$$

in the summation of Eq. (23), and p satisfies

$$p = \frac{\Delta E}{\hbar\bar{\omega}} \quad \Big(\Delta E = \sum_j \delta E_j\Big) \tag{30}$$

$\bar{\omega}$ is the average of ω_j, namely, $\sum_j \hbar\omega_j p_j = \hbar\bar{\omega}p$. Note that Eq. (25) takes the same form as Eq. (16). To simplify $\sum_{\{p_j\}} \prod_j I_{Vj}$ in Eq. (23), we consider two-mode ($N = 2$) case at first,

$$\begin{aligned} \sum_{p_1+p_2=p} I_{V1} I_{V2} &= \Big(\frac{\bar{n}_1+1}{\bar{n}_1}\Big)^{p_1/2} \Big(\frac{\bar{n}_2+1}{\bar{n}_2}\Big)^{p_2/2} \\ &\cdot \sum_{p_1} \exp\{-Q_1(2\bar{n}_1+1)\} J_{p_1}(2Q_1\sqrt{\bar{n}_1(\bar{n}_1+1)}) \\ &\cdot \exp(-Q_2(2\bar{n}_2+1)) J_{p-p_1}(2Q_2\sqrt{\bar{n}_2(\bar{n}_2+1)}) \end{aligned} \tag{31}$$

By use of the asymptotic formula for Bessel function, Eq. (20), we deduce

$$\begin{aligned} &\sum_{p_1} \exp(-z_1) J_{p_1}(z_1) \exp(-z_2) J_{p-p_1}(z_2) \\ &= \frac{1}{2\pi\sqrt{z_1z_2}} \int_{-\infty}^{\infty} \mathrm{d}p_1 \exp\Big(-\Big(\frac{p_1^2}{2z_1} + \frac{(p-p_1)^2}{2z_2}\Big)\Big) \\ &= \frac{1}{\sqrt{2\pi}} \frac{1}{\sqrt{z_1+z_2}} \exp\Big(-\frac{p^2}{2(z_1+z_2)}\Big) \end{aligned} \tag{32}$$

Under $\bar{n}_j$ >1, by inserting (32) into (31), we have

$$\begin{aligned} \sum_{p_1+p_2=p} I_{V_1} I_{V_2} &= \frac{1}{\sqrt{2\pi}} \frac{1}{\sqrt{z_1+z_2}} \\ &\cdot \exp\Big(-\frac{p^2}{2(z_1+z_2)}\Big) \cdot \exp\Big(\frac{\Delta E}{2k_{\mathrm{B}}T}\Big) \end{aligned} \tag{33}$$

where

$$Z_j = (\delta\theta_j)^2 \cdot \frac{k_{\mathrm{B}}T}{\hbar^2} I_j \tag{34}$$

Eq. (33) takes the same form as Eq. (21) for the single-mode.

Next, by using the similar deduction, we obtain the conclusion that if

$$\begin{aligned} \sum_{p_1+\cdots+p_{n-1}=q} &\exp(-z_1) J_{p_1}(z_1) \exp(-z_2) J_{p_2}(z_2) \cdots \\ &\cdot \exp(-z_{n-1}) J_{p_{n-1}}(z_{n-1}) \\ = \frac{1}{\sqrt{2\pi}} &\frac{1}{\sqrt{z_1+\cdots+z_{n-1}}} \exp\Big\{-\frac{q^2}{2(z_1+\cdots+z_{n-1})}\Big\} \end{aligned} \tag{35}$$

holds, then it leads to

$$\begin{aligned}&\sum_{p_1+\cdots+p_n=p}\exp(-z_1)J_{p_1}(z_1)\exp(-z_2)J_{p_2}(z_2)\cdots\\&\quad\cdot\ \exp(-z_n)J_{p_n}(z_n)\\&=\frac{1}{\sqrt{2\pi}}\frac{1}{\sqrt{z_1+\cdots+z_n}}\exp\Big(-\frac{p^2}{2(z_1+\cdots+z_n)}\Big)\end{aligned}\tag{36}$$

