Review

Quantum conformational transition in biological macromolecule

Liaofu Luo[1],* and Jun Lv[2],*

[1] School of Physical Science and Technology, Inner Mongolia University, Hohhot 010021, China
[2] Center for Physics Experiment, College of Science, Inner Mongolia University of Technology, Hohhot 010051, China
* Correspondence: lolfcm@imu.edu.cn, lujun@imut.edu.cn

Background: Recently we proposed a quantum theory on the conformational change of biomolecule, deduced several equations on protein folding rate from the first principles and discussed the experimental tests of the theory. The article is a review of these works.
Methods: Based on the general equation of the conformation-transitional rate several theoretical results are deduced and compared with experimental data through bioinformatics methods.
Results: The temperature dependence and the denaturant concentration dependence of the protein folding rate are deduced and compared with experimental data. The quantitative relation between protein folding rate and torsional mode number (or chain length) is deduced and the obtained formula can be applied to RNA folding as well. The quantum transition theory of two-state protein is successfully generalized to multi-state protein folding. Then, how to make direct experimental tests on the quantum property of the conformational transition of biomolecule is discussed, which includes the study of protein photo-folding and the observation of the fluctuation of the fluorescence intensity emitted from the protein folding/unfolding event. Finally, the potential applications of the present quantum folding theory to molecular biological problems are sketched in two examples: the glucose transport across membrane and the induced pluripotency in stem cell.
Conclusions: The above results show that the quantum mechanics provides a unifying and logically simple theoretical starting point in studying the conformational change of biological macromolecules. The far-reaching results in practical application of the theory are expected.

Keywords: conformational change; quantum transition; protein folding; RNA folding; temperature dependence

INTRODUCTION

Since the birth of molecular biology, the rationalization of life science has happened one wave after another. The post-genomic era calls for the deep integration of physical science and life science. Based on our personal experience and perspective the integration can occur on several important areas, for example, on the folding mechanism of protein and nucleic acid molecules, on the evolution of an informational system such as a genome, on the molecular basis of cognition. It is interesting to note that many of these studies are closely related to quantum mechanics. Recently, protein and RNA folding was explained by the quantum transition between conformational states [1–4]. The genome evolutionary equation was written in terms of nucleotide k-mer frequencies varying in time and the equation was transformed into a quantum-mechanical one to describe the stochastic speciation events [5]. The quantum bit was proposed to process the information in brain based on microtubules or nuclear spins and the consciousness was interpreted as a quantum effect [6–9]. In all above studies the role of quantum was emphasized. In his famous book *What is Life?* Schrödinger wrote that we must be prepared to find a new type of physical law prevailing in the living matter and the new principle is not a non-physical but "a genuinely physical one: it is nothing else than the principle of quantum theory over again." Hence from

the points of the existence of several unifying principles between life and non-life and the deep integration of physics and biology we learned about the importance of quantum. However, in the present article we will focus on reviewing our recent works on the quantum folding problem, the problem of the conformational change of biomolecules viewed as quantum transition. In the section of *General formula for conformational transition rate*, a general formula for conformational transition rate will be introduced. Then the temperature dependence and denaturant concentration dependence of protein folding rate will be discussed in next section. The dependence of conformational transition rate on torsion mode number N of macromolecule and the test of the theoretical formula in protein and RNA folding experiments will be studied in the section of *N-dependence of the folding rate for protein and RNA molecule*. In the subsequent section several discussions will be given, including how to generalize the theory of two-state protein folding to multi-state, the protein photo-folding, the quantum coherence and the direct experimental test on the quantum property of the conformational change of macromolecules. Finally, as two examples of the application of the quantum folding theory, the glucose transport across membrane and the induced pluripotency in stem cell will be sketched in the last section.

GENERAL FORMULA FOR CONFORMATIONAL TRANSITION RATE

The protein folding/unfolding is a typical example of the conformational change of biomolecules. The two-state protein folding can be looked at as a one-step quantum transition between conformational states. The main dynamical variables of the system include torsion (dihedral) angles, bond lengths and bond angles which determine the molecular shape. The former (torsion namely) is the slow-varying variable, denoted as $\{\theta\}$ and the latter two are fast-varying variables, denoted as $\{x\}$. For a conformational change problem, if the electronic variables, chemical bonds, hydrogen bonds and other forms of energy such as hydrophobic interaction etc. are important and should be considered, they all can be included in the fast-varying variable set. The wave function $M(\theta, x)$ satisfies

$$\left(H_{tor}\left(\theta, \frac{\partial}{\partial\theta}\right) + H_{fv}\left(x, \frac{\partial}{\partial x}; \theta\right)\right)M(\theta, x) = EM(\theta, x),$$

(1)

$$H_{tor} = -\sum \frac{\hbar^2}{2I_j}\frac{\partial^2}{\partial\theta_j^2} + U_{tor}(\theta).$$

(2)

H_{tor} and H_{fv} are slow- and fast-variable Hamiltonian

respectively. In adiabatic approximation the wave function is expressed as

$$M(\theta, x) = \psi(\theta)\varphi(x, \theta),$$

(3)

and the two factors satisfy

$$H_{fv}\left(x, \frac{\partial}{\partial x}; \theta\right)\varphi_\alpha(x, \theta) = \varepsilon_\alpha(\theta)\varphi_\alpha(x, \theta),$$

(4)

$$\left\{H_{tor}\left(\theta, \frac{\partial}{\partial\theta}\right) + \varepsilon_\alpha(\theta)\right\}\psi_{kn\alpha}(\theta) = E_{kn\alpha}\psi_{kn\alpha}(\theta),$$

(5)

respectively where α denotes the quantum number of fast-variable wave function φ, and (k, n) refers to the conformational (indicating which minimum the wave function is localized around) and the vibrational state of torsion wave function ψ.

Because $M(\theta, x)$ is not a rigorous eigenstate of Hamiltonian $H_{tor} + H_{fv}$, there exists a transition between adiabatic states that results from the off-diagonal elements

$$\langle k' n' \alpha' |H'|kn\alpha\rangle$$

$$= -\int \psi_{k' n' \alpha'}^+(\theta)$$

$$\cdot \sum_j \frac{\hbar^2}{2I_j}\left\{\int \varphi_{\alpha'}^+\left(\frac{\partial^2\varphi_\alpha}{\partial\theta_j^2} + 2\frac{\partial\varphi_\alpha}{\partial\theta_j}\frac{\partial}{\partial\theta_j}\right)dx\right\}\psi_{kn\alpha}(\theta)d\theta.$$

(6)

Through tedious calculation we obtain the rate of conformational transition [1,2]

$$W = \frac{2\pi}{\hbar^2\overline{\omega}'}I_V' I_E',$$

$$I_V' = \frac{\hbar}{\sqrt{2\pi}\delta\theta}\exp\left(\frac{\Delta G}{2k_BT}\right)\exp\left(-\frac{(\Delta G)^2}{2\zeta k_BT}\right)(k_BT)^{\frac{1}{2}}\left(\sum_j^N I_j\right)^{\frac{1}{2}},$$

$$I_E' = \sum_j^M |a_{\alpha'\alpha}^{(j)}|^2 = M\overline{a}^2,$$

(7)

in which $\zeta = \overline{\omega}^2(\delta\theta)^2\sum_j^N I_j$ as $z = \frac{k_BT}{\hbar^2}(\delta\theta)^2 I_0 \gg 1$ and the condition $z \gg 1$ is generally satisfied in protein folding problem. I_V' is slow-variable factor and I_E' fast-variable factor of the transitional rate, N is the number of torsion modes participating coherently in a quantum transition, I_j denotes the inertial moment of the atomic group of the j-th torsion mode (I_0 denotes its average), $\overline{\omega}$ and $\overline{\omega}'$ are respectively the initial and final frequency parameters ω_j and ω_j' of torsion potential averaged over N torsion modes, $\delta\theta$ is the averaged angular shift between initial

and final torsion potential, ΔG is the free energy decrease per molecule between initial and final states, M is the number of torsion angles correlated to fast variables, $\overline{a}^2$ is the square of the matrix element of the fast-variable Hamiltonian operator, or, more accurately, its change with torsion angle, averaged over M modes,

$$\overline{a}^2 = \frac{1}{M} \sum_j^M |a_{\alpha'\alpha}^{(j)}|^2,$$

$$a_{\alpha'\alpha}^{(j)} = \frac{i\hbar}{I_j^{1/2}} \frac{h_{\alpha'\alpha}^{(j)}}{\varepsilon_\alpha^{(0)} - \varepsilon_{\alpha'}^{(0)}}, \quad \alpha' \neq \alpha,$$

$$h^{(j)}\left(x, \frac{\partial}{\partial x}\right) = \left(\frac{\partial H_{fv}\left(x, \frac{\partial}{\partial x}; \theta\right)}{\partial \theta_j}\right)_0, \quad (8)$$

$\varepsilon_\alpha^{(0)}$, $\varepsilon_{\alpha'}^{(0)}$ are the eigenvalues of $H_{fv}\left(x, \frac{\partial}{\partial x}; \theta_0\right)$. Equations (7–8) are basic equations for conformational transition.

Note 1: The adiabatic approximation is the only important approximation in the above deduction. The approximation is applicable because the energy gap between the eigenvalues $\varepsilon_\alpha^{(0)}$ of fast-variable Hamiltonian is generally larger than the torsion energy [10].

Note 2: The torsion potential U_{tor} is a function of a set of torsion angles $\theta = \{\theta_j\}$. Its form is dependent of solvent environment of the molecule. Suppose the interaction between water (or other solvent, ion and denaturant) molecules (their coordinates denoted by r) and macromolecule is $V(r, \theta, x)$, its average over r in a given set of experimental conditions (including chemical denaturants, solvent conditions etc.) [11] can be expressed by $\langle V(r, \theta, x)\rangle_{av} = V_1(\theta) + V_2(x, \theta)$ where $V_1(\theta)$ is x-independent part of the average interaction. Then we define $U_{tor}(\theta) = U_{tor, vac}(\theta) + V_1(\theta)$ and $H_{fv}\left(x, \frac{\partial}{\partial x}; \theta\right)$
$= H_{fv,vac}\left(x, \frac{\partial}{\partial x}; \theta\right) + V_2(x, \theta)$ in the basic Equations (1) and (2) with $U_{tor,vac}$ the torsion potential in vacuum and $H_{fv,vac}$ the fast-variable Hamiltonian in vacuum. Therefore, although the influence of solvent is a difficult problem in molecular dynamics approach, it has been taken into account automatically in the present theory by redefining the torsion potential U_{tor} and fast-variable Hamiltonian H_{fv} in the basic equations.

