Physica A 493 (2018) 286–300

Contents lists available at ScienceDirect

Physica A

journal homepage: www.elsevier.com/locate/physa

Supersymmetric quantum mechanics method for the Fokker–Planck equation with applications to protein folding dynamics

Franciele Polotto [a], Elso Drigo Filho [a], Jorge Chahine [a], Ronaldo Junio de Oliveira [b,*]

[a] Departamento de Física, Instituto de Biociências, Letras e Ciências Exatas, Universidade Estadual Paulista "Júlio de Mesquita Filho", Rua Cristóvão Colombo, 2265, 15054-000, São José do Rio Preto, SP, Brazil

[b] Laboratório de Biofísica Teórica, Departamento de Física, Instituto de Ciências Exatas, Naturais e Educação, Universidade Federal do Triângulo Mineiro, Av. Dr. Randolfo Borges Junior, 1400, 38064-200, Uberaba, MG, Brazil

HIGHLIGHTS

- The Fokker–Planck equation is mapped onto a Schrödinger-type equation.
- It is solved for the double well potential with supersymmetric quantum mechanics.
- Time-dependent probability distributions are obtained with the variational method.
- Kinetics of the cold shock protein is characterized with a coarse-grained C_α model.

ARTICLE INFO

Article history:
Received 18 May 2017
Received in revised form 19 October 2017
Available online 10 November 2017

Keywords:
Quantum mechanics
Schrödinger equation
Folding rates
Structure-based model

ABSTRACT

This work developed analytical methods to explore the kinetics of the time-dependent probability distributions over thermodynamic free energy profiles of protein folding and compared the results with simulation. The Fokker–Planck equation is mapped onto a Schrödinger-type equation due to the well-known solutions of the latter. Through a semi-analytical description, the supersymmetric quantum mechanics formalism is invoked and the time-dependent probability distributions are obtained with numerical calculations by using the variational method. A coarse-grained structure-based model of the two-state protein *Tm*CSP was simulated at a C_α level of resolution and the thermodynamics and kinetics were fully characterized. Analytical solutions from non-equilibrium conditions were obtained with the simulated double-well free energy potential and kinetic folding times were calculated. It was found that analytical folding time as a function of temperature agrees, quantitatively, with simulations and experiments from the literature of *Tm*CSP having the well-known 'U' shape of the Chevron Plots. The simple analytical model developed in this study has a potential to be used by theoreticians and experimentalists willing to explore, quantitatively, rates and the kinetic behavior of their system by informing the thermally activated barrier. The theory developed describes a stochastic process and, therefore, can be applied to a variety of biological as well as condensed-phase two-state systems.

* Corresponding author.
E-mail address: ronaldo.oliveira@uftm.edu.br (R.J. Oliveira).

https://doi.org/10.1016/j.physa.2017.10.021
0378-4371/

1. Introduction

The Fokker–Planck (FP) equation is one of the basic tools suitable to manage fluctuations in a large range of dynamic systems [1,2]. It has been widely applied in diverse and, on some occasions, non-correlated areas such as physics, chemistry, biology, finance, economics, forecasting among many others [3–6]. Several branches of physics make use of the FP equation to describe and to understand their underlying complex dynamics, for instance, quantum mechanics, astrophysics, statistical physics [7–10], and, specifically in this work, it is applied to the protein folding biophysics problem [11–16].

Solutions of the FP equation provide the time-dependent probability density distributions to find a particle for a given position and velocity and a variety of methods to obtain exact or approximate solutions have been developed over the last few decades [1,2,17,18]. Specifically regarding this study, statistical physics has borrowed from quantum mechanics theoretical methods such as supersymmetry, originally developed to study particle physics, in order to solve the FP equation [19–31]. By using appropriate tools, one can transform the FP equation into a Schrödinger-type equation and apply the supersymmetric quantum mechanics formalism to solve the resulting equation. Recently, this resulting equation was solved analytically for both the generalized Morse potential and the Hulthén potentials [32].

The variational method allied with supersymmetric quantum mechanics is used in order to analyze the Schrödinger-type equation. This approach has been successfully used in the study of quantum potentials [25,33–38]. In this context, the formalism was first discussed in Ref. [33] and applied to analyze the Morse potential in Ref. [25], in an independent way. Subsequently, this approach was extended for different kind of problems such as confined quantum systems [34,35,38]. It is possible to use this approach to study stochastic problems by taking advantage of the relationship between the Schrödinger and the Fokker–Planck equations. The viability of this treatment was tested by the study of Fokker–Planck equation with a bistable potential [38].

Bistable double-well potentials are very commonly employed to study many biophysical processes [39], as is the case of protein folding dynamics [40–42], since the reaction rate theory proposed by the Kramers' seminal paper in the 40's [43]. In order to numerically solve the corresponding Schrödinger differential equation with a one-dimensional bistable potential, one choice would be the variational method [25,29,31,44]. The eigenvalues and eigenfunctions can be used to calculate the time-dependent transition probability distributions and, as a consequence, the biologically relevant relaxation rates and first passage times [22,24,30,31,39,45–47].

There are many studies of protein folding where the reaction is described by a one-dimensional diffusion process. Although the one-dimensional diffusion view of the folding process may be considered a simplified description of the folding reaction, several other studies, in which the folding reaction was described by the projection of the multidimensional energy landscape onto the one-dimensional reaction coordinate, have contributed to enhance our understanding of the protein folding dynamics [48,49]. In this regard, the energy landscape theory [49,50] has shown to be a useful framework to correlate theoretical and experimental data, such folding rates [51] and transition state studies [52]. Recent papers [53–55] describe experiments where folding trajectories of molecules were held in optical tweezers in order to compare the conditional probability of being on a transition path, obtained from the trajectory, with the prediction for ideal one-dimensional diffusion. The authors found good agreement between the experimental data and the theoretical prediction.

Protein folding, being a diffusive process, can be an application of the FP stochastic time evolution equation. Folding of small globular two-state proteins are reasonably described by the concepts of a one-dimensional diffusive process in a double-well potential along the most relevant reaction coordinate or order parameter [56–58]. The protein folding thermodynamics and kinetics can be described by the most accepted energy landscape theory [59–62]. According to this theory, an ensemble of unfolded protein conformations with high energy and entropy diffuses through many pathways en route to the lowest energy native state [57]. The protein energy landscape is intrinsically multidimensional and resembles a funnel of structures partially folded with a kinetic and thermodynamic bottleneck narrowed at the transition state lying between the folded/unfolded ensembles [51] (see Fig. 1). Fig. 1 shows the folding funnel landscape of the cold-shock protein *Tm*CSP, similar to the density of states of the previous study [63], which will be further and better explored within this study. Understanding the mechanism by which a protein folds to its native structure is crucial to create new treatments to prevent many pathological conditions, including Alzheimer's, Parkinson's, spongiform encephalopathies, type II diabetes and many others disorders [64].

Although a more general understanding of how a protein folds has emerged from many experimental and theoretical studies, a general principle that would explain rate mechanisms applicable to a broad range of proteins different in size and fold motifs is still a puzzle to solve [13]. Folding rates (k_f) or its inverse, the folding time (τ_f), spans about 6 orders of magnitude (from microseconds to seconds) for proteins with around 100 residues of amino acid [51,65]. Such diversity of scenarios demands more quantitative theoretical methods to predict folding rates that could shed light on the specific factors shaping the protein motion along the energy landscape funnel.

Folding times of many two-state proteins produces exponential distributions. Thus, the multidimensional energy landscape can be projected onto an effective one-dimensional free energy profile along a reaction coordinate with the folded and unfolded states, similar to a double-well separated by the energy barrier [42]. Based on the energy landscape theory, experimental and theoretical frameworks are able to determine the protein free energy profile and rates, at least for small globular proteins [58,66–71]. Recently, it was demonstrated that folding rates strongly correlate with folding and binding funnel energy landscape topography parameters associated with each particular protein: native to the transition state energy gap, landscape roughness and configurational entropy [42,72]. In this manuscript, an analytical approach, based

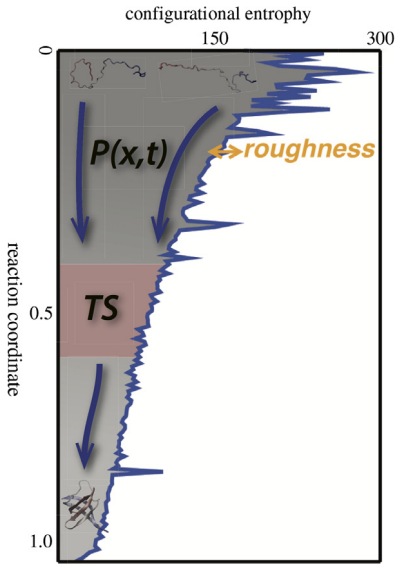

Fig. 1. The folding funnel landscape of the protein *Tm*CSP is represented by the configurational entropy (S) as a function of the reaction coordinate (x) of the system: $S = log\,[n\,(x)]$, where n is the density of states. High entropic ensemble of configurations at the top of the funnel diffuses through many pathways crossing the transition state (TS) ensemble in which a few are selected en route to the native ensemble at the bottom. Integration of the time-dependent probability distribution of an ensemble of configurations in x ($P(x, t)$) for a given effective potential ($F(x)$) along the folding funnel results in the protein analytical folding time (τ_f).

on quantum mechanics supersymmetric theory, is developed to connect the biophysical experimental/computational free energy transition barrier to the theoretically predicted folding times of proteins.