Therefore, for multimode transition, we obtain

$$\begin{aligned}\sum_{\{p_j\}}\prod_j I_{Vj}&=\frac{1}{\sqrt{2\pi}}\exp\Big(\frac{\Delta E}{2k_{\rm B}T}\Big)\Big(\sum_j Z_j\Big)^{-\frac{1}{2}}\\&\quad\cdot\ \exp\Big(-\frac{p^2}{2\sum\limits_j Z_j}\Big)\end{aligned}\tag{37}$$

The electronic matrix elements Eq. (14) or Eq. (24) can be estimated as follows:

$$\begin{aligned}I_E&=\left(\sum_j\frac{\hbar^2}{2I_j}\langle l_j^2\rangle\right)^2\\&=\frac{\hbar^4}{4}\left(\sum_j\frac{a_j}{I_j}\right)^2,\quad a_j=\langle l_j^2\rangle>0\end{aligned}\tag{38}$$

l_j is the jth magnetic quantum number (with respect to θ_j) of electronic wave function $\varphi_\alpha(x,\theta)$, $a_j=\langle l_j^2\rangle\approx O(1)$; here, $O(1)$ means a number in the order of magnitude of 1.

Inserting Eqs. (37) and (38) into Eq. (23) and using Eqs. (30) and (34), we obtain the final results of multimode transition [9–11]:

$$\begin{aligned}W&=\frac{\hbar^3\sqrt{\pi}}{2\sqrt{2}\delta\theta\bar{\omega}'}(k_{\rm B}T)^{-1/2}\exp\left(\frac{\Delta E}{2k_{\rm B}T}\right)\left(\sum_j I_j\right)^{-1/2}\\&\quad\cdot\left(\sum_j\frac{a_j}{I_j}\right)^2\exp\left(\frac{-(\Delta E)^2}{2\bar{\omega}^2(\delta\theta)^2k_{\rm B}T\sum\limits_j I_j}\right)\\\delta\theta&=\sqrt{\langle(\delta\theta_j)^2\rangle_{\rm av}}\end{aligned}\tag{39}$$

5 Protein folding rate in frequency varying case

In above deduction, the same frequency for initial and final states ($\omega_j=\omega_j'$) has been assumed in calculating overlap integral and obtaining Eqs. (16) and (25). Now, we shall generalize the results to the case of nonequal frequencies between initial and final states. From statistical mechanics theory, the free energy of a system of oscillators is expressed by

$$\begin{aligned}G_{\rm os}&=\frac{1}{\beta}\sum_j\Big[\ln(1-\exp(-\beta\omega_j))+E_j\Big]\\\beta&=\frac{1}{k_{\rm B}T}\end{aligned}\tag{40}$$

where E_j means the potential minimum of the jth oscillator. By using

$$\frac{\partial G_{\rm os}}{\partial(\hbar\omega_j)}=\langle n_j\rangle=\frac{1}{\exp(\beta\hbar\omega_j)-1}\tag{41}$$

(n_j: phonon number of the jth mode), we obtain the free energy variation

$$\begin{aligned}\delta_\omega G_{\rm os}&=\sum_j\int_{\omega_j}^{\omega_j'}\frac{\partial G_{\rm os}}{\partial\omega_j}{\rm d}\omega_j\\&=\sum_j\int_{\omega_j}^{\omega_j'}\frac{\hbar}{\exp(\beta\hbar\omega_j)-1}{\rm d}\omega_j\end{aligned}\tag{42}$$

due to the frequency shift from ω_j to ω_j'. It leads to

$$\delta_\omega G_{\rm os}=\sum_j\frac{1}{\beta}\ln\frac{\omega_j'}{\omega_j}\tag{43}$$

as $\beta\hbar\omega_j\ll 1$. Therefore, from Eq. (40), the free energy difference between torsion initial state (frequency $\{\omega_j\}$) and final state (frequency $\{\omega_j'\}$) is

$$\Delta G_{\rm os}=-\delta_\omega G_{\rm os}+\sum_j\delta E_j=\Delta E+\sum_j\frac{1}{\beta}\ln\frac{\omega_j}{\omega_j'}\tag{44}$$