Note 3: The torsion mode number N describes the coherence degree of multi-torsion transition in the folding. For protein we assume that N is obtained by numeration of all main-chain and side-chain dihedral angles on the polypeptide chain except those residues on its tail which does not belong to any contact. Each residue in such contact fragment contributes 2 main-chain dihedral angles and, for non-alanine and-glycine, it contributes 1–4 additional side-chain dihedral angles. For nucleic acid the torsion number can be estimated by chain length. Following IUB/IUPAC there are 7 torsion angles for each nucleotide, namely

$$\alpha(O3' - P - O5' - C5'),$$
$$\beta(P - O5' - C5' - C4'),$$
$$\gamma(O5' - C5' - C4' - C3'),$$
$$\delta(C5' - C4' - C3' - O3'),$$
$$\varepsilon(C4' - C3' - O3' - P),$$
$$\varsigma(C3' - O3' - P - O5'),$$

and

$$\chi(O4' - C1' - N1 - C2) \quad \text{(for Pyrimidine)},$$

or $\chi(O4' - C1' - N9 - C2)$ (for Purine),

of which many have more than one advantageous conformations (potential minima). If each nucleotide has q torsion angles with multi-minimum in potential then the torsion number $N = qL$, where L is chain length.

TEMPERATURE AND DENATURANT CONCENTRATION DEPENDENCE OF PROTEIN FOLDING

From the rate Equation (7) we have deduced a temperature dependence law of protein folding rate $W(T)$

$$\ln W(T) = \frac{S}{T} - RT + \frac{1}{2}\ln T + const, \quad (9)$$

where *const* means temperature-independent term. The law explains the curious non-Arrhenius behavior of the rate — temperature relationships in protein folding/unfolding experiments [1,3]. It gives an example in Figure 1.

The comprehensive comparisons of the theoretical predictions with experimental rates were made in [3] for all two-state proteins whose temperature dependence data were available. The strong curvature of folding rate on Arrhenius plot is due to the R term in Equation (9) which comes from the square free energy $(\Delta G)^2$ in rate Equation (7) for theoretical $\ln W$. Since the factor $(\Delta G)^2$ occurs only in rate equation of quantum theory the good agreement between theory and experiments affords support to the concept of quantum folding. Moreover, in this theory the universal non-Arrhenius characteristics of folding rate are described by only two parameters S and R and these

FBP28(Y11R/W30F)(1E0L)
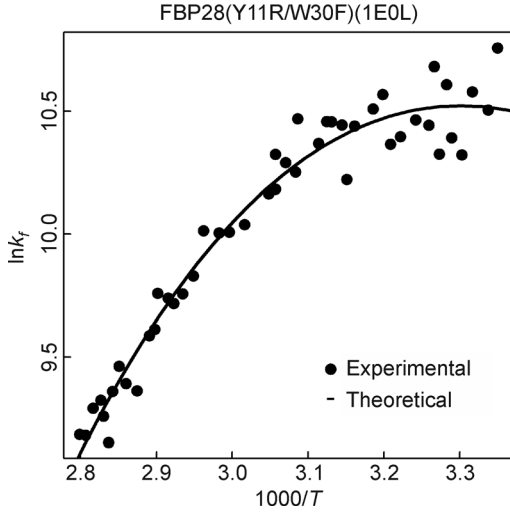
Figure 1. Model fit to overall folding rate k_f vs temperature 1000/T for protein FBP28. Experimental logarithm folding rates are shown by "●" (data taken from [12]), and solid lines are theoretical model fits to the folding rate (k_f in unit s^{-1}, T in unit Kelvin).

parameters are related to the known folding dynamics. All parameters related to torsion potential defined in this theory (such as torsion frequency $\overline{\omega}$ and $\overline{\omega}'$, averaged angular shift $\delta\theta$ and energy gap ΔE between initial and final torsion potential minima, etc.) can be determined, calculated consistently with each other for all studied proteins [3]. Furthermore, in this theory the folding and unfolding rates are correlated with each other, needless of introducing any further assumption as in [13]. An interesting relation we obtained is

$$R'/R = (\overline{\omega}/\overline{\omega}')^2 \qquad (10)$$

(slope S and curvature R in folding are denoted as S', R' in unfolding). The equation explains why for some proteins the plots of ln$W$ versus $1/T$ are strongly curved but almost linear for their unfolding by $\overline{\omega} < \overline{\omega}'$.

Assuming the free energy change ΔG in a temperature interval lower than melting temperature T_c has been measured and expressed as

$$\Delta G = \Delta G_0 + \Delta G_1 (T - T_c) = \alpha + \beta T,$$

$$\alpha = \Delta G_0 - \Delta G_1 T_c, \ \beta = \Delta G_1. \qquad (11)$$

We know that the above linear relation was tested by experiments for many proteins [12,14–24] (Supplementary Figure S1). By using α and β given from experiments we can re-deduce the temperature dependence of folding rate. In fact, by inserting Equation (11) into Equation (7) one easily obtains Equation (9) and in the equation the slope and curvature parameters S and R are given by

$$S = \frac{(\Delta G_0 - \Delta G_1 T_c)}{2k_B} \left(1 - \frac{\Delta G_0 - \Delta G_1 T_c}{\zeta}\right) = \frac{\alpha}{2k_B}\left(1 - \frac{\alpha}{\zeta}\right),$$

$$R = \frac{(\Delta G_1)^2}{2\zeta k_B} = \frac{\beta^2}{2\zeta k_B}. \qquad (12)$$

Eliminating ζ in Equation (12) a universal relation between R and S is deduced,

$$R = \left(\frac{\alpha}{2k_B} - S\right)\frac{\beta^2}{\alpha^2}. \qquad (13)$$

Following the similar deduction the relation between slope and curvature parameters for unfolding rate is obtained,

$$R' = -\left(\frac{\alpha}{2k_B} + S'\right)\frac{\beta^2}{\alpha^2}. \qquad (14)$$

With the aid of known α and β from the temperature dependence of free energy, constraints on slope and curvature parameters can be obtained by Equations (13) and (14) for folding/unfolding rate.

In addition to Equations (13) and (14), from the relation between folding rate W and unfolding rate W_u

$$\ln\left\{\frac{W}{W_u}\right\} = \frac{\Delta G}{k_B T} + \frac{(\Delta G)^2}{2k_B T \varepsilon}\left(\frac{\overline{\omega}^2 - \overline{\omega}'^2}{\overline{\omega}'^2}\right) + \ln\frac{\overline{\omega}}{\overline{\omega}'},$$

$$\varepsilon = N\overline{\omega}^2(\delta\theta)^2 I_0 \qquad (15)$$

(ε is defined as ζ in Equation (7) as $z \gg 1$) we obtain the relation between R and R'

$$\frac{\Delta G_0}{k_B T_c} + \frac{(\Delta G_0)^2}{2k_B T_c \varepsilon}\left(\frac{R'}{R} - 1\right) + \frac{1}{2}\ln\frac{R'}{R} = 0. \qquad (16)$$

Here T_c is defined by $W = W_u$ and as $T = T_c$ one has $\Delta G = \Delta G_0$.

Equation (13) was proved at precision higher than 90% by using experimental data on S, R, α and β of 15 two-state proteins [4]. Equations (14) and (16) can also be tested by experimental data. It is interesting to note that $\Delta G(T_c) \equiv \Delta G_0 \neq 0$, which has been proved by the linear temperature-dependence of free energy (Supplementary Figure S1). If ΔG_0 were 0 then R' would equal R from Equation (16). So the inequality between R' and R indicated by experiments would require $\Delta G_0 \neq 0$. Of course, ΔG_0 is a small quantity in the order of $k_B T_c |\ln\frac{\overline{\omega}}{\overline{\omega}'}|$ that can be seen from Equation (15). Usually the free energy ΔG at given temperature T was measured through $\Delta G = k_B T\ln(k_f/k_u)$ (the experimental value of W and W_u denoted as k_f and k_u respectively) in literature. However from Equation (15), this determination of free

energy is not accurate as T near T_c where $\ln(k_f/k_u)$ is a small quantity and the term proportional to $\ln(\overline{\omega}/\overline{\omega}')$ cannot be neglected. So, we argue that the present measurement of free energy at temperature T near T_c, namely ΔG_0, is not accurate. We expect Equation (16) will be experimentally tested by more precise measurement data of ΔG_0 and R' value.

Note: The temperature dependence Equation (9) for protein folding rate can be generalized to other macromolecular conformational changes as long as the slow variable is torsion in the process. We have demonstrated if only electrons serve as the fast variables the folding rate of a torsion-electron system takes the nearly same form of temperature dependence as Equation (9) except the unimportant term ½ ln$T$ changes to $-$½ ln$T$ [1,2]. Generally speaking, for fast-variables with energy level spacing $\Delta\varepsilon_a \gg k_B T$ the temperature variation of several tens of degrees would not markedly change the statistical distribution of fast-variable quantum states and therefore, the temperature dependence Equation (9) still holds. The unusual temperature response of intrinsically disordered protein may be attributed to the change of protein helicity [25]. The thermal sensitivity of secondary structure content induces the shift of the energy level of the protein and causes the change of the temperature dependence of its folding rate.