In this article, the thermodynamic free energy profile is calculated via alpha carbon (C_α) coarse-grained simulations with the structure-based model [52,73]. Kinetic folding times obtained with computer simulations were recovered with the analytical approach, given the protein free energy profile. The two-state cold-shock protein *Tm*CSP [74–77] was chosen for the simulations. The study is presented in the following way. In the first section, the theory necessary to calculate rates analytically is developed. The Fokker–Planck equation is modified to resemble the Schrödinger equation and the supersymmetric quantum mechanics approach is applied to solve the resulting differential diffusion equation. The variational method [29] is used to solve the Schrödinger-type equation for the bistable potential, and to compute time-evolution distribution functions of the system for a range of relevant temperatures. The second section presents the simulation details and the thermodynamic and kinetic analysis employed in the *Tm*CSP hyperthermophilic protein. The last section is devoted to showing the comparison between the analytical and the computational folding times (τ_f), which is related to the mean first passage time (mfpt) and the reaction rate theory [16,39,45,46]. The experimental 'U' shape behavior for the kinetic folding times and rates were also obtained from the analytical theory and simulations. These results provide the bridge between theory and experiment. The theory could be extended to other proteins and rate mechanisms where one has information about the thermodynamic free energy barrier.

2. Theory

2.1. The Fokker–Planck equation associated to a Schrödinger-type equation

The FP equation can be studied by converting it to a Schrödinger-type equation [2]. Following this approach, the FP equation is given by

$$\frac{\partial}{\partial t}P\,(x,\,t) = -\frac{\partial}{\partial x}\left[f\,(x)\,P\,(x,\,t)\right] + \frac{\Gamma}{2}\frac{\partial^2}{\partial x^2}P\,(x,\,t) \tag{1}$$

where $f(x)$ is a function which resembles the "force" and Γ is a constant related to the diffusion coefficient. The solution of Eq. (1) gives the time evolution of the probability distribution and can be written as [1]

$$P\,(x,\,t) = \psi_0\,(x)\sum_{l=0}^{\infty} a_l\psi_l\,(x)\,e^{-t|\Lambda_l|} \tag{2}$$

with the a_l coefficients inside the sum given by $a_l = \frac{\psi_l(x_0)}{\psi_0(x_0)}$. The time-dependent probability of Eq. (2) is subjected to the initial condition x_0 where the system starts its diffusive motion and, therefore, defines the conditional probability $P\,(x,\,t|x_0)$ [39].

x_0 will be omitted for clarity of the following equations, yet its value used in the numerical calculations of P will be informed when necessary. The eigenfunctions ψ_l and the eigenvalues Λ_l represent the solutions of the Schrödinger-type equation [18] obtained by substituting Eq. (2) in (1) and results in

$$-\frac{\Gamma}{2}\frac{\partial^2\psi_l}{\partial x^2}+\frac{1}{2}\left\{\frac{\partial f(x)}{\partial x}+\frac{[f(x)]^2}{\Gamma}\right\}\psi_l=\Lambda_l\psi_l. \tag{3}$$

Supersymmetric methods in quantum mechanics and statistical physics allow one to solve the Schrödinger-type Eq. (3) [78,79]. Using the supersymmetric formalism, a given Hamiltonian H_+ can be factored in terms of bosonic operators $a^{\pm}$

$$H_+=-\frac{\hbar^2}{2m}\frac{d^2}{dx^2}+V(x)=a^+a^-+E_{+,0} \tag{4}$$

where $a^{\pm}=\mp\frac{d}{dx}+W(x)$, $W(x)$ is the superpotential, $V(x)$ is the potential of the system investigated and $E_{+,0}$ is the eigenvalue of the ground state. Eq. (4) results in a Riccati equation [80], the solution of which gives $E_{+,0}$. It is possible to obtain the ground state wave function once the superpotential is determined.

The algebraic structure of the supersymmetry formalism gives a hierarchy of Hamiltonians related to each other [78,79]. It is possible to determine the relationship between eigenvalues and eigenfunctions of the hierarchical members by associating them with the original Hamiltonian H_+. This method can be applied to any potential that has an exact solution.

Unfortunately, the Schrödinger equation only has an exact solution for rather a few potentials and even the eigenfunction expansion method is quite restricted [2]. Alternatively, Schrödinger equations with potentials that do not have exact analytical solutions can be solved numerically by approximated methods such as the variational method [25]. The variational method is employed to estimate a good test function, written in terms of a set of parameters that minimizes the energy. The variational method can be written as

$$\frac{\int\psi_\mu^*(x)\,H\psi_\mu(x)\,dx}{\int\left|\psi_\mu(x)\right|^2dx}\geq E \tag{5}$$

where H is the Hamiltonian operator of the problem, E is the energy eigenvalue of the state in study and $\psi_\mu(x)$ is the test function to be chosen in terms of a set of μ parameters. The eigenfunctions obtained by applying the supersymmetric formalism can be used as the test functions. The closer the test function is to the real wave function, the better are the results for E. This method described to solve the one-dimensional FP equation by exploiting the correspondence between an arbitrary potential and the supersymmetric quantum mechanics formalism can be used to study, for example, a particle diffusing in a viscous medium [19]. Trial wave functions for the variational method and the construction of the effective Hamiltonians hierarchy can be used to obtain eigenfunctions and energies of the excited states, necessary to compute the probability distributions densities [29,30]. The supersymmetry theory can be applied to investigate the kinetics mechanisms of the protein folding process, which, apropos, have a strong resemblance to the diffusing particle problem [12,66,70].

Within the protein folding scenario, $f(x)$ in Eq. (1) is rewritten as

$$f(x)=-\frac{1}{\gamma}\frac{\partial}{\partial x}F(x) \tag{6}$$

where $F(x)$ is the free energy as a function of the one-dimensional reaction coordinate (x) and γ is the friction coefficient. By putting Eq. (6) into Eq. (1), it results in

$$\frac{\partial}{\partial t}P(x,t)=-\frac{\partial}{\partial x}\left[\left(-\frac{1}{\gamma}\frac{\partial}{\partial x}F(x)\right)P(x,t)\right]+D\frac{\partial^2}{\partial x^2}P(x,t) \tag{7}$$

where D is the diffusion coefficient, which is related to the constant $\frac{\Gamma}{2}$ in Eq. (1) and x describes the reaction coordinate, which is the number of native contacts (normalized to one, or the fraction) formed during the folding reaction. The domain of x, which is between 0 and 1, is shifted to the left, so that the barrier is at $x=0$. That is the reason to have values, for example, between -0.45 and 0.55. Eq. (7) is the diffusion equation used to describe the time evolution of the probability distribution $(P(x,t))$ of a protein for a given x diffusing over a free energy profile $F(x)$ [57,66,81–84]. Eq. (7) results in the stationary solution of the Fokker–Planck equation given by

$$P(x)=Nexp\left\{-\frac{F(x)}{\gamma D}\right\} \tag{8}$$

where N is the normalization constant. By using the Einstein relation $\gamma D\sim kT$ [7], the probability distribution is rewritten as

$$P(x)=Nexp\{-\beta F(x)\} \tag{9}$$

where $\beta=\frac{1}{kT}$, k is the Boltzmann constant and T is the temperature of the system.

In the Schrödinger-type Eq. (3), $f(x)$ is related to an effective potential, the free energy, and a suitable choice can be

$$V_{ef}(x) = \frac{D}{2}\left\{\frac{\beta^2}{2}\left(-\frac{\partial}{\partial x}F(x)\right)^2 - \beta\frac{\partial^2}{\partial x^2}F(x)\right\}. \tag{10}$$

The solution of the FP equation can be obtained from the associated Schrödinger-type equation. In order to study the protein folding process, effective potentials of the type of Eq. (10) are used in the Schrödinger-type equation to obtain the eigenfunctions $\psi_l(x)$ and the eigenvalues Λ_l, which compose the probability distribution $P(x, t)$ in Eq. (2), i.e., the solution of the FP equation. Thus, the Schrödinger-type Eq. (3) with the effective potential given by Eq. (10) becomes

$$-D\frac{\partial^2}{\partial x^2}\psi_l(x) + V_{ef}(x)\,\psi_l(x) = \Lambda_l\psi_l(x). \tag{11}$$

The formalism of the supersymmetric quantum mechanics will be applied to solve Eq. (11). The solution is obtained state by state by invoking the Hamiltonian hierarchy [78,79].

2.2. Bistable potential solution of the Schrödinger-type equation with the supersymmetric quantum mechanics and the variational method

The solution of the FP equation is obtained in this manuscript by using the supersymmetric quantum mechanics formalism along with the variational method. This approach determines approximated wave function and energy eigenvalues of the system. In order to do that, the double-well protein free energy profile will be given by the polynomial

$$F(x) = ax^4 + bx^3 + cx^2 \tag{12}$$

where a, b and c are constants. The free energy shown in Eq. (12) was chosen to display a bistable function and to fit appropriately the free energy profile calculated via computer simulations. $F(x)$ is the free energy function and bistable potentials are broadly reported to describe many two-state folders [51,57–62,64–66]. The shape of the double-well potential depends of the values chosen for the constants of Eq. (12). Constant a inverts the two minima to a two maxima potential and constant c changes the wells depth. Considering $a < 0$ for the double-well potential and assuming that Eq. (12) is centered at $x = 0$, if $b > 0$, the deepest minimum appears for negative values of x; otherwise ($b < 0$), the asymmetry implies a more pronounced minimum for positive values of x.