Correspondingly, the net change of phonon number p in Eqs. (21) and (37) will be replaced by

$$p=\sum_j\frac{\delta E_j}{\hbar\omega_j}+\sum_j\frac{1}{\beta}\frac{\ln\frac{\omega_j}{\omega_j'}}{\hbar\omega_j}\tag{45}$$

Considering that the contribution of frequency variation to folding rate comes mainly from the torsion vibration term, instead of Eq. (39), we obtain a generalized equation of folding rate [12]

$$\begin{aligned}W&=\frac{\hbar^3\sqrt{\pi}}{2\sqrt{2}\delta\theta\bar{\omega}'}(k_{\rm B}T)^{-1/2}\exp\Big(\frac{\Delta G_{\rm os}}{2k_{\rm B}T}\Big)\left(\sum_j I_j\right)^{-1/2}\\&\quad\cdot\left(\sum_j\frac{a_j}{I_j}\right)^2\exp\left(\frac{-(\Delta G_{\rm os})^2}{2\bar{\omega}^2(\delta\theta)^2k_{\rm B}T\sum\limits_j I_j}\right)\end{aligned}\tag{46}$$

in the frequency variation case, where $\Delta G_{\rm os}$ is given by Eq. (44). The application of free energy gap $\Delta G_{\rm os}$ instead of energy gap ΔE was first proposed in the photosynthetic electron transfer when the frequency variation is small [13]. From Eq. (44), we know that when

$$\left|\sum_j\ln\frac{\omega_j}{\omega_j'}\right|\ll\frac{|\Delta E|}{k_{\rm B}T}\tag{47}$$

the effect of frequency variation can be neglected and Eq. (46) returns to Eq. (39).

6 Application of quantum folding theory to interpretation of experiments

The following problems about the application of the proposed rate theory have been discussed in Ref. [12], and here, we give a brief summary. All these studies support the view that the protein folding is essentially a quantum transition between conformational states.

1) The torsion inertial moment dependences of protein folding rate are calculated. The inertial moments are dynamical variable describing torsion motion of main-chain and side-chain dihedral angles. In principle, they can be estimated by the number of dihedral angles and the mass of related atomic groups on polypeptide chain for each contact. A model of dynamical contact order was proposed for calculating inertial moment dependences of folding rate, and consistent results were obtained by comparison with 100-protein experimental dataset [14]. The correlation coefficient of inertial moment term in lnW with experimental folding rate ln k_f is 0.85, higher than or comparable with other calculations and predictions for protein folding rate based on the same concept of contact order [15, 16].

2) The unified theory on the two-state and multistate protein folding is worked out based on the concept of quantum transition. Under the assumption of the successive transitions of multistate proteins taking place in different torsion modes, the multistate protein folding rate is simply expressed as the rate of two-state protein folding of same topology multiplied by a time delay factor $\exp(-\tau)$. From a set of 80 proteins (45 two-state and 35 non-two-state), we compared the results of theoretical calculations of folding rate with experimental data and the distribution of the deviation between theoretical and experimental folding rates is given in Fig. 1 [14]. In the calculation of multistate protein folding the time delay factor has been introduced. The peak in Fig. 1 indicates the theory basically consistent with experiments and demonstrates the correctness of the above idea on multistate protein folding [14, 17].

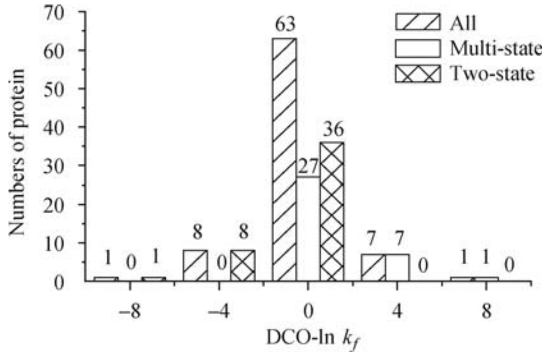

Fig. 1 Distribution of the deviation between theoretical and experimental folding rates. Horizontal scale value is the median value of DCO-ln k_f value range, such as 0 indicate $-2 <$ DCO-ln $k_f < 2$, etc. ln k_f means logarithm experimental folding rate, and DCO means theoretical folding rate. In theoretical calculation, the inertial moment of a protein is estimated by the number of dihedral angles and the mass of related atomic groups on polypeptide chain for each contact, the torsion vibration free energy of all proteins is classified into three categories – exergonic, endergonic and neutral, and the best-fit time $\tau = 3.5$ in the delay factor $\exp(-\tau)$ for multistate protein has been introduced.