The denaturant concentration dependence of protein folding rate can be discussed in the same framework. By setting $\Delta G = \Delta G^{(0)} + mc$, $c =$ [denaturant] , $m = \dfrac{\partial(\Delta G)}{\partial c}$ at given temperature we obtain

$$
\begin{aligned}
\ln W &= \ln W^{(0)} + m_f c + m_f' c^2, \\
\ln W_u &= \ln W_u^{(0)} + m_u c + m_u' c^2,
\end{aligned}
\tag{17}
$$

from Equation (7). Here $W^{(0)}$ and $W_u^{(0)}$ represent the folding and unfolding rates in water, m_f and m_u are the respective slopes of the folding and unfolding arms, m_f' and m_u' describe their curvature

$$
m_f = \frac{m}{2k_B T}\left(1 - \frac{2\Delta G^{(0)}}{\zeta}\right); \; m_u = -\frac{m}{2k_B T}\left(1 + \frac{2\Delta G^{(0)}}{\zeta'}\right);
$$

$$
m_f' = -\frac{m^2}{2k_B T\zeta}; \; m_u' = -\frac{m^2}{2k_B T\zeta'};
$$

$$
\zeta = NI_0\overline{\omega}^2(\delta\theta)^2, \; \zeta' = NI_0\overline{\omega}'^2(\delta\theta)^2.
\tag{18}
$$

One may use the four parameters, m_f, m_u, m_f' and m_u' to fit the experimental data of denaturant dependence. Setting $\dfrac{\zeta'}{\zeta} = \dfrac{\overline{\omega}'^2}{\overline{\omega}^2} = 1 + \delta$ ($|\delta|$ is a small quantity) one has

$$
\begin{aligned}
m_f + m_u &= \frac{-m\Delta G^{(0)}}{k_B T\zeta}(2-\delta), \\
m_f' - m_u' &= \frac{-m^2\delta}{2k_B T\zeta}.
\end{aligned}
\tag{19}
$$

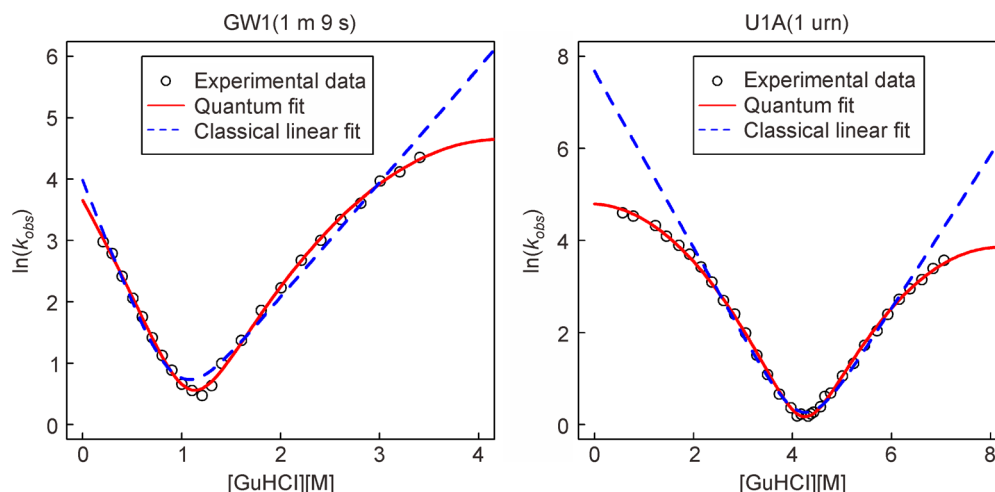If δ is near 0 and neglectable then $m_f' = m_u' = m'$ and $\zeta = \zeta' = 2\Delta G^{(0)}\dfrac{m_u - m_f}{m_f + m_u}$. In this case only three parameters m_f, m_u, and m' are needed for fitting the experimental data. By using chevron plot (the relation of the relaxation rate constant $\ln k_{obs} = \ln(W + W_u)$ vs denaturant concentration c) to fit the experiments [11] two examples are given in Figure 2 (details can be found in Supplementary Table S1). In Figure 2 the classical fits are taken from [11] where only the linear terms are retained in the chevrons, while the quantum fits with three or four parameters are based on the abovementioned Equations (17) to (19). From Figure 2 one finds the quantum fits are obviously better than the classical linear fits. In the meantime we notice that the polynomial fitting Equation (17) was reported recently on the frataxin folding in denaturant urea which is in good agreement with experiments [26]. The polynomial fitting is easily understood in our quantum approach.

N-DEPENDENCE OF THE FOLDING RATE FOR PROTEIN AND RNA MOLECULE

The protein and RNA folding rate is dependent of the number N of the torsion modes participating in the transition. As seen from Equation (7) the dependence comes mainly from two factors: the fast-variables factor I_E' and the free energy ΔG.

For a large class of conformational change problems such as the common protein and RNA folding, the chemical reaction and electronic transition are not involved and the fast variables include only bond lengths and bond angles of the macromolecule. In this case an approximate relation of the fast-variable factor I_E' with respect to torsion number N can be deduced. When the kinetic energy in $H_{fv}\left(x, \dfrac{\partial}{\partial x}; \theta\right)$ is neglected as compared with interaction potential U_{fv} one has

$$
a_{\alpha'\alpha}^{(j)} = \frac{i\hbar}{I_j^{1/2}}\frac{1}{\varepsilon_\alpha^{(0)} - \varepsilon_{\alpha'}^{(0)}}\int \varphi_{\alpha'}^*(\mathbf{r}, \mathbf{r}')\left(\frac{\partial U_{fv}(|\mathbf{r}-\mathbf{r}'|, \theta)}{\partial\theta_j}\right)_{\theta_0}\varphi_\alpha(\mathbf{r}, \mathbf{r}')d^3\mathbf{r}d^3\mathbf{r}' = \frac{i\hbar}{I_j^{1/2}}\frac{1}{V(\varepsilon_\alpha^{(0)} - \varepsilon_{\alpha'}^{(0)})}\int\left(\frac{\partial U_{fv}(|\mathbf{r}-\mathbf{r}'|, \theta)}{\partial\theta_j}\right)_{\theta_0}d^3(\mathbf{r}-\mathbf{r}').
\tag{20}
$$

Figure 2. **The statistical analysis on the denaturant concentration dependence of folding /unfolding rates for protein GW1 and U1A.** Experimental logarithm relaxation rates $\ln k_{obs}$ are shown by "o". Quantal and classical model fits to the experimental rates are given by solid and dashed lines respectively. The data of experimental folding /unfolding rates and classical fits are taken from [11].

In the above deduction of the second equality the fast-variable wave function $\varphi_\alpha(\mathbf{r}, \mathbf{r}')$ has been assumed to be a constant and normalized in the volume V. As the energy and volume V are dependent of the size of the molecule one may assume energy $\varepsilon_\alpha^{(0)}$ and U_{fv} proportional to the interacting-pair number (namely N^2) and V proportional to N. However, because only a small fraction of interacting-pairs are correlated to given $\theta_j (j=1, ..., N)$, $\left(\dfrac{\partial U_{fv}(|\mathbf{r} - \mathbf{r}'|, \theta)}{\partial \theta_j}\right)_{\theta_0}$ does not increase with N. So, one estimates $a_{\alpha'\alpha}^{(j)} \approx N^{-3}$. On the other hand, the integral $\int \left(\dfrac{\partial U_{fv}(|\mathbf{r} - \mathbf{r}'|, \theta)}{\partial \theta_j}\right)_{\theta_0} d^3(\mathbf{r} - \mathbf{r}')$ may depend on the molecular structure. For example, the high helix content makes the integral increasing. It was indicated that a protein with abundant α helices may have a quite oblong or oblate ellipsoid, instead of spheroid, shape and this protein has higher folding rate [3,27]. Therefore, apart from the factor N^{-3} there is another independent structure-related factor in $a_{\alpha'\alpha}^{(j)}$. Furthermore, assuming M proportional to N, one obtains

$$M\overline{a}^2 = cfN^{-5}, \qquad (21)$$

where f is a structure-related shape parameter. It means the fast-variable factor I_E' is inversely proportional to N^5. The result is consistent with the direct fit of experimental data of 65 proteins to a power law $M\overline{a}^2 \approx N^{-d}$ that gives d equal about 5.5 to 4.2 [3]. Finally from Equations (7) and (21) we obtain the N-dependence of folding rate

$$\ln W = \frac{\Delta G}{2k_B T} - \frac{(\Delta G/k_B T)^2}{2\rho N} - 5.5\ln N + \ln c_0 f, \qquad (22)$$
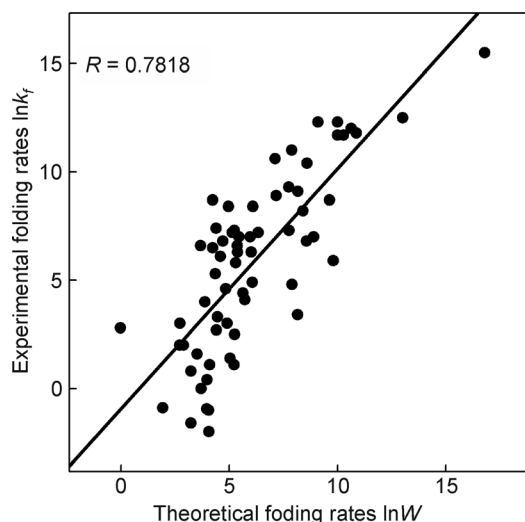
where

$$\rho = I_0 \overline{\omega}^2 (\delta\theta)^2 / (k_B T)$$

is a torsion-energy-related parameter and c_0 is an N-independent constant, proportional to c. The relationship of $\ln W$ with N given by Equation (22) can be tested by the statistical analyses of 65 two-state protein folding rates k_f, which is shown in Figure 3. The details can be found in Ref. [3,4]. The theoretical logarithm rate $\ln W$ is in good agreement with the experimental $\ln k_f$.

To find the relation between free energy ΔG and torsion number N we consider the statistical relation of free energy combination $\dfrac{\Delta G}{2k_B T} - \dfrac{(\Delta G)^2}{2(k_B T)^2 \rho N}$ that occurs in rate Equations (7) or (22). Assuming the free energy differences are measured under a "standard" set of experimental conditions [11] and setting

$$\frac{\Delta G}{2k_B T} - \frac{(\Delta G)^2}{2(k_B T)^2 \rho N} = y, \quad -\frac{1}{N} = x \qquad (23)$$

for the 65-protein dataset [27] we find a good linear relation $y = A + Bx$ existing where A and B are two statistical parameters describing the free energy distribution. The correlation R in the linear regression is near to 0.8 for $\rho = 0.065 \sim 0.075$ and reaches maximum $R = 0.7966$ at $\rho = 0.069$. Thus, by single-ρ-fit we obtain the best-fit statistical relation of free energy for two-state

Figure 3. Comparison of theoretical folding rates lnW with experimental folding rates lnk_f for 65 two-state proteins. Experimental rates for 65 two-state proteins are taken from the database published in [27] where the folding experiments were carried out at temperature around 25°C and the rates were extrapolated to denaturant-free case. The details of the parameter choice in theoretical calculation can be found in [3,4]. The linear regression is given by the solid line with correlation $R = 0.7818$.

proteins as

$$y = 4.306 + 541.1x, \quad (\rho = 0.069) \tag{24}$$

(Figure 4A). As the variation of ρ for different proteins is taken into account the linear regression between free energy combination y and torsion number x will be further improved.