Substituting Eq. (12) into Eq. (10), the effective potential is written as

$$\frac{V_{ef}(x)}{D} = \beta^2\left\{4a^2x^6 + 6abx^5 + \left(4ac + \frac{9b^2}{4}\right)x^4 + 3bcx^3 + \left(c^2 - \frac{6a}{\beta}\right)x^2 - \frac{3b}{\beta}x - \frac{c}{\beta}\right\} \tag{13}$$

where $\beta = \frac{1}{kT}$ and k is the Boltzmann constant. From the Schrödinger-type Eq. (11), the initial Hamiltonian with the effective potential is given by

$$H_{+,1} = -\frac{d^2}{dx^2} + \beta^2\left\{4a^2x^6 + 6abx^5 + \left(4ac + \frac{9b^2}{4}\right)x^4 + 3bcx^3 + \left(c^2 - \frac{6a}{\beta}\right)x^2 - \frac{3b}{\beta}x - \frac{c}{\beta}\right\}. \tag{14}$$

The initial Hamiltonian can be factored in the terms of

$$H_{+,1} = -\frac{d^2}{dx^2} + W_1^2(x) - W_1'(x) + E_0^{(1)}. \tag{15}$$

In order to guarantee the equality between Eqs. (14) and (15), it is sufficient to write the superpotential $W_1(x)$ as

$$W_1(x) = \beta\left(2ax^3 + \frac{3}{2}bx^2 + cx\right) \tag{16}$$

with the lowest energy eigenvalue being

$$E_0^{(1)} = 0. \tag{17}$$

The ground state wave function of the superpotential given by Eq. (16) can be written as

$$\psi_0^{(1)}(x) = exp\left\{-\frac{\beta}{2}\left(ax^4 + bx^3 + cx^2\right)\right\}. \tag{18}$$

This case has an analytical solution for the ground state. However, the supersymmetric partner of $H_{+,1}$ should be determined to obtain the solution of the Schrödinger-type equation for the first excited state

$$H_{-,1} = -\frac{d^2}{dx^2} + W_1'(x) + W_1^2(x) + E_0^{(1)} \tag{19}$$

or in terms of the new operators

$$H_{+,2} = a_2^+ a_2^- + E_0^{(2)} = -\frac{d^2}{dx^2} - W_2'(x) + W_2^2(x) + E_0^{(2)}. \tag{20}$$

The equality between Eqs. (19) and (20) requires that $W_2(x)$ must satisfy the following equation

$$\beta^2 \left\{ 4a^2x^6 + 6abx^5 + \left(4ac + \frac{9b^2}{4}\right)x^4 + 3bcx^3 + \left(c^2 + \frac{6a}{\beta}\right)x^2 + \frac{3b}{\beta}x + \frac{c}{\beta} \right\} = W_2^2(x) - W_2'(x) + E_0^{(2)}. \tag{21}$$

Unfortunately, it is not possible to determine the exact solution of this equation. Therefore, a modified superpotential is suggested of the type

$$W_2(x) = \alpha_2 x^3 + \mu_2 x^2 + k_2 x + \eta_2 \tag{22}$$

where α_2, μ_2, k_2 and η_2 are the parameters to be determined by the variational method. The ground state wave function of the superpotential given by Eq. (22) is written as

$$\psi_0^{(2)}(x) = exp\left\{ -\alpha_2 \frac{x^4}{4} - \mu_2 \frac{x^3}{3} - k_2 \frac{x^2}{2} - \eta_2 x \right\}. \tag{23}$$

It is sufficient to apply the bosonic operator $a_1^+ = -\frac{d}{dx} + W_1(x)$ in $\psi_0^{(2)}$ to obtain the eigenfunction of the first excited state $\psi_1^{(1)}$ [78]. The eigenfunction, the original Hamiltonian H_0 and the variational method are used in Eq. (5) in order to calculate the energy eigenvalues for the first excited state with the four variational parameters α_2, μ_2, k_2 e η_2.

By repeating this process, it is possible to build a hierarchy of Hamiltonians and eigenvalues for higher excited states. This procedure requires superpotentials that generate test wave functions for each excited state. The suggested superpotentials adopt the same form of Eq. (22) with different variational parameters α_n, μ_n, k_n e η_n. These wave functions are used by the variational method to obtain the energy eigenvalues [29].

3. Simulation and numerical details

The protein used in this study was the cold-shock protein from the hyperthermophilic bacterium *Thermotoga maritima* (*Tm*CSP Protein Data Bank entry 1G6P [85]), with its three dimensional structure shown in Figs. 1 and 2. *Tm*CSP is a small globular 66-amino-acid β-barrel protein that is known to have well-defined two-state folding behavior by experiments [74–77] and theory [42,63,66,67,86]. The simulations were executed with the widely used structure-based model at the protein-coarse-grained-in–alpha-carbon (C_α-SBM) level of resolution [52,73]. The simulation inputs were generated via the SMOG webserver [87] with the shadow map algorithm to create the contact map list [88]. The functional form of the simulated potential energy and protocol used to perform kinetics (mean first passage times or folding times) and thermodynamics (free energy profiles) calculations were accomplished in the same way as in previous studies except for the shadow map contact list and the latest SMOG version 2.0 [42,63,66,67,86]. The reaction coordinate (x), chosen to monitor folding/unfolding events, was that commonly employed for protein folding studies i.e. the native contacts [89] formed in a given structure normalized between −0.5 and 0.5, which are the numerical boundaries for the analytical approach. Computational folding times (τ_f) were obtained averaging over 1000 folding runs with *Tm*CSP initiating the simulations from completely open random conformations and stopping when it first reached the native state at $x_{folded} = +0.4$ (examples of opened and closed structures are shown in Figs. 1 and 2). Free energy profiles were calculated with the Weighted Histogram Analysis Method (WHAM) [90] over 24 runs of different temperatures spanning from lower temperatures (fully folded states) to higher temperatures (unfolded states), including the transition temperature from folded to unfolded state (T_f). Each run was performed long enough, 1×10^9 steps (with time step $\Delta t = 0.5$ fs), which is 3 orders of magnitude higher than one folding event at T_f, to ensure proper accurate sampling. The configurational entropy (S), was calculated using the logarithm of the density of states as a function of the reaction coordinate ($S = log\,[n(Q)]$) obtained with the WHAM algorithm.

Numerical integrations were executed by using the software Wolfram Mathematica® licensed to Universidad de Valladolid.

4. Results and discussions

4.1. Temporal evolution of probability distributions and folding transition times

It is assumed that the polynomial Eq. (12) can fit the thermodynamic free energy profile ($F(x)$) of the protein simulation data. The Schrödinger-type equation will be solved analytically with this double-well potential for the ground state and higher excited states with the supersymmetric quantum mechanics theory and the variational method. Then, the temporal evolution of the probability distributions ($P(x, t)$, Eq. (2)) will be integrated to obtain the kinetic protein folding transition times. Fig. 1 presents the probability distribution ($P(x, t)$ diffusing along the *Tm*CSP folding landscape with the

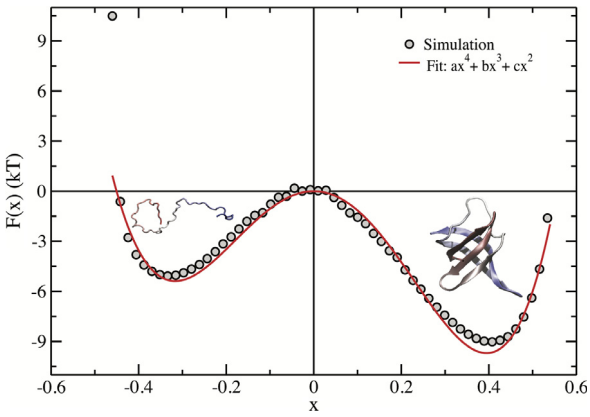

Fig. 2. Thermodynamic free energy profile ($F(x)$) as function of the reaction coordinate native contacts (x). Filled circles represent $F(x)$ obtained with C_α-SBM simulations and the continuum curve is the polynomial fit ($F(x) = ax^4 + bx^3 + cx^2$) over the simulation data with parameters $a = 460$, $b = -48$ and $c = -115$. It also shows the unfolded ($x_{unfolded} \sim -0.4$) and folded ($x_{folded} \sim +0.4$) structures of *Tm*CSP (Protein Data Bank entry 1G6P). $F(x)$ is shown at temperature $T = 1.080$ (in reduced units), which is lower than the folding transition temperature ($T_f = 1.112$). Energy is in units of kT.

configurational entropy (S) as the horizontal axis and the reaction coordinate (x) as the vertical axis. S is calculated with the density of states (n) computed with the *Tm*CSP simulated thermodynamic data by $S = log\,[n\,(x)]$ (see the Simulation details section). The kinetic folding times over the *Tm*CSP folding landscape will be further calculated analytically (integration of $P(x, t)$) and computationally (simulation runs of mean first passage time).