The distribution is plotted for 80 protein dataset [14]. In addition to this dataset, by using 27 two-state folding proteins from the "standard" set given by Maxwell *et al.* [18], the result can be further improved. (The peak of protein numbers at DCO-ln $k_f = 0$ increases from 63 to 90 [14].)

3) The relationship of folding and unfolding rates versus denaturant concentration is studied. One may assume δE_j, and therefore ΔE, is a linear function of denaturant concentration. This interprets the observed slope of the arm of the chevron plot, a plot of $\ln W$ versus denaturant concentration [18]. Another factor is the change of the inertial moment I_j due to the varying viscosity of the solvent. One may assume the effective inertial moment I_j is approximately a linear function of denaturant concentration. This can interpret the observed deceleration of folding as the viscosity of the solvent is increased [19].

4) The exergonic and endergonic foldings are distinguished through the comparison between theoretical and experimental rates for each protein. Exergonic means the torsion vibration free energy decreasing in the reaction, $\Delta G_{\mathrm{os}} > 0$, while endergonic means the torsion vibration free energy increasing in the reaction, $\Delta G_{\mathrm{os}} < 0$. We have predicted exergonic/endergonic folding of a protein, and these predictions are basically consistent with experimental data [14].

5) The ultrafast folding problem [20–22] is viewed from the point of quantum folding theory, and a new folding speed limit is deduced from uncertainty relation. From the point of quantum transition, we demonstrate that the necessary conditions for ultrafast folding are given as follows: first, the folding is exergonic rather than endergonic; second, the large initial/final energy gap ΔE in torsion potential; and third, the large initial/final frequency ratio of torsion vibration. On the other hand, considering that the high-speed folding generally occurs in downhill processes [23], we have discussed the implication of downhill folding in quantum theory and proved that $(\delta\theta)^2$ should be small in downhill process (Fig. 2). This is another factor to speed up the folding rate. The problem of protein folding "speed limit" is viewed from quantum uncertainty relation, and we suggest a new physical speed limit of picoseconds as the limitation for the local and transient fold formation on the polypeptide chain.

6) The temperature dependence of folding rate is deduced, and the non-Arrhenius behaviors of temperature dependence are interpreted in a natural way. From Eq. (46), one obtains

$$\ln W(T) = \frac{\Delta E(1-\frac{\Delta E}{\varepsilon})}{2k_{\mathrm{B}}T} - \frac{1}{2}\ln T - \frac{\lambda^2}{2\varepsilon}k_{\mathrm{B}}T + const \quad (48)$$

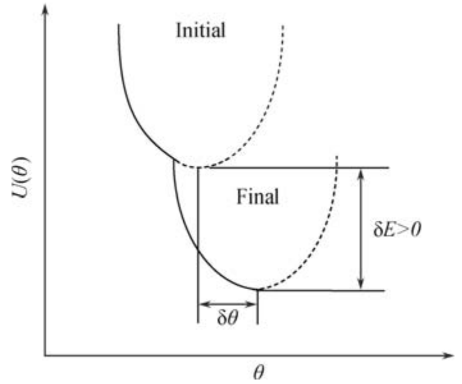

Fig. 2 $U(\theta)$–θ relation for downhill folding. Figure 2 is plotted for typical torsion mode j for down-hill folding where the barrier disappears. The subscript j in θ_j and E_j has been dropped for simplicity. Suppose the intersection of two potential at θ_c the downhill requires θ_c smaller than both potential minima $\theta_1^{(0)}$ and $\theta_2^{(0)}$. When initial frequency equals final frequency, θ_c has a simple expression, and one has $(\delta\theta)^2 < \frac{2\delta E}{I\omega^2}$.