Because N increases linearly with the length L of polypeptide chain, instead of Equation (23), by setting

$$\frac{\Delta G}{2k_B T} - \frac{(\Delta G)^2}{2(k_B T)^2 \rho_L L} = y, \quad -\frac{1}{L} = x, \tag{25}$$

we obtain the best-fit statistical relation of free energy for two-state proteins (Figure 4B) as

$$y = 3.626 + 83.76x, \quad (\rho_L = 0.28) \tag{26}$$

and the correlation coefficient $R = 0.781$. (Figure 4B)

About the relationship of free energy ΔG with torsion number N or chain length L several proposals were proposed in literature. One statistics was done by assuming the linear relation between ΔG and $\sqrt{N}$, $\Delta G = a\sqrt{N} - b$ ($b \neq 0$) [3]. Another was based on the assumed relation of ΔG vs ($Lg + \sigma B_L L^{2/3}$) [27]. By the statistics on 65 two-state proteins in the same dataset we demonstrated that the correlation R between free energy and N or L is 0.67 for the former and 0.69 for the latter [3], both lower than the correlation shown in Figure 4. We shall use the statistical relation of the free-energy combination, Equation (23), versus N in the following studies on RNA folding.

In virtue of Equations (22–24) we obtain an approximate expression for transitional rate lnW versus N for protein folding
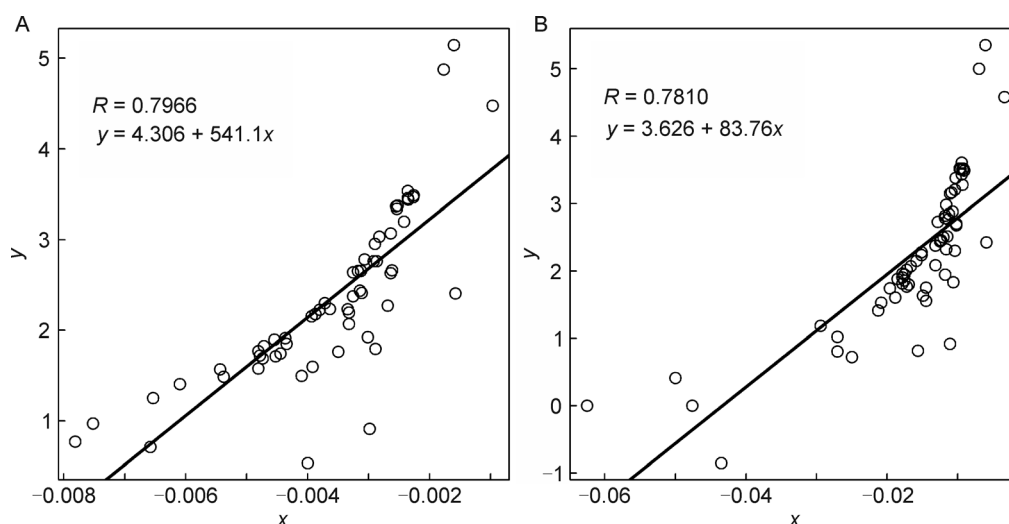
Figure 4. Statistical relation of free energy for two-state proteins. Experimental data are taken from 65- protein set [3,27]. Five proteins in the set denatured by temperature have been omitted in our statistics. (A) y and x are defined as Equation (23). (B) y and x defined as Equation (25). The linear regression between y and x is plotted by the solid line and R means the correlation coefficient.

$$\ln W = A - \frac{B}{N} - D\ln N + const, \qquad (27)$$

$$(D = 5.5,\ A = 4.306,\ B = 541.1).$$

The quantum folding theory of protein is applicable in principle for each step of the conformational transition of RNA molecule. Although recent experiments have revealed multi-stages in RNA collapse, the final search for the native structure within compact intermediates seems a common step in the folding process. Moreover, the step exhibits strong cooperativity of helix assembly [28,29]. Because the collapse transition prior to the formation of intermediate is a fast process and the time needed for the former is generally shorter than the latter, the calculation of the transition from intermediate to native fold can be directly compared with the experimental data of total rate. By using $N = qL$ (L is the chain length of RNA), instead of Equation (27), we have

$$\ln W = A - \frac{B'}{L} - D\ln L + c',$$

$$(B' = B/q,\ c' = const - D\ln q), \qquad (28)$$

as an alternative expression of folding rate versus chain length. Equation (28) is deduced from quantum folding theory with some statistical consideration and it predicts the rate W increasing with L, attaining the maximum at $L_{\max} = B'/D$, then decreasing with power law L^{-D}.

In a recent work Hyeon and Thirumalai [30] indicated that the chain length determines the folding rates of RNA. They obtained a good empirical relation between folding rate and chain length L in a dataset of 27 RNA sequences. Their best-fit result is

$$\ln W_H = 14.3 - 1.15 \times L^{0.46}. \qquad (29)$$

Both Equations (28) and (29) give us the relation between RNA folding rate and chain length. Comparing the theoretical folding rates $\ln W$ or $\ln W_H$ with the experimental folding rates $\ln k_f$ in 27 RNA dataset the results are shown in Figure 5. We found that the theoretical Equation (28) can fit the experimental data on RNA folding rate equally well as the empirical Equation (29). By using the best-fit value of B' and D the correlation between $\ln W$ (calculated from Equation (28)) and $\ln k_f$ is $R = 0.9729$ (Figure 5A), while the correlation between $\ln W_H$ (calculated from Equation (29)) and $\ln k_f$ is $R = 0.9752$ (Figure 5B). However, in Figure 5B the slope of the regression line is 1.03 and the line deviates from origin by -0.36, while in Figure 5A the slope is 1.0001, very close to 1 and the line deviates from origin only by -0.0012. The reason lies in: although the two equations have the same overall accuracy in fitting experimental data, for large L the errors $Er = |\log W - \log k_f|$ calculated from Equation (28) are explicitly lower than $Er_H = |\log W_H - \log k_f|$ from Equation (29) (details can be found in Ref. [4]). It means the folding rate lowers down with increasing L as L^{-D} ($D \cong 5.5$) at large L (a long-tail existing in the W-L curve) rather than a short tail as $\exp(-\lambda\sqrt{L})$ assumed in [30]. The long-tail form of folding rate can be used to explain some small-probability events in pluripotency conversion of gene (see the section of *Applications of quantum folding theory in discussion*).

There are two independent parameters in RNA folding rate Equation (28), B' and D, apart from the additive
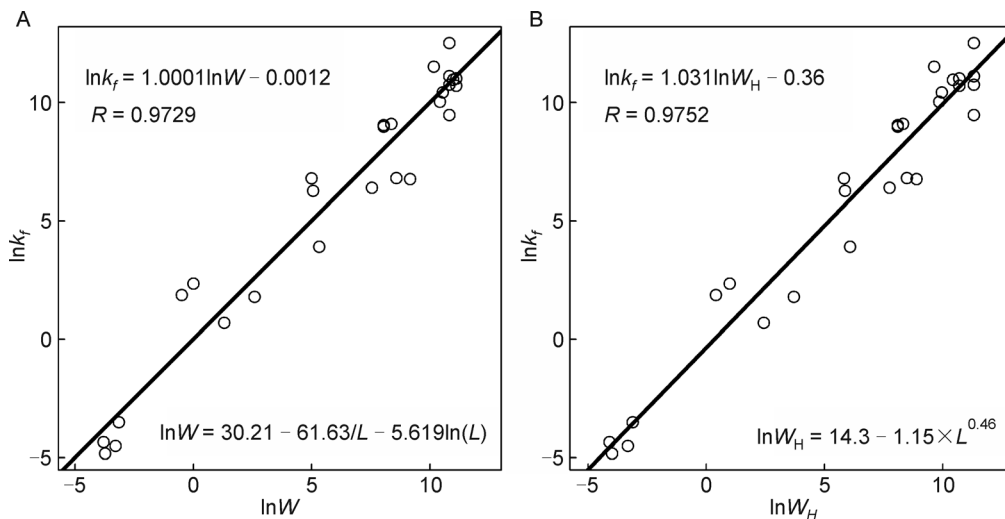
Figure 5. **Comparison of experimental folding rates** $\ln k_f$ **with theoretical folding rates** $\ln W$ (**Panel A**) **or** $\ln W_H$ (**Panel B**) **for 27 RNA molecules.** Experimental rates are taken from Ref. [30]. Theoretical rates are calculated from Equations (28) and (29) and shown on the lower right of two panels. The regression equation between $\ln k_f$ and $\ln W$ (or $\ln W_H$) is given on the upper left of the panel and plotted by the solid line. R means the correlation coefficient of the regression analysis.

constant. As seen from Figure 5A we obtain the best-fit D value $D_f = 5.619$ on the 27-RNA dataset, close to $D = 5.5$ predicted from a general theory of quantum folding. Simultaneously we obtain the best-fit B' value $B'_f = 61.63$. The B'_f value derived from RNA folding can be compared with the $B = 541.1$ from protein folding (Equation (27) if each nucleotide in RNA containing 7 torsion angles and the free energy difference between protein and RNA are taken into account .

DISCUSSION

Multi-state protein folding

For a long time the multi-state folding mechanism was unclear in theory. Kamagata *et al.* indicated that the folding rates of non-two-state proteins show the similar dependence on the native backbone topological parameters as for the two-state proteins [31]. Quantum folding theory provides a unified point on the folding mechanism of non-two-state and two-state proteins. Zhang and Luo proposed that the multi-state folding can be viewed as a joint of several quantum transitions with independent degrees of freedom of torsion angle. So the total collapse rate can be expressed by the formula of two-state folding but with an additional factor indicating the time delay in intermediate state [32].