Fig. 2 shows the thermodynamic free energy profile ($F(x)$) as a function of the reaction coordinate native contacts (x) of the simulated protein *Tm*CSP. In Fig. 2, the free energy results calculated with the coarse-grained C_α-model are shown using circles for the simulated temperature $T = 1.080$ (in reduced units of simulation temperature), which is lower than the transition temperature for folding ($T = T_f = 1.112$). The simulated $F(x)$ has a thermodynamic free energy barrier of $\Delta F \sim 6.5kT$ centered at $x = 0$, consistent with previous experimental [74–77] and theoretical [42,63,66,86] studies of this hyperthermophilic protein. At lower temperature, folded states ($x_{folded} \sim +0.4$) are more stabilized than unfolded states ($x_{unfolded} \sim -0.4$). Also, Fig. 2 presents the simulated protein *Tm*CSP in its native folded and unfolded conformations. The simulated data are used to fit the analytical $F(x)$ by determining the constants a, b and c from Eq. (12). The adjusted values for $F(x)$ were $a = 460$, $b = -48$ and $c = -115$. The asymmetry of the double-well minima guarantees that the probability distribution function migrates from the unfolded to the folded state by the analytical method developed. It is possible to obtain the solution to the Schrödinger-type Eq. (3), where the effective potential (Eq. (10)) is written in terms of the analytical free energy profile $F(x)$ from Fig. 2 by following the procedure outlined in the previous sections.

The ground state energy eigenvalue and eigenfunction are determined as there is an exact analytical solution (Eqs. (17) and (18)) with constants a, b and c, previously computed. However, excited state eigenfunctions are calculated in terms of the variational parameters α_n, μ_n, k_n and η_n as they do not have exact analytical solution. As described in the Theory section, eigenvalues for the excited states were obtained by building a hierarchy of Hamiltonians [78]. The superpotentials generate wave functions for each excited state through the relation between the eigenfunctions, when applying the bosonic operator $a_n^+ = -\frac{d}{dx} + W_n(x)$ successively [78], as

$$\begin{aligned}
&\psi_0^{(1)}\\
&\psi_1^{(1)} = a_1^+\psi_0^{(2)}\\
&\psi_2^{(1)} = a_1^+\psi_1^{(2)} = a_1^+a_2^+\psi_0^{(3)}\\
&\psi_3^{(1)} = a_1^+\psi_2^{(2)} = a_1^+a_2^+\psi_2^{(3)} = a_1^+a_2^+a_3^+\psi_0^{(4)}\\
&\quad\vdots\\
&\psi_n^{(1)} = a_1^+a_2^+a_3^+\ldots a_n^+\psi_0^{(n+1)}.
\end{aligned} \tag{24}$$

Table 1 shows the results using the described iteration procedure in Eq. (24). It contains the minimized parameters α_n, μ_n, k_n, η_n and the energy eigenvalues E_n obtained by the iteration of the wave functions for excited states with the variational method and the free energy profile from Fig. 2. The energy eigenvalue calculations were truncated at $n = 5$ since the probability distributions for $n = 4$ and $n = 5$ are practically the same for long times ($t > 0.08$) and do not include any new contribution to $P(x, t)$. Also, the analytical computational time increases a great deal for higher order eigenvalues.

Fig. 3 presents the time-dependent probability distribution of the reaction coordinate x ($P(x, t)$, Eq. (2)). It shows P for three examples of increasing analytical times ($P(x, t = 0.046)$ in green, $P(x, t = 0.050)$ in red and $P(x, t = 0.080)$ in blue) with the initial condition of the reaction coordinate set at $x_0 = -0.35$ and with time in absolute units of the numerical

Table 1
Minimized parameters α_n, μ_n, k_n, η_n and the energy eigenvalues E_n obtained by the supersymmetric quantum mechanics formalism with the variational method using $a = 460$, $b = -48$, $c = -115$ and $\beta = \frac{1}{1.080}$.

n	α_n	μ_n	k_n	η_n	E_n
1	889.308	511.96	140.363	6.55981	188.159
2	−40.689	15.0036	84.7965	3.84468	190.815
3	−4.7722	−45.6839	69.8482	5.03949	297.025
4	866.976	−377.05	78.0321	−1.26092	368.834
5	386.904	−83.9753	53.774	−1.83139	404.222

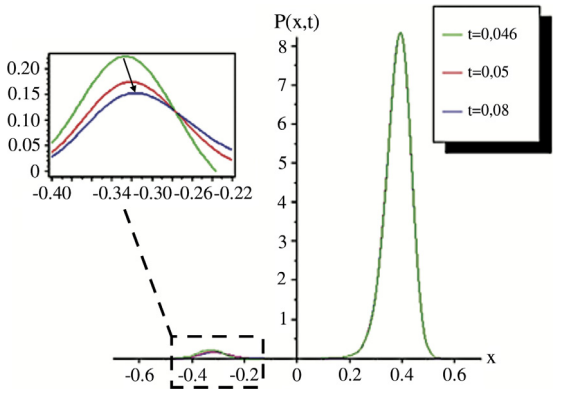

Fig. 3. Probability distribution ($P(x, t)$) as a function of the reaction coordinate (x) for different times (t) of P migrating from the unfolded free energy minimum ($x_{unfolded} \sim -0.4$) to the folded valley ($x_{folded} \sim +0.4$). $P(x, t)$ reaches the steady state for t higher than 0.08. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

integrations. The $P(x, t)$ was calculated with the eigenfunctions (Eq. (24)) and eigenvalues from Table 1. The series was truncated at $n = 5$, since at $t = 0.08$, with $n = 4$ and $n = 5$, both P distribution curves almost coincide. Fig. 3 shows that, for the three curves, the probability distribution decreases at the unfolded state ($x_{unfolded} \sim -0.4$) for increasing time and P concentrates in the region of x corresponding to the protein native state ($x_{folded} \sim +0.4$), the second minimum in the free energy $F(x)$ in Fig. 2. The temporal evolution of the system and the transition is observed between the two minima of the potential. Also, $P(x, t)$ diffuses fast over the thermodynamic free energy barrier at $x = 0$, even for low values of t. This is a consequence of the numerical calculations performed with the minimized variational parameters from Table 1 and the potential $F(x)$, which has its global minimum at the folded state. The probability of finding the system in the region corresponding to the minimum of the potential in $x > 0$ (the deeper minimum) is greater than the probability to find the system in the region corresponding to the shallow minimum in $x < 0$. The probability distribution reaches equilibrium, the stationary state, for higher time values ($t > 0.08$).

In order to observe the temporal evolution of the system, the probability as a function of time of the model protein being in the unfolded state were calculated by integration of Eq. (2) on the left side of the quadratic potential ($x < 0$)

$$N_1(t) = \int_{-1}^{0} P(x, t)\, dx \tag{25}$$

and the probability of the system being in the folded state by integrating P on the right side of the potential ($x > 0$)

$$N_2(t) = \int_{0}^{1} P(x, t)\, dx. \tag{26}$$

The next step was to introduce the quantity that calculates the mentioned temporal evolution of the system [31,39,56]

$$\Delta N(t) = N_1(t) - N_2(t) = \int_{-1}^{0} P(x, t)\, dx - \int_{0}^{1} P(x, t)\, dx \tag{27}$$

that is the difference of the areas close to each potential minimum. The relaxation rate of the system, initially at x_0, is completely accounted for by the exponential decay of $\Delta N(t)$ [38]. Thus, $\Delta N(t|x_0) = \exp[-t/\tau(x_0)]$, where τ is the relaxation time. τ measures the time that the system takes to reach equilibrium, the stationary state, and is related to the mean first passage time (mfpt) and the protein folding time (τ_f) [39,46,56,91].

Fig. 4 shows ΔN as a function of time (t) with the starting point set at $x_0 = -0.30$ and temperature $T = 1.080$, same as Fig. 2. In Fig. 4, the symbols represent ΔN as a function of t obtained by numerical calculations of Eq. (27) and the continuous

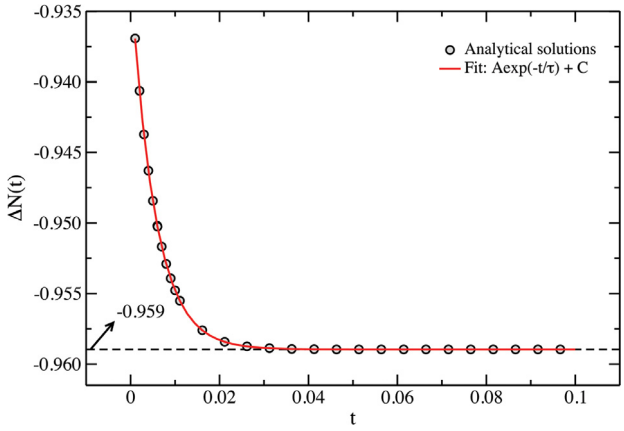

Fig. 4. Temporal evolution of the probability distributions. Difference between the integral of the probability distributions from folded and unfolded states (ΔN) as a function of time (t). For the filled circles, ΔN is obtained by numerically solving $\Delta N\ (t)$ from Eq. (27). The continuous curve is the exponential decay fit to the numerical data ($\tau = 5.42 \times 10^{-3}$ in units of inverse absolute time). τ is the relaxation time of the system, which is related to the protein folding transition time (τ_f). The broken line represents the convergence of the system approaching the stationary state. The system is at temperature $T = 1.080$ (in reduced units).

curve is the exponential decay adjustment ($\Delta N = A \exp(-t/\tau) + C$) and ($\tau = 5.42 \times 10^{-3}$ in units of inverse of absolute time). In Eq. (27) the initial condition is $\Delta N\ (t \to 0) \to 1$ due to the initial high concentration of the probability distribution at $x_0 = -0.30$ in the left well of the potential which represents the unfolded state. The probability concentration converges to $\Delta N \to -1$ ($N_1 \to 0$ and $N_2 \to 1$) as t increases and the system moves to the second minimum at $x > 0$, which is the folded state. Actually, the probability distribution is not all concentrated in the second minimum. There is a residual probability in the unfolded well, so the probability concentration $N_1\ (t)$ is $\sim 5\%$ of the total probability. This difference can be observed in Fig. 4. The exponent τ is the only parameter to be considered in this analysis. The parameters A and C in ΔN are not relevant for the present analysis. The exponent τ is extracted from the exponential fit to the curve as shown in Fig. 4. The value of τ corresponds to the folding transition time (τ_f) obtained from the folding simulations and also to the experimental folding rates of *Tm*CSP [92,93].