$$\lambda = \sum_j \ln \frac{\omega_j}{\omega'_j}, \quad \varepsilon = \bar{\omega}^2(\delta\theta)^2 \sum I_j \tag{49}$$

Therefore, the relation of $\ln W$ versus $1/T$ is not linear in the Arrhenius plot. The slope–temperature relation is given by

$$\frac{\mathrm{d}\ln W}{\mathrm{d}(\frac{1}{T})} = S + \frac{1}{2}T + RT^2 \tag{50}$$

$$S = \frac{\Delta E(1 - \frac{\Delta E}{\varepsilon})}{2k_{\mathrm{B}}}$$

$$R = \frac{k_{\mathrm{B}}\lambda^2}{2\varepsilon} \tag{51}$$

Note that if the structural phase transition occurs for some proteins then the potential parameter may change with temperature and an additional term will introduced in slope S and R, but Eq. (50) remains unchanged.

The experiments on rate–temperature relationships in protein folding exhibit the following characteristics of non-Arrhenius behavior, and they all can be interpreted by the theory above: i) All nonlinearity in Arrhenius plot of $\ln W$ versus $1/T$ shows the decrease in their folding rates upon increase in temperature at high temperature (slightly curved downward at high temperature) [24–26]. This can be explained by two temperature-dependent terms, $-\frac{1}{2}\ln T - \frac{\lambda^2}{2\varepsilon}k_{\mathrm{B}}T$, in Eq. (48). ii) In terms of activation energy barrier, the normal positive barrier means that the slope S of Arrhenius plot is negative. However, the Arrhenius plot of folding rate versus temperature for some proteins shows abnormal negative activation energy barriers. In this theory, the negative activation energy barriers are explained by the energy gap between the initial and final states $\Delta E > 0$, which makes slope positive in Arrhenius plot. The case occurs more generally in ultrafast folders. iii) For some proteins, the crossover from normal positive barrier to abnormal negative at high-temperature occurs. That is, a maximum is observed in Arrhenius plot. The crossover can be interpreted by $\Delta E < 0$ ($S < 0$) of the protein at lower temperatures, but the symbol of $\frac{\mathrm{d}\ln W}{\mathrm{d}(1/T)}$ changes at high temperature due to the temperature-dependent terms in Eq. (50). Note that the $\frac{1}{2}T$ term alone is not enough for interpreting the strong nonlinearity, and RT^2 is a key term for the occurrence of crossover. iv) The plots of $\ln W$ versus $\frac{1}{T}$ are strongly curved for refolding of some proteins but almost linear for their unfolding under denaturant. This can be explained by the denaturant concentration effect. The denaturant possibly strengthens the torsion force field and increases the energy gap ΔE between initial and final states in the unfolding experiments. Therefore, $|S|$ increases with denaturant concentration, and the relative contribution of non-Arrhenius term RT to the rate $\ln W$ decreases [see Eq. (48)]. v) Recent experiments in λ-repressor fragments folding exhibited another characteristic of rate–temperature relationships. The set of mutants containing more glycines versus alanines exhibit stronger non-Arrhenius behavior [24]. This may be due to the change of torsion potential in the mutation of amino acid. The residue dependence can also be interpreted by smaller inertia moment of glycine than alanine since the non-Arrhenius term RT^2 is greater in glycine-rich mutant [see the dependence of ε on inertia moment, Eq. (49)].

To conclude, the fundamental physics underlying protein folding may be much simpler than the extremely complexity inherent in the protein structure, as proposed by Baker ten years ago [27]. Here, our analysis indicates that a key point on the surprising logical simplicity may be in the numerous types of protein folding obey the same universal kinetics of quantum transition between conformational states. The proposed theoretical formulation in the article can, in principle, interpret all basic experimental facts on protein folding. Moreover, the view of quantum transition among torsion states gives deeper insights into the folding event of polypeptide chain.

Acknowledgements The author was indebted to Dr. Jun Lu and Dr. Ying Zhang for their helpful discussions.