Consider 3-state protein folding. Assuming polypeptide chain divided into to 2 parts and denoting initial state as $|i_1 i_2\rangle$ and final state as $\langle f_1 f_2|$. The transitional matrix element of 1st order perturbation $\langle f_1 f_2|(H'_1 + H'_2)|i_1 i_2\rangle = 0$ where the perturbation Hamiltonian $H'_j (j = 1, 2)$ acts on the j-th part only. The matrix element of 2nd order perturbation is proportional to

$$\sum_{m_1 m_2} \frac{1}{E_{i_1 i_2} - E_{m_1 m_2}} \langle f_1 f_2|(H'_1 + H'_2)|m_1 m_2\rangle \langle m_1 m_2|(H'_1 + H'_2)|i_1 i_2\rangle$$

$$= \left(\frac{1}{E_{i_1} - E_{f_1}} + \frac{1}{E_{i_2} - E_{f_2}} \right) \langle f_1|H'_1|i_1\rangle \langle f_2|H'_2|i_2\rangle.$$

So the transitional probability W of the 3-state protein is proportional to the product of those of the two partial 2-state proteins, W_1 from i_1 to f_1 and W_2 from i_2 to f_2 respectively. By using Equation (22) for two-state protein, namely

$$\ln W_i = \frac{\Delta G_i}{2k_B T} - \frac{(\Delta G_i/k_B T)^2}{2\rho N_i}$$

$$- 5.5\ln N_i + \ln c_0 f_i, \quad i = 1, 2, \quad (30)$$

neglecting the difference between structure-related shape parameters f_1 and f_2 in Equation (30) and denoting N and

ΔG of the 3-state protein as

$$N = N_1 + N_2, \quad N_1 = rN, \quad N_2 = (1-r)N;$$

$$\Delta G = \Delta G_1 + \Delta G_2, \quad \Delta G_1 = r\Delta G, \quad \Delta G_2 = (1-r)\Delta G,$$

we obtain

$$\ln W = \frac{\Delta G}{2k_B T} - \frac{(r\Delta G/k_B T)^2}{2\rho rN} - \frac{((1-r)\Delta G/k_B T)^2}{2\rho(1-r)N}$$

$$- 5.5\ln N - 5.5(\ln r + \ln(1-r)) + \ln f + const. \quad (31)$$

Equation (31) is the theoretical folding rate for any 3-state protein. The formula can easily be generalized to other m-state proteins with $m > 3$. The statistical comparisons of the theoretical rates $\ln W_f$ of multi-state protein folding with experimental rates $\ln k_f$ in a dataset of 38 multi-state proteins [33–67] are given in Figure 6. Since most of the multi-state proteins in databases are 3-state proteins we use Equation (31) to calculate $\ln W_f$ directly. It gives the average error MAE($\langle |\ln k_f - \ln W_f| \rangle$) = 1.68 and the correlation between theoretical and experimental rates $R = 0.8098$.

The success of the above calculation shows that a unified quantum folding mechanism does exist for multi-state and two-state protein and the folding rate of multi-state protein can be obtained in a simple way through calculating the product of the rates of two-state proteins.
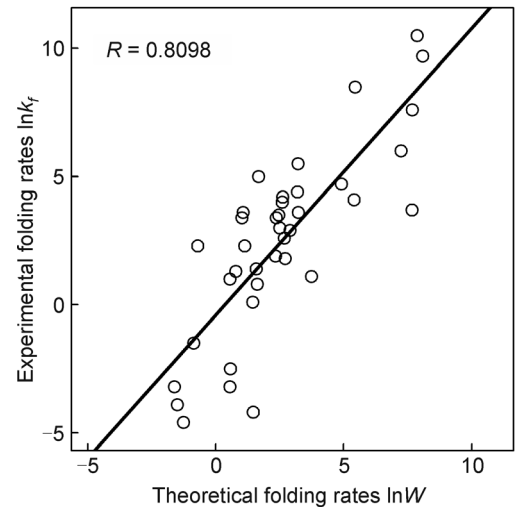
Figure 6. Comparison of theoretical folding rates $\ln W_f$ with experimental folding rates $\ln k_f$ for 38 multi-state proteins. Experimental rates for 38 proteins are taken from refs [33–67]. In calculating $\ln W_f$ the parameters $r = 0.35$ and $\rho = 0.07$ are assumed. Details can be found in Supplementary Table S2. The regression between $\ln k_f$ and $\ln W_f$ is plotted by the solid line and R means the correlation coefficient.

Protein photo-folding

To explore the fundamental physics underlying protein folding more deeply and to clarify the quantum nature of the folding mechanism more clearly we have studied the protein photo-folding processes, namely, the photon emission or absorption in protein folding and the inelastic scattering of photon on protein (photon-protein resonance Raman scattering). In these processes the emission or absorption of a photon by atomic electrons is coupled to protein's conformational change. To simplify the discussion we assume only electrons serve as the fast variables of the protein. After the first-principle-calculation based on quantum electrodynamics and quantum folding theory the rates and cross sections of these processes have been deduced [68]. Moreover, these photo-folding processes can be compared with common protein folding without interaction of photons (non-radiative folding). It is demonstrated that there exists a common factor (thermo-averaged overlap integral of vibration wave function) $I_V = \frac{1}{k_B T} I_V'$ for protein folding and protein photo-folding. Based on this finding it is predicted that the stimulated photo-folding rates and the resonance fluorescence cross section show the same temperature dependence as protein folding, namely $\frac{d \ln I_V}{d(1/T)} = S + \frac{1}{2}T + RT^2$.

Due to the coupling between protein structure and electron motion, the electronic transition will inevitably lead to the structural relaxation or conformational changes of the protein. Therefore the spectrum of protein photo-folding includes information of several kinds of quantum transitions: the electron energy-level transition, the transition between vibrational energy-levels of the molecule, the transition between rotational energy-levels of the molecule and the transition between different molecular conformations. Conformational transitions are somewhat like the rotational transitions, but the rotational transition refers to the whole molecule, while the conformational transition is related only to the dihedral angle rotation of local atomic groups. Because of the coupling with vibration and torsional transition the spectral line of electronic transition is broadened to a band which includes abundant vibration spectrum without and with conformational transition. The width of the spectral band is determined by the torsion vibration frequency, in the order of $10^{13}\ s^{-1}$, one hundredth or thousandth of the electronic transition frequency. Each spectral band includes a large amount of spectral lines and forms an abundant structure. The transition between torsion-vibration states in different conformations is a kind of forbidden transition since the overlap integral

between initial and final torsional wave functions is very small, about 10^{-5}. So, the width of these extra-narrow spectral lines is five orders smaller than the natural linewidth. This is an important prediction of quantum folding theory [68].

From the experimental point of view, to observe the extra-narrow spectral line a high-precision and high-resolution spectroscopy is needed. The spectral resolution of femtosecond Raman spectroscopy (FSRS) is 10 cm^{-1}, corresponding to $\Delta \nu = 3 \times 10^{11} s^{-1}$ [69]. This resolution is already close to the range of the width that the spectral line of ultra-narrow conformational transitions can be searched.

The particular form of the same temperature dependence for protein nonradiative folding and photo-folding and the abundant structure of the photo-folding spectral band consisting of many narrow lines are two main results deduced from protein photo-folding theory. These results are closely related to the fundamental concepts of quantum mechanics. First, they imply the existence of a set of quantum oscillators in the transition process and these oscillators are mainly of torsion vibration type of low frequency. Second, they imply in protein folding the quantum tunneling does exist which means the non-locality of state and the quantum coherence of conformational-electronic motion. More experimental tests on above two predictions are waited for.

Quantum coherence and experimental tests on quantum property of macromolecular conformational transition

So far we have discussed the quantum transition between torsion states in macromolecules. A fundamental problem is: due to quantum entanglement with the environment the decoherence possibly makes the quantum picture ceasing to be effective for a macromolecular system. It was estimated that the decoherence time $\tau_D = \tau_R \left(\frac{\hbar}{\Delta x \sqrt{2m k_B T}} \right)^2$ in a simple model where τ_R means the relaxation time, $\frac{\hbar}{\sqrt{2m k_B T}}$ the thermal de Broglie wavelength, Δx the separation of position and m the particle mass. This leads to the decoherence time inversely proportional to particle mass m [70]. For atoms or molecules in water the stochastic collisions with water molecules cause the decoherence of the objects. It was indicated that the decoherence time is lower than the dissipation time by a factor m_w/m (the mass ratio of water molecule to solute) [71]. However the above estimates hold only for the center-of-mass motion of particles. The protein or nucleic acid molecule is a compact aggregate of atoms. The coherence of the motion in internal degrees of

freedom may be preserved for a much longer time. How to estimate the coherence of the constituents which are bound in the macromolecule? Set the range of movement of the constituent denoted by d and the thermal de Broglie wavelength of the constituent denoted by λ. For electrons in hydrogen atom, $\lambda = 1$ nm and d = 0.05 nm, one has d/λ = 0.05 and the motion is coherent. For C atoms bounded in a simple organic molecule, $\lambda = 0.007$ nm and $d = 0.15$ nm (d calculated from the C-C or C = C bond length), one has $d/\lambda = 21$. Although the d/λ ratio in this case is much higher than the atom's electrons the quantum coherence of C atoms has been well established by the observed molecular vibrational spectrum. For torsion of atomic groups in protein, $\lambda = 0.01$ (inertial moment 10^{-37} gram cm^2 and room temperature are taken) and $d = 0.1$ (from the averaged angular shift between two minima of torsion potential), one has $d/\lambda = 10$. For N or C atoms bounded in RNA, $\lambda = 0.007$ nm and $d = 0.1$ nm, one has $d/\lambda = 14$. Since they have d/λ value smaller than the case of bounded C atoms in organic molecule, it is reasonable to assume the quantum coherence existing in the latter two cases. However, for C atoms freely moving in a macroscopic scale (say 1 cm) the ratio d/λ is about 10^9, much larger than above cases, and the quantum coherence is definitely destroyed. Therefore, the ratio d/λ can serve as a measure to determine the boundary between quantum and classical motions and by use of the measure one may recognize that the macromolecular conformational motion is basically a quantum event.

What should be cut away by Occam's Razor, classical or quantum, in studying the conformational change of biological macromolecules? What rules, classical or quantum, are obeyed by the macromolecular conformational motion? We believe and showed that the quantum mechanics provides a unifying and logically simple theoretical starting point. Moreover, as stated above, use of ratio d/λ as a criterion to observe the bound atoms in a molecule also supports the view of the quantum folding. Of course, the final solution of the problem still needs more direct experimental evidences.

Whether protein folding is quantum or classical can be directly tested by the observation of the instantaneous nature of the folding event. As is well known that the change of electronic state in atomic radiation is a kind of quantum transition and the transition is instantaneous. The instantaneousness is characteristic of the quantum transition. When one says the folding rate 1 ms it does not mean the folding continues one millisecond but rather means on average 1000 instantaneous folding events are observed stochastically in one second. So, the observation of the instantaneous change of the torsion angle in protein folding provides a clue to solve the puzzling problem.