The temporal evolution of the system can also be evaluated for different initial conditions of the protein motion (x_0). The difference between the time-dependent probability concentrations in the double-well potential (Eq. (27)) is also dependent on the reaction coordinate starting point of the distribution ($\Delta N\ (t|x_0)$). Thus, ΔN was calculated by the numerical procedure (the same as Fig. 4) for different values of x_0 and the exponential decay as a function of the initial condition ($\tau(x_0)$) was extracted from each exponential fit ($\Delta N\ (t|x_0) = \exp[-t/\tau(x_0)]$). Fig. 5, on the left axis, compares the relaxation times as a function of the starting configuration of the system obtained by the numerical analysis ($\tau_f(x_0)$ in diamonds) with the folding times calculated with the computer simulations of *Tm*CSP ($\tau_f(x_0)$ in circles). The analytical folding transition times were rescaled with the simulation time step ($\Delta t = 0.5$ fs) in order to compare with the computational τ_f. Fig. 5 also shows, superposed, on the right axis, the free energy profile as a function of the native contacts $F(x)$ (shaded curve). The simulated $\tau_f(x_0)$ shows large error bars due to the large dispersion of the values of first passage times (that is, the time required for a random conformation to visit x_{fold} for the first time) of the unfolded model protein to overcome the energy barrier. Each value of τ, starting from a particular $x = x_0$, were averaged over a diverging number of fast and slow folding trajectories and this is a consequence of the stochasticity of the diffusion process of the protein folding problem studied. Nevertheless, the analytical and simulated $\tau_f(x_0)$ agrees reasonably over the entire reaction coordinate. It is worth noting in Fig. 5 that τ_f coincides with the $F(x)$ profile, τ_f decreases abruptly as the initial position of the motion approaches the thermodynamic energy barrier at $x = 0$. Earlier lattice, off-lattice and analytical theoretical works have also reported this behavior, an abrupt decrease in the kinetic folding times as the initial position of the motion approaches the thermodynamic barrier [31,66,68].

Exhaustive computational and experimental sampling of the protein folding thermodynamics and kinetics is replaced by an efficient numerical calculation. The numerical solutions of the FP equation with the double-well potential are shown to be a valuable tool to describe the kinetics of fixed temperature protein folding over a thermodynamic energy barrier. The probability distributions diffuse through the free energy barrier and the supersymmetric quantum mechanics method applied to the Schrödinger-type equation was sufficient to accurately calculate the folding times for the given protein *Tm*CSP. The estimates from the analytical results capture the dynamic process of the protein, as it escapes from local traps along the energy landscape.

4.2. Temperature-dependent thermodynamic profiles and kinetic folding times

This section extends the analytical and computational studies for a range of temperatures commonly explored in protein folding studies. The thermodynamic free energy profile ($F(x)$) and other related parameters are transformed into

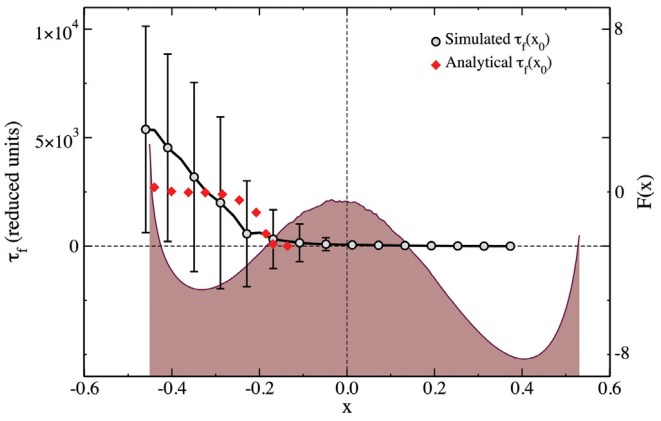

Fig. 5. On the left axis, the folding times (τ_f) versus the starting point of the system (x_0) of the reaction coordinate at $T = 1.080$ (in reduced unit of temperature). Filled circles are the simulated folding times and diamonds represent the analytical folding times. The relaxation time is extracted from the exponential decay fit of the analytical expression $\Delta N\,(t|x_0)$ (Eq. (27)) for each initial condition of the motion (x_0). On the right axis, the free energy profile ($F(x)$) is shown as a function of the reaction coordinate (x). Error bars of the simulated $\tau_f(x_0)$ are the standard deviation of the mean over 1000 runs for each x_0. The vertical broken line marks the position of the thermodynamic free energy barrier.

temperature-dependent parameters. The numerical methodology developed to solve the Schrödinger-type equation with the variational method has to be modified in order to properly shift the difference between the two minima of $F(x)$ and to account for the time-dependent probabilities diffusing from one minimum to another over the energy barrier.

Accordingly, the double-well free energy as a function of the reaction coordinate adopted in this work (Eq. (12)) for a given temperature (T) is rewritten as

$$F_T\,(x) = ax^4 + b\left(\frac{T}{T_f}\right)x^3 + cx^2 \tag{28}$$

where T_f is the transition temperature between folded and unfolded states. Specifically for further calculations, $a > 0$ and $c < 0$. The term $b\left(T/T_f\right)$ in Eq. (28) is a function that depends of the given temperature (T) of the system. To model b in Eq. (28), the hyperbolic function

$$b\left(\frac{T}{T_f}\right) = C\tanh\left(\frac{T}{T_f} - 1\right) \tag{29}$$

was chosen where C is a constant. The above function in Eq. (29) incorporates the properties to numerically solve the FP equation and provides the elements for the free energy profile to vary smoothly with the temperature of the system.

The polynomial adjustments to the values of the free energy obtained from the C_α structure-based model of *Tm*CSP in the proposed model were made in Eqs. (28) and (29). Fig. 6 shows the free energy profile as a function of the reaction coordinate ($F_T\,(x)$) for three relevant temperatures: $T = 1.080 < T_f$ (circles), $T = T_f = 1.112$ (squares) and $T = 1.130 > T_f$ (triangles). The fit to the simulation data is represented by the continuum curves with the polynomial function from Eq. (29) which have been adjusted with the parameters $T_f = 1.112$ (folding temperature) and $C = 1669$. F represents the balance between the total energy (E) and the configurational entropy ($-TS$) of the system. A lower temperature contributes to diminishing the entropic term and, as a consequence, the probability to find the system in the native basing at $x_{folded} \sim +0.4$ is higher ($T < T_f$ in Fig. 6). On the other hand, a higher temperature plays the role of stabilizing high energetic states and the unfolded state ($x_{unfolded} \sim -0.4$) becomes more probable ($T > T_f$ in Fig. 6). At the transition temperature (T_f), both free energy terms, the energetic and the entropic, are balanced along the reaction coordinate except at the transition state around $x = 0$; an effective thermodynamic barrier arises in between the folded and the unfolded states. These temperature-dependent thermodynamic effects are mapped by the Schrödinger-type equation and the time-dependent probability distribution motions can be investigated analytically for different temperatures.

The folding times (τ_f), i.e. the characteristic times of the protein folding process, were calculated for different thermodynamic conditions following the same protocol presented in the previous section. The free energy profiles obtained with the thermodynamic algorithm WHAM for different temperature runs of the C_α-model were fitted with the polynomial function to extract the potentials input into the FP equation. The resulting probability distribution curves were used to obtain the concentration probabilities; ΔN from Eq. (27) was calculated for each temperature. The procedure used to calculate ΔN is the same as that presented in the previous section. The exponential fit of the type $\Delta N = e^{-t/\tau}$ to the numerical calculations of ΔN versus t was performed and τ was extracted for each temperature. Each τ was obtained with the initial condition of the probability concentrations: $x_0 = -0.3$, if $T < T_f$, and $x_0 = +0.3$, if $T > T_f$. Therefore, the concentrations are able to flow to the deeper potential minimum, a condition required for the numerical calculations performed with the temporal evolutions of the Schrödinger-type equation.