References

1. L. F. Luo, Int. J. Quant. Chem., 1987, 32: 435
2. D. Shepelyansky, Symposium Anderson Localization in Nonlinear and Many-Body Systems, Dresden, 2009
3. J. D. Bryngelson, J. N. Onuchic, N. D. Socci, and P. G. Wolynes, Proteins, 1995, 21:167
4. K. Huang and A. Rhys, Proc. Roy. Soc. (London), 1950, A204: 406
5. D. Devault, Quart. Rev., Biophysics, 1980, 13: 387
6. J. Jortner, J. Chem. Phys., 1976, 64: 4860
7. M. Abramowitz and I. A. Stegun, Handbook of Mathemat-

ical Functions, 10th printing with corrections, National Bureau of Standards, Applied Mathematics Series 55, 1972

8. G. N. Watson, A Treatise on the Theory of Bessel Functions, 2nd Ed., Series: Cambridge Mathematical Library, 1995
9. L. F. Luo, Int. J. Quant. Chem., 1995, 54: 243
10. L. F. Luo, Theoretic-Physical Approach to Molecular Biology, Shanghai: Shanghai Scientific & Technical Publishers, 2004: 437, 457
11. L. F. Luo, arXiv: qbio/0906.2452, 2009
12. L. F. Luo, arXiv: qbio/1008.0237, 2010
13. T. Kakitani and H. Kakitani, Biochim. et Biophys. Acta, 1981, 635: 498
14. Y. Zhang and L. F. Luo, Scientia Sinica Vitae, 2010, 40: 887, doi:10.1360/052010-337
15. K.W Plaxco, T. Simons, and D. Baker, J. Mol. Biol., 1998, 277(4): 985
16. D. N. Ivankov and A. V. Finkelstein, Proc. Natl. Acad. Sci. USA, 2004, 101: 8942
17. K. Kamagata, M. Arai, and K. Kuwajima, J. Mol. Biol., 2004, 339: 951
18. K. L. Maxwell, D. Wildes, A. Zarrine-Afsar, M. A. De Los Rios, A. G. Brown, C. T. Friel, L. Hedberg, J. C. Horng, D. Bona, E. J. Miller, A. Vallée-Bélisle, E. R. Main, F. Bemporad, L. Qiu, K. Teilum, N. D. Vu, A. M. Edwards, I. Ruczinski, F. M. Poulsen, B. B. Kragelund, S. W. Michnick, F. Chiti, Y. Bai, S. J. Hagen, L. Serrano, M. Oliveberg, D. P. Raleigh, P. Wittung-Stafshede, S. E. Radford, S. E. Jackson, T. R. Sosnick, S. Marqusee, A. R. Davidson, and K. W. Plaxco, Protein Sci., 2005, 14(3): 602
19. M. Jacob, T. Schindler, J. Balbach, and F. X. Schmid, Proc. Natl. Acad. Sci. USA, 1997, 94 (11): 5622
20. L. Qiu, S. A. Pabit, A. E. Roitberg, and S. J. Hagen, J. Am. Chem. Soc., 2002, 124(44): 12952
21. Y. Zhu, D. O. V. Alonso, K. Maki, C. Y. Huang, S. J. Lahr, V. Daggett, H. Roder, W. F. DeGrado, and F. Gai, Proc. Natl. Acad. Sci. USA, 2003, 100: 15486
22. H. Neuweiler, C. M. Johnson, and A. R. Fersht, Proc. Natl. Acad. Sci. USA, 2009, 106(44): 18569
23. M. M. Garcia-Mira, M. Sadqi, N. Fischer, J. M. Sanchez-Ruiz, and V. Munoz, Science, 2002, 298: 2191
24. M. L. Scalley and D. Baker, Proc. Natl. Acad. Sci. USA, 1997, 94: 10636
25. W. Y. Yang and M. Gruebele, Biochemistry, 2004, 43: 13018
26. K. Ghosh, B. Ozkan, and K. A. Dill, J. Am. Chem. Soc., 2007, 129: 11920
27. D. Baker, Nature, 2000, 405: 39