Qiu *et al*. used laser temperature-jump spectroscopy to measure the folding rate of the 20-residue Trp-cage protein. They found the fluorescence intensity (FI) increasing rapidly from 11.5 mV to 14 mv in 4 μs and determined the folding rate 4 μs [19]. Which law, classical or quantum, the folding/unfolding event obeys in the duration of 4 μs? Let's consider the case in terms of the gradual decrease of relevant molecular concentrations. In the beginning, the fluorescence intensity will be weakened accompanying with the lowering of Trp-cage concentration,but the shape of the FI-t (fluorescence intensity versus time) curve remains unchanged. As the concentration decreased to very low, the single-molecule motion can be observed and the fluctuation appears. For quantum folding, the torsion takes only two possible values corresponding to folding and unfolding states respectively. The Trp fluorescence can be measured only in unfolding state. So, in the duration of 4 μs the fluorescence randomly appears and each occurrence corresponds to one unfolding event. However, if the folding/unfolding obeys the laws of classical physics the torsional angle changes continuously and the Trp fluorescence can be recorded only when the protein reaches the unfolding state. Thus the fluorescence will be measured near the end of 4 microsecond unfolding process. Two pictures are different from each other. We suggest making the observation of fluorescence fluctuation to test whether the folding obeys the classical or quantum law.

This is an important experiment. If the test has a positive result, then firstly, the boundary between classical and quantum physics will be modified and the applicability of quantum mechanics will be expanded to the internal degrees of freedom of macromolecules; secondly, many strange phenomena of sudden change that is of great importance in molecular biology and molecular genetics will be explained from the idea of quantum transition.

Applications of quantum folding theory

Example 1: Glucose transport across membrane. The binding of a ligand to a membrane receptor results in a conformational change, which then causes a specific programmed response. Recently, the crystal structures of several bacterial and human monosaccharide transporters were reported. Through sequential and structural comparison with other members of the sugar porter subfamily, the basic transport mechanism of the human glucose GLUT1 is clarified [72]. It was proposed that the successive conformational changes of the transporter occur in the glucose transport process and form a complete cycle, from ligand free occluded conformation (A), changed to outward open (B), ligand bound occluded (C), and

inward open (D), then to the ligand free occluded of the next cycle. The conformation A is connected to the intracellular side and the conformation C to the extracellular side. (Figure 7)
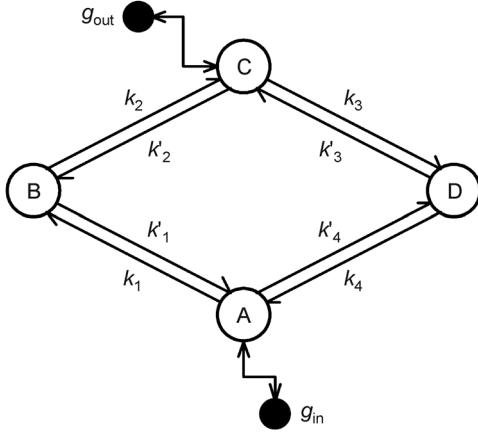
Figure 7. Quantum transition between conformational states of human glucose transporter GLUT1. A = ligand-free occluded, B = outward open, C = ligand-bound occluded, D = inward open, g_{out} and g_{in} denote the glucose molecule extracellular or intracellular respectively.

The human glucose transporter GLUT1 works in the following cycle:

(i) $A \leftrightarrow B$

(ii) $B + g_{out} \leftrightarrow C$

(iii) $C \leftrightarrow D$

(iv) $D \leftrightarrow A + g_{in}$

Following mass-action kinetics the reaction equations of the macromolecular concentration in these conformational states can be written. Through solution of the set of equations we deduced all steady states in the cycle and demonstrated that the glucose transport across membrane is globally stable [73].

The rate constants k_i, k_i' are key parameters in the theory that can be calculated from quantum theory of conformational transition. We notice that all rates

$$K_1 = k_1 = W(A \to B), \quad K_2 = k_2 v_0 = W(B \to C),$$

$$K_3 = k_3 = W(C \to D), \quad K_4 = k_4 = W(D \to A)$$

and their reverses

$$K_1' = k_1' = W(B \to A), \quad K_2' = k_2' = W(C \to B),$$

$$K_3' = k_3' = W(D \to C), \quad K_4' = k_4' u_0 = W(A \to D)$$

(v_0 and u_0 are steady state concentration of intracellular and extracellular glucose molecule respectively) are in the dimension of (1/time) and they can be calculated from quantum folding rate Equation (7).

The glucose molecules are hydrogen-bonded to several residues (for example, Gln, Asn, Trp) of the transporter GLUT1. Set the binding energy of one sugar molecule being E_b. The N- and C-domain of GLUT1 are connected by an intracellular helical bundle (ICH). The ICH domain serves as a latch that tightens the intracellular gate. The conformational change between inward-open (or outward-open) and occluded is related to the rigid-body rotation of the N and C domains which may be achieved through the bond-length and bong-angle variation of ICH residues [72]. It is reasonable to assume there exist two or more minima in the bond-length and bong-angle potential and the conformation change occurs through the quantum transition among them, jumping from one minimum to another. Set the conformational energy of the outward open (or inward open) relative to the occluded denoted by E_p (neglecting the difference between the outward open and inward open). Thus, we have an approximate estimate of the free energy difference $\Delta G^{(i)}$ for the i-th step quantum transition,

$$\Delta G^{(1)} = -E_p, \quad \Delta G^{(2)} = E_p + E_b,$$

$$\Delta G^{(3)} = -E_p, \quad \Delta G^{(4)} = E_p - E_b.$$

The reverse rate constants K_i' are related to K_i as $\ln\dfrac{K_i}{K_i'} \cong \dfrac{\Delta G^{(i)}}{k_B T}$ (see Equation (15)). Since both glucose binding energy E_b (typically several ev) and conformational energy E_p (typically 0.13 ev for one stretching-bending degree of freedom) are much larger than $k_B T$ (0.026 ev at room temperature) we have $K_2' \ll K_2$, $K_1' \gg K_1$, $K_3' \gg K_3$. It means the reverse transition from the conformation C to B can be neglected, while the reverse transitions from A to B and from C to D are important in the glucose transport cycle across membrane.

The total time needed for glucose transport across membrane in a cycle is called the mean transport time of glucose. We obtained the mean transport time for GLUT1

$$\tau = \frac{1}{K_1} + \frac{1}{K_2} + \frac{1}{K_3} + \frac{1}{K_4} \cong \frac{1}{K_2},$$

(for positive transport)

$$\tau' = \frac{1}{K_1'} + \frac{1}{K_2'} + \frac{1}{K_3'} + \frac{1}{K_4'} \cong \frac{1}{K_2'},$$
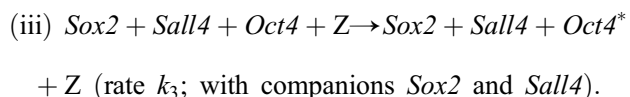
(for negative transport).

by the rate constant estimate [73]. Therefore, we

concluded that the glucose transport time in a cycle is determined mainly by the step of $B + g_{out} \leftrightarrow C$ (with rate K_2 or K_2'). Moreover, the positive transport time is smaller than the negative due to $K_2' \ll K_2$. This means the net transport of glucose is in the direction from extracellular to intracellular.

The rate of conformational transition for biological macromolecule is temperature dependent. Through direct measurement of the temperature relation of the mean transport time in a cycle we are able to give deeper insight into the mechanism for the glucose transmembrane transport. Simultaneously we expect other membrane receptors and membrane transport problems can be studied in the framework of the present theory.

Example 2: Induced pluripotency in stem cell
Recently Deng's group reported that the induced pluripotent stem cells (iPSC) can be created chemically without any gene modification [74]. They used a cocktail of seven small-molecule compounds to induce the mouse somatic cells into stem cells (called Chemically iPSC or CiPSC) with a higher efficiency up to 0.2%. There are seven pluripotency genes, namely *Oct4*, *Sox2*, *Nanog*, *Sall4*, *Sox17*, *Gata4* and *Gata6*, involved in the pluripotency circuitry during chemical reprogramming. The circuitry is divided into 3 sub-circuits. The first sub-circuit includes the action of small molecules CHIR(C), FSK(F), 616452(6) and DZNep(Z) on pluripotency genes *Oct4*, *Sox2* and *Sall4*. The small molecule action makes the gene to undergo a conformational transition from its torsion-ground to torsion-exited state (denoted by star * after the gene name)

(i) $C6F + Sox2 \rightarrow Sox2^* + C6F$ (rate k_1);

(ii) $C6F + Sall4 \rightarrow Sall4^* + C6F$ (rate k_2);

(iii) $Sox2 + Sall4 + Oct4 + Z \rightarrow Sox2 + Sall4 + Oct4^*$

$+ Z$ (rate k_3; with companions *Sox2* and *Sall4*).

The next 7 equations describe the torsion transfers (the conformational transitions under gene interaction) of the second and third sub-circuit. The second sub-circuit includes the interaction between pluripotency genes *Sall4*, *Oct4*, *Sox17*, *Gata4* and *Gata6*. The third sub-circuit includes the interaction between pluripotency genes *Oct4*, *Sox2* and *Nanog*. The gene interactions in these two circuits are generally in double direction. For gene interaction $A \leftrightarrow B$ we assume a conformational transition from torsion-ground to torsion-exited state of gene A (gene B) or its reverse in companion of gene B (gene A), namely,

$$A + B \leftrightarrow A^* + B, \quad A + B \leftrightarrow A + B^*$$
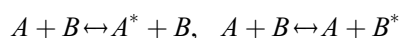
(details of 10 equations in the pluripotency circuitry can be found in [75]).

A set of reaction equations for pluripotency gene concentrations are deduced from mass-action kinetics. From these equations one obtains the steady states immediately and proves the stability of these steady states. Moreover, the relaxation times for attaining the steady states are deduced.