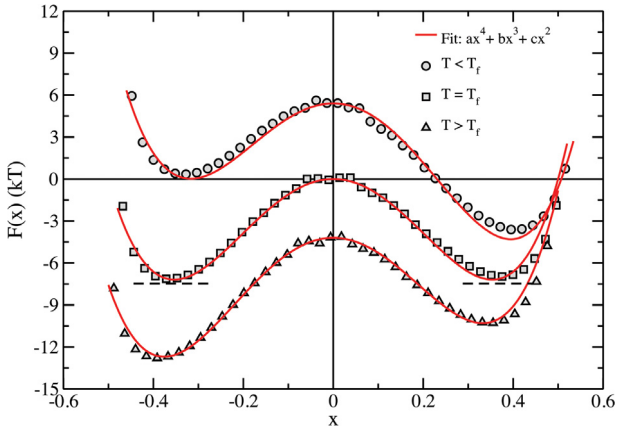

Fig. 6. Thermodynamic free energy profile ($F(x)$) as a function of the reaction coordinate native contacts (x) for different temperatures (T). Filled shapes represent $F(x)$ obtained by the C_α-SBM simulations and the continuum curves are the polynomial fit ($F(x) = ax^4 + bx^3 + cx^2$) over the simulation data with parameters $a = 460$, $c = -115$ and $b = -48$ (for $T = 1.080 < T_f$)), 0 (for $T = T_f = 1.112$) and 27 (for $T = 1.130 > T_f$). T_f is defined as the folding transition temperature in which the system has equal probabilities of being folded ($x > 0$) and unfolded ($x < 0$), marked by the aligned dashed lines in $T = T_f$. Energy is in units of kT.

Fig. 7 presents the folding time (τ_f) as a function of temperature (T) rescaled by the *Tm*CSP folding temperature (T_f). On a logarithmic scale, Fig. 7 shows a good agreement between τ_f calculated with the folding time runs of the C_α-SBM simulations and the analytical folding times given by the supersymmetric quantum mechanics method, previously described. In Fig. 7, error bars of the folding time simulations were calculated as the standard deviation of the mean over diverse 1000 first passage time runs, same effect as in Fig. 5 for the $\tau_f(x_0)$. Fig. 7 shows the characteristic 'U' shape average kinetic behavior in temperature or denaturant concentration of many earlier investigations: i.e. theoretical (analytical and computational) folding times and experimental Chevron Plots for folding rates [15,57,66,94–96]. This occurs because the folded state is unstable at a high temperature, high energetic states are more populated and, thus folding to the lowest energy native state is slow. At the other end, the system has low thermal energy at a low temperature, becomes kinetically trapped in the underlying energy landscape roughness and thus, folding abruptly slows down. At the fastest folding time temperature ($T \approx 0.85$), the thermodynamic free energy barrier is lower than that in T_f, nevertheless with enough thermal energy to overcome the landscape roughness and thus, folding is fast. The temperature region that is interesting to investigate is that around T_f and the analytical calculations were focused on this vicinity. Exactly at T_f, it was not possible to calculate the analytical τ_f due to numerical instabilities. The approach developed in this study is based on the diffusion of the time-dependent probability distributions between both wells of the bistable potential and, as the potential is symmetrical at T_f, there is no effective migration of the probability distribution from one minimum to the other. Even so, there is reasonable agreements between analytical and simulation τ_f results.

Despite the fact that the temperature-dependent kinetic rates, or flux, calculated with the analytical method outlined in this paper were in a good agreement with the simulations, inside the mean first passage time dispersion, τ_f were overestimated for $T < T_f$ and, τ_f were underestimated for $T > T_f$ if the average τ_f is considered. This is a direct reflection of the eigenfunction expansion method of the Schrödinger equation, which is fairly restrictive as previously quoted [2]. Five terms were used for the eigenfunction iteration to reach the numerical solutions as stated in Table 1 and more terms could be considered in a much bigger numerical effort to increase accuracy. Another less computationally expensive improvement to the analytical method would be to compute a coordinate-dependent diffusion coefficient ($D(x)$) in the Schrödinger-type equation, instead of the constant diffusion D employed in this paper. $D(x)$ takes into account the effective kinetic free energy barrier of folding (about $1.0kT$), which increases the thermodynamic free energy barrier (originally $5–7kT$) and shifts the transition state [66]. Also, one might consider the funnel landscape roughness (as represented in Fig. 1) in the drift of the probability distributions, which is on average $1.7kT$ [66]. The folding landscape roughness varies as a function of position (as seen in Fig. 1) and time, which is described by the diffusion coefficient [66]. Previous studies have suggested that not only coordinate but also time-dependent diffusion plays a significant role in the folding of *Tm*CSP and other proteins as well [49,82,84], but this was not included in this study. As seen, there is a lot of room to improve the accuracy of the supersymmetric quantum mechanics approach to the kinetic estimations using thermodynamic profiles as an input for the Schrödinger-type equation.

5. Conclusion

This work developed both analytical and computational methods to explore the kinetics of the time-dependent probability distributions over thermodynamic free energy profiles of protein folding. The analytical results were in complete

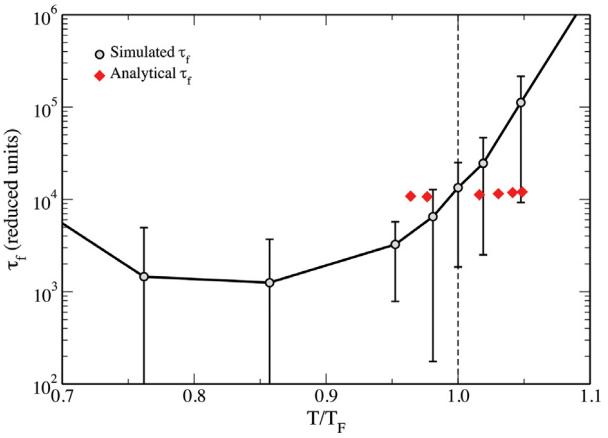

Fig. 7. On a logarithmic scale, the folding times (τ_f) versus temperature rescaled by the folding temperature (T/T_f). Filled circles are the simulated folding times obtained by kinetic C_α-SBM runs of the protein *Tm*CSP and diamonds represent the analytical folding times calculated by the exponential decay fit of the expression $\Delta N\,(t)$ (Eq. (27)) for each temperature (T). Error bars of the simulated τ_f were calculated as the standard deviation of the mean over 1000 runs for each T. The vertical broken line marks the folding temperature.

agreement with the simulations. The analytical method developed was to solve the Fokker–Planck equation with the supersymmetric quantum mechanics formalism mapped onto a converted Schrödinger-type equation and the variational method was used in the numerical calculations. The coarse-grained structure-based model of the protein *Tm*CSP was executed at a C_α level of resolution and the protein thermodynamics and kinetics were fully characterized by the simulations.

The computational double-well free energy profiles were modeled (Figs. 2 and 5) and used as the input for the diffusion equation of the Schrödinger-type. Analytical expressions of probabilities of concentrations migrating between the potential minima (Fig. 3) described reasonably the folding kinetics as revealed by the analytical folding times extracted from the concentration differences (Fig. 4). Analytical and computational τ_f were coincident and strongly dependent on the initial condition of the motion (x_0). $\tau(x_0)$ follows exponential decay distributions as the thermodynamic barrier is approached (Fig. 5), a result similar to that of previous works [31,66,68].

This paper also developed temperature-dependent parameters for the numerical calculations of the probabilities (Fig. 7). Temperature has an important role in theoretical and experimental protein folding studies. Folding times as a function of temperature has the well-known 'U' shape and analytical and simulation results converged to similar values within error boundaries. It was outlined previously by analytical works that folding times or rate mechanisms are strongly affected by the temperature range [31,66]. Low to intermediate temperatures where the system has enough energy to overcome the activation barrier gives the Arrhenius behavior, non-exponential and non-Poissonian kinetic distributions. At lower and higher temperatures, the system deviates from the Arrhenius behavior and exponential and Poissonian kinetic distributions better describe the rates. This temperature-dependent phenomenon has implications to downhill (low or none thermodynamic folding barrier) and activated folding dynamics of several proteins as seen in both theory and experiment [67,97]. The analytical method allows the description of the temperature range for the Arrhenius behavior (in Fig. 7), non-exponential and non-Poissonian kinetic distributions, around the transition temperature for folding.

Folding is governed by the diffusion of $P(x, t)$ throughout the funneled landscape surface (Fig. 1). Landscape roughness plays an important role in the protein motion or distributions down the funnel and, therefore, the diffusion coefficient as a function of the reaction coordinate ($D(x)$) might influence the analytical folding times in Fig. 7. However, the analytical τ_f obtained with the Schrödinger equation solutions with constant diffusion coefficient, was in line with simulations and experiments. For this reason, this study was successful in its first attempt to build simple and numerically cheap model to describe the kinetics of activated dynamics. One could enhance the accuracy by including $D(x)$ to the model and/or by accounting for more terms in the eigenfunction iteration method. The laborious supersymmetric quantum mechanics method applied to the Schrödinger equation resulted in a rather simple analytical expression to investigate folding kinetics mechanisms.

The theory and methodology developed in this study provide the basis for connecting theory to experiments. The methodology can be employed in studies where theoreticians and experimentalists have determined the thermodynamic barrier and are interested in predictions of the kinetic behavior of the system. The theory can be applied to a wide variety of biological as well as condensed-phase two-state systems in order to investigate their kinetics mechanisms.

Acknowledgments

The authors thank Profs. Vitor B.P. Leite, Leandro C. Oliveira and Sidney J. Carvalho for new insights and helpful discussions. This research was supported by resources supplied by the Núcleo de Computação Científica (NCC/GridUNESP) of

the Universidade Estadual Paulista (UNESP), Centro Nacional de Processamento de Alto Desempenho (CENAPAD-SP) of the Universidade Estadual de Campinas (UNICAMP) and Prof. Leite (FAPESP grant 2011/17658-3). This research was supported by the Brazilian agencies FAPESP, FAPEMIG (APQ-00941-14), CAPES and CNPq (441730/2014-0).

Authors' contributions

EDF, JC, and RJO conceived and designed the theory/experiments. FP and RJO performed the calculations. Data was analyzed by FP, EDF, JC and RJO. FP, EDF, JC and RJO wrote the manuscript. All authors read and approved the final version of the manuscript.

Competing interests

The authors declare that they have no competing interests.

References

[1] H. Risken, The Fokker–Planck Equation: Methods of Solution and Applications, Springer, 1989.
[2] C.-L. Ho, R. Sasaki, Quasi-exactly solvable Fokker–Planck equations, Ann. Phys. 323 (2008) 883–892. http://dx.doi.org/10.1016/j.aop.2007.12.001.
[3] C.F. Curtiss, R.B. Bird, Fokker–Planck equation for the one-molecule distribution function in polymer mixtures and its solution, J. Chem. Phys. 106 (1997) 9899–9921. http://dx.doi.org/10.1063/1.473878.
[4] K. Lee, W. Sung, Ion transport and channel transition in biomembranes, Phys. Statist. Mech. Appl. 315 (2002) 79–97. http://dx.doi.org/10.1016/S0378-4371(02)01247-5.
[5] D. Brigo, F. Mercurio, Lognormal-mixture dynamics and calibration to market volatility smiles, Int. J. Theor. Appl. Finance 05 (2002) 427–446. http://dx.doi.org/10.1142/S0219024902001511.
[6] C. Montagnon, A closed solution to the Fokker–Planck equation applied to forecasting, Phys. Statist. Mech. Appl. 420 (2015) 14–22. http://dx.doi.org/10.1016/j.physa.2014.10.079.
[7] T.J. Murphy, J.L. Aguirre, Brownian motion of N interacting particles. I. Extension of the Einstein diffusion relation to the N-particle case, J. Chem. Phys. 57 (1972) 2098–2104. http://dx.doi.org/10.1063/1.1678535.
[8] H.T. Davis, K. Hiroike, S.A. Rice, On the quantum-mechanical Fokker—Planck equation, J. Chem. Phys. 43 (1965) 2633–2642. http://dx.doi.org/10.1063/1.1697188.
[9] S.A. Hughes, R.D. Blandford, Black hole mass and spin coevolution by mergers, Astrophys. J. Lett. 585 (2003) L101. http://dx.doi.org/10.1086/375495.
[10] F. Reif, Fundamentals of Statistical and Thermal Physics, McGraw-Hill, 1985.
[11] J.D. Bryngelson, P.G. Wolynes, Intermediates and barrier crossing in a random energy model (with applications to protein folding), J. Phys. Chem. 93 (1989) 6902–6915. http://dx.doi.org/10.1021/j100356a007.
[12] P. Faccioli, M. Sega, F. Pederiva, H. Orland, Dominant pathways in protein folding, Phys. Rev. Lett. 97 (2006) 108101. http://dx.doi.org/10.1103/PhysRevLett.97.108101.
[13] K.A. Dill, J.L. MacCallum, The protein-folding problem, 50 years on, Science 338 (2012) 1042–1046. http://dx.doi.org/10.1126/science.1219021.
[14] R. Du, V.S. Pande, A.Y. Grosberg, T. Tanaka, E.S. Shakhnovich, On the transition coordinate for protein folding, J. Chem. Phys. 108 (1998) 334–350. http://dx.doi.org/10.1063/1.475393.
[15] J.D. Bryngelson, J.N. Onuchic, N.D. Socci, P.G. Wolynes, Funnels, pathways, and the energy landscape of protein folding: a synthesis, Proteins 21 (1995) 167–195. http://dx.doi.org/10.1002/prot.340210302.
[16] C.-L. Lee, G. Stell, J. Wang, First-passage time distribution and non-Markovian diffusion dynamics of protein folding, J. Chem. Phys. 118 (2003) 959. http://dx.doi.org/10.1063/1.1527672.
[17] S.K. Banik, B.C. Bag, D.S. Ray, Generalized quantum Fokker–Planck, diffusion, and Smoluchowski equations with true probability distribution functions, Phys. Rev. E 65 (2002) 051106. http://dx.doi.org/10.1103/PhysRevE.65.051106.
[18] T. Tomé, M.J. de Oliveira, Stochastic Dynamics and Irreversibility, Springer International Publishing, Cham, 2015. http://dx.doi.org/10.1007/978-3-319-11770-6.
[19] M. Bernstein, L.S. Brown, Supersymmetry and the Bistable Fokker–Planck equation, Phys. Rev. Lett. 52 (1984) 1933–1935. http://dx.doi.org/10.1103/PhysRevLett.52.1933.
[20] L. é Gendenshteĭn, I.V. Krive, Supersymmetry in quantum mechanics, Sov. Phys. Uspekhi. 28 (1985) 645–666. http://dx.doi.org/10.1070/PU1985v028n08ABEH003882.
[21] F. Marchesoni, P. Sodano, M. Zannetti, Supersymmetry and bistable soft potentials, Phys. Rev. Lett. 61 (1988) 1143–1146. http://dx.doi.org/10.1103/PhysRevLett.61.1143.
[22] F. Marchesoni, C. Presilla, P. Sodano, Supersymmetry and activation rates, Phys. Lett. A 134 (1989) 465–468. http://dx.doi.org/10.1016/0375-9601(89)90686-5.
[23] F. Cooper, A. Khare, U. Sukhatme, Supersymmetry and quantum mechanics, Phys. Rep. 251 (1995) 267–385. http://dx.doi.org/10.1016/0370-1573(94)00080-M.
[24] G. Junker, Quantum and classical stochastic dynamics: Exactly solvable models by supersymmetric methods, in: III Int. Workshop Class. Quantum Integrable Syst., JINR Publishing, Yerevan, Armenia, 1998, pp. 94–103 http://arxiv.org/abs/quant-ph/9810070 (accessed 09.09.15).
[25] E. Drigo Filho, R.M. Ricotta, Morse potential energy spectra through the variational method and supersymmetry, Phys. Lett. A 269 (2000) 269–276. http://dx.doi.org/10.1016/S0375-9601(00)00267-X.
[26] C.V. Sukumar, Supersymmetric quantum mechanics and its applications, in: R. Bijker, O. Castanos, L. Urrutia, D. Fernandez, H. MoralesTecoti, C. Villarreal (Eds.), Lat.-Am. Sch. Phys. - Xxxv Elaf Supersymmetries Phys. Its Appl., Amer Inst Physics, Melville, 2005, pp. 166–235.
[27] A. Schulze-Halberg, J.M. Rivas, J.J. Peña Gil, J. García-Ravelo, P. Roy, Exact solutions of the Fokker–Planck equation from an nth order supersymmetric quantum mechanics approach, Phys. Lett. A 373 (2009) 1610–1615. http://dx.doi.org/10.1016/j.physleta.2009.03.005.
[28] V.P. Berezovoj, G.I. Ivashkevych, M.I. Konchatnij, Exactly solvable diffusion models in the framework of the extended supersymmetric quantum mechanics, Phys. Lett. A 374 (2010) 1197–1200. http://dx.doi.org/10.1016/j.physleta.2009.12.066.
[29] G.R.P. Borges, E. Drigo Filho, R.M. Ricotta, Variational supersymmetric approach to evaluate Fokker–Planck probability, Physica A 389 (2010) 3892–3899. http://dx.doi.org/10.1016/j.physa.2010.05.027.
[30] Y. Basak, O. Ozer, Variational method and supersymmetric techniques for the Fokker-Plank probability, BPL 20 (2012) 231–242.