The rate constants in reaction equations describe the pluripotency transition in DNA. Since the pluripotency transition is essentially a torsion transition we use Equations (27) or (28) to estimate the rate constants. From Equation (28) we predict the transition rate decreases with increasing L as $L^{-D}(D \cong 5.5)$ at large L. The long-tail existing in the W-L curve can be used to explain why the pluripotency conversion of gene is a small-probability event since the gene sequence length is much longer than protein. On the other hand, we assume the torsion transition from differentiate to pluripotent state is of uphill-type (from torsion-ground to torsion excited) while the reverse is of downhill-type (from torsion-excited to torsion-ground). The parameter A (= 4.306) in Equation (27) is related to the free energy change ΔG in a dataset of two-state protein folding. However for DNA, ΔG has different symbols for positive downhill and negative uphill reactions. So, upon comparison with protein folding we predict the rate of DNA uphill transition is smaller than downhill by about a factor $e^{8.612} \approx 5500$. The torsion transfer between genes is generally a slow process but the reverse transition from torsion-excited to ground state in the circuitry is faster that provides a positive feedback mechanism to the pluripotency circuitry establishment.

From the rates of ten torsion transitions one can define three characteristic times τ_a, τ_b and τ_c corresponding to three sub-circuits. For example, $\tau_a = \dfrac{1}{k_1} + \dfrac{1}{k_2} + \dfrac{1}{k_3}$ for the first sub-circuitry. The time needed for CiPSC is estimated by $\tau = \tau_a + \tau_b + \tau_c$. Through numerical calculation we estimated $\tau_c \ll \tau_a$ and as the second circuit is switched off we have [75]

$$\tau \cong \tau_a \cong \frac{1}{k_2} = e^{30.31}\tau_{pt} = 27d/0.2\%,$$

which is in consistency with the experimental data of pluripotent stem cells generated at a frequency up to 0.2% on day 30–40 [74].

So far we have shown that the pluripotency conversion time in chemical reprogramming is calculable by using quantum folding theory. Now we examine how other physical factors influence the rate of the acquisition of pluripotency. As inferred from the theory, we have:

a) Transition rate depends on temperature
Assuming the free energy decrease ΔG in pluripotency

genes is linearly dependent of temperature T as in protein folding we obtain the temperature dependence of the reprogramming transition for the pluripotency gene as

$$\ln W(T) = \frac{S}{T} - RT + \frac{1}{2}\ln T + const.$$

b) Transition rate depends on pH

Following the biochemical principle the free energy change ΔG of a reaction is linearly dependent on the logarithm of ion concentration $[H^+]$. We have

$$\frac{d}{d\text{pH}}\ln W(\text{pH}) < 0, \quad (\text{for } \Delta G < 0).$$

When pH decreases from 7 to 6 (or to 5, to 4, …) the rate increases by a factor 3.16 (or 9.97, 31.5, …). It means one may observe the acidity-induced pluripotency by soaking the tissues in acidic medium below pH 6.0.

c) Transition rate depends on volume of the coherent domain

The conformational transitional rate is strongly dependent of the sequence length L of the gene if its nucleotides are coherent in the transition. Assuming the molecule is a compact aggregate having quantum coherence between its constituents, from Equation (28) one readily obtains

$$W(V) \approx fV^{-5.5}\exp\left(-\frac{a}{V}\right) \tag{32}$$

($a > 0$ is a V-independent constant) due to L proportional to the volume V of the gene. Equation (32) means $W(V)$ grows as V decreases (in the range $V > \frac{a}{5.5}$). The strong dependence of the transitional rate on coherent volume provides a regulation mechanism in the work of pluripotency circuitry. By deleting some unimportant redundant nucleotides in the gene and decreasing the sequence length one can lower down the volume of the coherent domain and increase the transitional rate.

SUMMARY

The quantum theory of conformation change of biomolecule is studied and a set of general equations on protein and RNA folding rate are proposed. From the basic equations the following quantitative results have been deduced: (i) the temperature dependence of the folding rate; (ii) the denaturant concentration dependence of the protein folding rate; (iii) the torsion mode number dependence of the protein folding rate; (iv) the chain length dependence of the RNA folding rate; (v) the folding rate of multi-state protein. Although these problems are difficult to study in classical theory they can be deduced easily from quantum equations and all quantitative results can be successfully tested by experimental data. Then, the quantum coherence of the internal

motion of a macromolecule is demonstrated and how to make the direct experimental test on the quantum property of the conformational transition are discussed. It is suggested that the spectrum structure of protein photofolding and the fluctuation of the fluorescence intensity emitted from folding/unfolding may serve as useful tools for verifying the quantum property of the conformation change. The idea of quantum folding is useful in dealing with practical application problems. Two examples show that the proposed idea and theory can help us to establish a set of reaction equations regarding the problem and determine the rate constants in the biological network.

SUPPLEMENTARY MATERIALS

ACKNOWLEDGEMENTS

COMPLIANCE WITH ETHICS GUIDELINES

The authors Liaofu Luo and Jun Lv declare that they have no conflict of interests.

This article does not contain any studies with human or animal subjects performed by any of the authors.

REFERENCES

1. Luo, L. F. (2014) Quantum theory on protein folding. Sci. China Phys. Mech. Astron., 57, 458–468

2. Luo, L. F. (2011) Protein Folding as a quantum transition between conformational states. Front. Phys., 6, 133–140

3. Lv, J. and Luo, L. (2014) Statistical analyses of protein folding rates from the view of quantum transition. Sci. China Life Sci., 57, 1197–1212

4. Luo, L. F. and Lv, J. (2015) Quantitative relations in protein and RNA folding deduced from quantum theory. bioRxiv: http://dx.doi.org/10.1101/021782

5. Luo, L. F. (2015) A model on avian genome evolution. bioRxiv: http://dx.doi.org/10.1101/034710; arXiv: 1411.2205, http://arxiv.org/abs/1411.2205

6. Hameroff, S. and Penrose, R. (2014) Consciousness in the universe: a review of the 'Orch OR' theory. Phys. Life Rev., 11, 39–78

7. Fisher, M. P. A. (2015) Quantum cognition: The possibility of processing with nuclear spins in the brain. Ann. Phys., 362, 593–602

8. Melkikh, A. V. (2014) Congenital programs of the behavior and nontrivial quantum effects in the neurons work. Biosystems, 119, 10–19

9. Gauger, E. M., Rieper, E., Morton, J. J. L., Benjamin, S. C. and Vedral, V. (2011) Sustained quantum coherence and entanglement in the avian compass. Phys. Rev. Lett., 106, 040503

10. Eyring, H., Lin, S. H. and Lin, M. (1980) Basic Chemical Kinetics.

New York: Wiley

11. Maxwell, K. L., Wildes, D., Zarrine-Afsar, A., De Los Rios, M. A., Brown, A. G., Friel, C. T., Hedberg, L., Horng, J. C., Bona, D., Miller, E. J., *et al.* (2005) Protein folding: defining a "standard" set of experimental conditions and a preliminary kinetic data set of two-state proteins. Protein Sci., 14, 602–616

12. Nguyen, H., Jager, M., Moretto, A., Gruebele, M. and Kelly, J. W. (2003) Tuning the free-energy landscape of a WW domain by temperature, mutation, and truncation. Proc. Natl. Acad. Sci. USA, 100, 3948–3953

13. Ghosh, K., Ozkan, S. B. and Dill, K. A. (2007) The ultimate speed limit to protein folding is conformational searching. J. Am. Chem. Soc., 129, 11920–11927

14. Dimitriadis, G., Drysdale, A., Myers, J. K., Arora, P., Radford, S. E., Oas, T. G. and Smith, D. A. (2004) Microsecond folding dynamics of the F13W G29A mutant of the B domain of staphylococcal protein A by laser-induced temperature jump. Proc. Natl. Acad. Sci. USA, 101, 3809–3814

15. Kuhlman, B., Luisi, D. L., Evans, P. A. and Raleigh, D. P. (1998) Global analysis of the effects of temperature and denaturant on the folding and unfolding kinetics of the N-terminal domain of the protein L9. J. Mol. Biol., 284, 1661–1670

16. Mayor, U., Johnson, C. M., Daggett, V. and Fersht, A. R. (2000) Protein folding and unfolding in microseconds to nanoseconds by experiment and simulation. Proc. Natl. Acad. Sci. USA, 97, 13518–13522

17. Manyusa, S. and Whitford, D. (1999) Defining folding and unfolding reactions of apocytochrome *b5* using equilibrium and kinetic fluorescence measurements. Biochemistry, 38, 9533–9540

18. Bunagan, M. R., Yang, X., Saven, J. G. and Gai, F. (2006) Ultrafast folding of a computationally designed Trp-cage mutant: Trp2-cage. J. Phys. Chem. B, 110, 3759–3763

19. Qiu, L., Pabit, S. A., Roitberg, A. E. and Hagen, S. J. (2002) Smaller and faster: the 20-residue Trp-cage protein folds in 4 micros. J. Am. Chem. Soc., 124, 12952–12953

20. Yang, W. Y. and Gruebele, M. (2004) Rate-temperature relationships in λ-repressor fragment λ 6-85 folding. Biochemistry, 43, 13018–13025

21. Jäger, M., Nguyen, H., Crane, J. C., Kelly, J. W. and Gruebele, M. (2001) The folding mechanism of a beta-sheet: the WW domain. J. Mol. Biol., 311, 373–393

22. Wang, T., Zhu, Y. J. and Gai, F. (2004) Folding of a three-helix bundle at the folding speed limit. J. Phys. Chem. B, 108, 3694–3697

23. Zhu, Y., Alonso, D. O., Maki, K., Huang, C. Y., Lahr, S. J., Daggett, V., Roder, H., DeGrado, W. F. and Gai, F. (2003) Ultrafast folding of α3D: a *de novo* designed three-helix bundle protein. Proc. Natl. Acad. Sci. USA, 100, 15486–15491

24. Spector, S. and Raleigh, D. P. (1999) Submillisecond folding of the peripheral subunit-binding domain. J. Mol. Biol., 293, 763–768

25. Uversky, V. N. (2013) Unusual biophysics of intrinsically disordered proteins. Biochim. Biophys. Acta, 1834, 932–951

26. Bonetti, D., Toto, A., Giri, R., Morrone, A., Sanfelice, D., Pastore, A., Temussi, P., Gianni, S. and Brunori, M. (2014) The kinetics of folding of frataxin. Phys. Chem. Chem. Phys., 16, 6391–6397

27. Garbuzynskiy, S. O., Ivankov, D. N., Bogatyreva, N. S. and Finkelstein, A. V. (2013) Golden triangle for folding rates of globular proteins. Proc. Natl. Acad. Sci. USA, 110, 147–150

28. Thirumalai, D. and Hyeon, C. (2005) RNA and protein folding: common themes and variations. Biochemistry, 44, 4957–4970

29. Woodson, S. A. (2010) Compact intermediates in RNA folding. Annu. Rev. Biophys., 39, 61–77

30. Hyeon, C. and Thirumalai, D. (2012) Chain length determines the folding rates of RNA. Biophys. J., 102, L11–L13

31. Kamagata, K., Arai, M. and Kuwajima, K. (2004) Unification of the folding mechanisms of non-two-state and two-state proteins. J. Mol. Biol., 339, 951–965

32. Zhang, Y. and Luo, L. (2011) The dynamical contact order: protein folding rate parameters based on quantum conformational transitions. Sci. China Life Sci., 54, 386–392

33. Cavagnero, S., Dyson, H. J. and Wright, P. E. (1999) Effect of H helix destabilizing mutations on the kinetic and equilibrium folding of apomyoglobin. J. Mol. Biol., 285, 269–282

34. Golbik, R., Zahn, R., Harding, S. E. and Fersht, A. R. (1998) Thermodynamic stability and folding of GroEL minichaperones. J. Mol. Biol., 276, 505–515

35. Banachewicz, W., Johnson, C. M. and Fersht, A. R. (2011) Folding of the Pit1 homeodomain near the speed limit. Proc. Natl. Acad. Sci. USA, 108, 569–573

36. Marianayagam, N. J., Khan, F., Male, L. and Jackson, S. E. (2002) Fast folding of a four-helical bundle protein. J. Am. Chem. Soc., 124, 9744–9750

37. Löw, C., Weininger, U., Zeeb, M., Zhang, W., Laue, E. D., Schmid, F. X. and Balbach, J. (2007) Folding mechanism of an ankyrin repeat protein: scaffold and active site formation of human CDK inhibitor p19 (INK4d). J. Mol. Biol., 373, 219–231

38. Calosci, N., Chi, C. N., Richter, B., Camilloni, C., Engström, A., Eklund, L., Travaglini-Allocatelli, C., Gianni, S., Vendruscolo, M. and Jemth, P. (2008) Comparison of successive transition states for folding reveals alternative early folding pathways of two homologous proteins. Proc. Natl. Acad. Sci. USA, 105, 19241–19246

39. Schreiber, G. and Fersht, A. R. (1993) The refolding of cis- and trans-peptidylprolyl isomers of barstar. Biochemistry, 32, 11195–11203

40. Burns, L. L., Dalessio, P. M. and Ropson, I. J. (1998) Folding mechanism of three structurally similar beta-sheet proteins. Proteins, 33, 107–118

41. Dalessio, P. M. and Ropson, I. J. (2000) Beta-sheet proteins with nearly identical structures have different folding intermediates. Biochemistry, 39, 860–871

42. Gianni, S., Guydosh, N. R., Khan, F., Caldas, T. D., Mayor, U., White, G. W., DeMarco, M. L., Daggett, V. and Fersht, A. R. (2003) Unifying features in protein-folding mechanisms. Proc. Natl. Acad. Sci. USA, 100, 13286–13291

43. Gianni, S., Calosci, N., Aelen, J. M., Vuister, G. W., Brunori, M. and Travaglini-Allocatelli, C. (2005) Kinetic folding mechanism of PDZ2 from PTP-BL. Protein Eng. Des. Sel., 18, 389–395

44. Calloni, G., Taddei, N., Plaxco, K. W., Ramponi, G., Stefani, M. and Chiti, F. (2003) Comparison of the folding processes of distantly related proteins. Importance of hydrophobic content in folding. J. Mol. Biol., 330, 577–591

45. Liu, C., Gaspar, J. A., Wong, H. J. and Meiering, E. M. (2002) Conserved and nonconserved features of the folding pathway of hisactophilin, a beta-trefoil protein. Protein Sci., 11, 669–679

46. Parker, M. J., Dempsey, C. E., Lorch, M. and Clarke, A. R. (1997) Acquisition of native beta-strand topology during the rapid collapse phase of protein folding. Biochemistry, 36, 13396–13405

47. Forsyth, W. R. and Matthews, C. R. (2002) Folding mechanism of indole-3-glycerol phosphate synthase from *Sulfolobus solfataricus*: a

test of the conservation of folding mechanisms hypothesis in (beta (alpha))(8) barrels. J. Mol. Biol., 320, 1119–1133

48. Maki, K., Cheng, H., Dolgikh, D. A., Shastry, M. C. and Roder, H. (2004) Early events during folding of wild-type staphylococcal nuclease and a single-tryptophan variant studied by ultrarapid mixing. J. Mol. Biol., 338, 383–400

49. Parker, M. J., Spencer, J., Jackson, G. S., Burston, S. G., Hosszu, L. L., Craven, C. J., Waltho, J. P. and Clarke, A. R. (1996) Domain behavior during the folding of a thermostable phosphoglycerate kinase. Biochemistry, 35, 15740–15752

50. Parker, M. J., Spencer, J. and Clarke, A. R. (1995) An integrated kinetic analysis of intermediates and transition states in protein folding reactions. J. Mol. Biol., 253, 771–786

51. Ogasahara, K. and Yutani, K. (1994) Unfolding-refolding kinetics of the tryptophan synthase alpha subunit by CD and fluorescence measurements. J. Mol. Biol., 236, 1227–1240

52. Jennings, P. A., Finn, B. E., Jones, B. E. and Matthews, C. R. (1993) A reexamination of the folding mechanism of dihydrofolate reductase from *Escherichia coli*: verification and refinement of a four-channel model. Biochemistry, 32, 3783–3789

53. Matouschek, A., Kellis, J. T. Jr, Serrano, L., Bycroft, M. and Fersht, A. R. (1990) Transient folding intermediates characterized by protein engineering. Nature, 346, 440–445

54. Schymkowitz, J. W., Rousseau, F., Irvine, L. R. and Itzhaki, L. S. (2000) The folding pathway of the cell-cycle regulatory protein p13suc1: clues for the mechanism of domain swapping. Structure, 8, 89–100

55. Teilum, K., Thormann, T., Caterer, N. R., Poulsen, H. I., Jensen, P. H., Knudsen, J., Kragelund, B. B. and Poulsen, F. M. (2005) Different secondary structure elements as scaffolds for protein folding transition states of two homologous four-helix bundles. Proteins, 59, 80–90

56. Fowler, S. B. and Clarke, J. (2001) Mapping the folding pathway of an immunoglobulin domain: structural detail from Phi value analysis and movement of the transition state. Structure, 9, 355–366

57. Cota, E. and Clarke, J. (2000) Folding of beta-sandwich proteins: three-state transition of a fibronectin type III module. Protein Sci., 9, 112–120

58. Jemth, P., Day, R., Gianni, S., Khan, F., Allen, M., Daggett, V. and Fersht, A. R. (2005) The structure of the major transition state for folding of an FF domain from experiment and simulation. J. Mol. Biol., 350, 363–378

59. Melnik, B. S., Marchenkov, V. V., Evdokimov, S. R., Samatova, E. N. and Kotova, N. V. (2008) Multy-state protein: determination of carbonic anhydrase free-energy landscape. Biochem. Biophys. Res. Commun., 369, 701–706

60. Tang, K. S., Guralnick, B. J., Wang, W. K., Fersht, A. R. and Itzhaki, L. S. (1999) Stability and folding of the tumour suppressor protein p16. J. Mol. Biol., 285, 1869–1886

61. Laurents, D. V., Corrales, S., Elías-Arnanz, M., Sevilla, P., Rico, M. and Padmanabhan, S. (2000) Folding kinetics of phage 434 Cro protein. Biochemistry, 39, 13963–13973

62. Parker, M. J. and Marqusee, S. (1999) The cooperativity of burst phase reactions explored. J. Mol. Biol., 293, 1195–1210

63. Lowe, A. R. and Itzhaki, L. S. (2007) Rational redesign of the folding pathway of a modular protein. Proc. Natl. Acad. Sci. USA, 104, 2679–2684

64. Choe, S. E., Matsudaira, P. T., Osterhout, J., Wagner, G. and Shakhnovich, E. I. (1998) Folding kinetics of villin 14T, a protein domain with a central beta-sheet and two hydrophobic cores. Biochemistry, 37, 14508–14518

65. Muñoz, V., Lopez, E. M., Jager, M. and Serrano, L. (1994) Kinetic characterization of the chemotactic protein from *Escherichia coli*, CheY. kinetic analysis of the inverse hydrophobic effect. Biochemistry, 33, 5858–5866

66. Stagg, L., Samiotakis, A., Homouz, D., Cheung, M. S. and Wittung-Stafshede, P. (2010) Residue-specific analysis of frustration in the folding landscape of repeat beta/alpha protein apoflavodoxin. J. Mol. Biol., 396, 75–89

67. Ratcliff, K., Corn, J. and Marqusee, S. (2009) Structure, stability, and folding of ribonuclease H1 from the moderately thermophilic *Chlorobium tepidum*: comparison with thermophilic and mesophilic homologues. Biochemistry, 48, 5890–5898

68. Luo, L. (2012) Protein photo-folding and quantum folding theory. Sci. China Life Sci., 55, 533–541

69. Fang, C., Frontiera, R. R., Tran, R., Mathies, R. A. and Mathies, R. A. (2009) Mapping GFP structure evolution during proton transfer with femtosecond Raman spectroscopy. Nature, 462, 200–204

70. Zurek, W. H. (2002) Decoherence and the transition from quantum to classical − Revisited. In Quantum Decoherence. 16–31. Birkhäuser Basel, Doi:10.1007/978-3-7643-7808-0_1

71. Tegmark, M. (2000) Importance of quantum decoherence in brain processes. Phys. Rev. E Stat. Phys. Plasmas Fluids Relat. Interdiscip. Topics, 61, 4194–4206

72. Deng, D., Xu, C., Sun, P., Wu, J., Yan, C., Hu, M. and Yan, N. (2014) Crystal structure of the human glucose transporter GLUT1. Nature, 510, 121–125

73. Luo, L. F. (2014) Quantum theory on glucose transport across membrane. arXiv: 1407.7198 at http://arxiv.org/abs/1407.7198

74. Hou, P., Li, Y., Zhang, X., Liu, C., Guan, J., Li, H., Zhao, T., Ye, J., Yang, W., Liu, K., *et al.* (2013) Pluripotent stem cells induced from mouse somatic cells by small-molecule compounds. Science, 341, 651–654

75. Luo, L. F. (2015) From chemically to physically induced pluripotency in stem cell. arXiv: 1506.02053 at http://arxiv.org/abs/1506.02053