[31] D. Caldas, J. Chahine, E. Drigo Filho, The Fokker–Planck equation for a bistable potential, Physica A 412 (2014) 92–100. http://dx.doi.org/10.1016/j.physa.2014.06.009.
[32] R.C. Anjos, G.B. Freitas, C.H. Coimbra-Araújo, Analytical solutions of the Fokker–Planck equation for generalized morse and hulthén potentials, J. Stat. Phys. 162 (2016) 387–396. http://dx.doi.org/10.1007/s10955-015-1414-7.
[33] E. Gozzi, M. Reuter, W.D. Thacker, Variational methods via supersymmetric techniques, Phys. Lett. A 183 (1993) 29–32. http://dx.doi.org/10.1016/0375-9601(93)90883-2.
[34] E. Drigo Filho, R.M. Ricotta, Supersymmetric variational energies of 3D confined potentials, Phys. Lett. A 320 (2003) 95–102. http://dx.doi.org/10.1016/j.physleta.2003.11.014.
[35] Y.P. Varshni, Susy superpotentials for a confined hulthén potential, Modern Phys. Lett. A 19 (2004) 2757–2764. http://dx.doi.org/10.1142/S0217732304016202.
[36] D.J. Kouri, T. Markovich, N. Maxwell, E.R. Bittner, Supersymmetric quantum mechanics, excited state energies and wave functions, and the Rayleigh—Ritz variational principle: A proof of principle study, J. Phys. Chem. A 113 (2009) 15257–15264. http://dx.doi.org/10.1021/jp905798m.
[37] A. Vega, J. Flores, Heavy quarkonium properties from Cornell potential using variational method and supersymmetric quantum mechanics, Pramana 87 (2016) 73. http://dx.doi.org/10.1007/s12043-016-1278-7.
[38] F.R. Silva, E.D. Filho, A new type of confinement for the Morse potential, Chem. Phys. Lett. 498 (2010) 198–202. http://dx.doi.org/10.1016/j.cplett.2010.08.032.
[39] K. Schulten, Z. Schulten, A. Szabo, Dynamics of reactions involving diffusive barrier crossing, J. Chem. Phys. 74 (1981) 4426–4432. http://dx.doi.org/10.1063/1.441684.
[40] J.D. Bryngelson, P.G. Wolynes, Spin glasses and the statistical mechanics of protein folding, Proc. Natl. Acad. Sci. 84 (1987) 7524–7528.
[41] E.A. Lipman, B. Schuler, O. Bakajin, W.A. Eaton, Single-molecule measurement of protein folding kinetics, Science 301 (2003) 1233–1235. http://dx.doi.org/10.1126/science.1085399.
[42] J. Wang, R.J. Oliveira, X. Chu, P.C. Whitford, J. Chahine, W. Han, E. Wang, J.N. Onuchic, V.B.P. Leite, Topography of funneled landscapes determines the thermodynamics and kinetics of protein folding, Proc. Natl. Acad. Sci. 109 (2012) 15763–15768. http://dx.doi.org/10.1073/pnas.1212842109.
[43] H.A. Kramers, Brownian motion in a field of force and the diffusion model of chemical reactions, Physica 7 (1940) 284–304. http://dx.doi.org/10.1016/S0031-8914(40)90098-2.
[44] E. Drigo Filho, R.M. Ricotta, Supersymmetric variational energies for the confined Coulomb system, Phys. Lett. A 299 (2002) 137–143. http://dx.doi.org/10.1016/S0375-9601(02)00610-2.
[45] S.H. Northrup, J.T. Hynes, Reactive dynamics for diffusive barrier crossing, J. Chem. Phys. 69 (1978) 5246–5260. http://dx.doi.org/10.1063/1.436578.
[46] D.J. Bicout, A. Szabo, First passage times, correlation functions, and reaction rates, J. Chem. Phys. 106 (1997) 10292–10298. http://dx.doi.org/10.1063/1.474066.
[47] J. Dunkel, W. Ebeling, L. Schimansky-Geier, P. Hänggi, Kramers problem in evolutionary strategies, Phys. Rev. E 67 (2003) 061118. http://dx.doi.org/10.1103/PhysRevE.67.061118.
[48] W. Zheng, R.B. Best, Reduction of all-atom protein folding dynamics to one-dimensional diffusion, J. Phys. Chem. B 119 (2015) 15247–15255. http://dx.doi.org/10.1021/acs.jpcb.5b09741.
[49] R.B. Best, G. Hummer, Diffusion models of protein folding, Phys. Chem. Chem. Phys. 13 (2011) 16902–16911. http://dx.doi.org/10.1039/C1CP21541H.
[50] P. Wolynes, J. Onuchic, D. Thirumalai, Navigating the folding routes, Science 267 (1995) 1619–1620. http://dx.doi.org/10.1126/science.7886447.
[51] L.L. Chavez, J.N. Onuchic, C. Clementi, Quantifying the roughness on the free energy landscape: entropic bottlenecks and protein folding rates, J. Am. Chem. Soc. 126 (2004) 8426–8432. http://dx.doi.org/10.1021/ja049510+.
[52] C. Clementi, H. Nymeyer, J.N. Onuchic, Topological and energetic factors: what determines the structural details of the transition state ensemble and "en-route" intermediates for protein folding? an investigation for small globular proteins, J. Mol. Biol. 298 (2000) 937–953. http://dx.doi.org/10.1006/jmbi.2000.3693.
[53] K. Neupane, A.P. Manuel, M.T. Woodside, Protein folding trajectories can be described quantitatively by one-dimensional diffusion over measured energy landscapes, Nat. Phys. 12 (2016) 1–5. http://dx.doi.org/10.1038/nphys3677.
[54] K. Neupane, D.A.N. Foster, D.R. Dee, H. Yu, F. Wang, M.T. Woodside, Direct observation of transition paths during the folding of proteins and nucleic acids, Science 352 (2016) 239–242. http://dx.doi.org/10.1126/science.aad0637.
[55] K. Neupane, F. Wang, M.T. Woodside, Direct measurement of sequence-dependent transition path times and conformational diffusion in DNA duplex formation, Proc. Natl. Acad. Sci. 114 (2017) 1329–1334. http://dx.doi.org/10.1073/pnas.1611602114.
[56] B. Jun, D.L. Weaver, One-dimensional potential barrier model of protein folding with intermediates, J. Chem. Phys. 116 (2002) 418–426. http://dx.doi.org/10.1063/1.1425819.
[57] N.D. Socci, J.N. Onuchic, P.G. Wolynes, Diffusive dynamics of the reaction coordinate for protein folding funnels, J. Chem. Phys. 104 (1996) 5860. http://dx.doi.org/10.1063/1.471317.
[58] M. Hinczewski, Y. von Hansen, J. Dzubiella, R.R. Netz, How the diffusivity profile reduces the arbitrariness of protein folding free energies, J. Chem. Phys. 132 (2010) 245103. http://dx.doi.org/10.1063/1.3442716.
[59] H. Frauenfelder, S. Sligar, P. Wolynes, The energy landscapes and motions of proteins, Science 254 (1991) 1598–1603. http://dx.doi.org/10.1126/science.1749933.
[60] J.N. Onuchic, Z. Luthey-Schulten, P.G. Wolynes, Theory of protein folding: The energy landscape perspective, Annu. Rev. Phys. Chem. 48 (1997) 545–600. http://dx.doi.org/10.1146/annurev.physchem.48.1.545.
[61] V.I. Abkevich, A.M. Gutin, E.I. Shakhnovich, Free energy landscape for protein folding kinetics: Intermediates, traps, and multiple pathways in theory and lattice model simulations, J. Chem. Phys. 101 (1994) 6052–6062. http://dx.doi.org/10.1063/1.467320.
[62] K.A. Dill, H.S. Chan, From Levinthal to pathways to funnels, Nat. Struct. Mol. Biol. 4 (1997) 10–19. http://dx.doi.org/10.1038/nsb0197-10.
[63] R.J. Oliveira, P.C. Whitford, J. Chahine, J. Wang, J.N. Onuchic, V.B.P. Leite, The origin of nonmonotonic complex behavior and the effects of nonnative interactions on the diffusive properties of protein folding, Biophys. J. 99 (2010) 600–608. http://dx.doi.org/10.1016/j.bpj.2010.04.041.
[64] C.M. Dobson, Protein folding and misfolding, Nature 426 (2003) 884–890. http://dx.doi.org/10.1038/nature02261.
[65] N. Koga, S. Takada, Roles of native topology and chain-length scaling in protein folding: A simulation study with a Go-like model, J. Mol. Biol. 313 (2001) 171–180. http://dx.doi.org/10.1006/jmbi.2001.5037.
[66] R.J. Oliveira, P.C. Whitford, J. Chahine, V.B.P. Leite, J. Wang, Coordinate and time-dependent diffusion dynamics in protein folding, Methods 52 (2010) 91–98. http://dx.doi.org/10.1016/j.ymeth.2010.04.016.
[67] W. Xu, Z. Lai, R.J. Oliveira, V.B.P. Leite, J. Wang, Configuration-dependent diffusion dynamics of downhill and two-state protein folding, J. Phys. Chem. B (2012). http://dx.doi.org/10.1021/jp212132v.
[68] J. Chahine, R.J. Oliveira, V.B.P. Leite, J. Wang, Configuration-dependent diffusion can shift the kinetic transition state and barrier height of protein folding, Proc. Natl. Acad. Sci. 104 (2007) 14646–14651. http://dx.doi.org/10.1073/pnas.0606506104.
[69] S.E. Jackson, A.R. Fersht, Folding of chymotrypsin inhibitor 2. 1. Evidence for a two-state transition, Biochem. (Mosc.) 30 (1991) 10428–10435. http://dx.doi.org/10.1021/bi00107a010.

[70] H.S. Chung, J.M. Louis, W.A. Eaton, Experimental determination of upper bound for transition path times in protein folding from single-molecule photon-by-photon trajectories, Proc. Natl. Acad. Sci. 106 (2009) 11837–11844. http://dx.doi.org/10.1073/pnas.0901178106.
[71] C.D. Snow, E.J. Sorin, Y.M. Rhee, V.S. Pande, How well can simulation predict protein folding kinetics and thermodynamics? Annu. Rev. Biophys. Biomol. Struct. 34 (2005) 43–69. http://dx.doi.org/10.1146/annurev.biophys.34.040204.144447.
[72] X. Chu, L. Gan, E. Wang, J. Wang, Quantifying the topography of the intrinsic energy landscape of flexible biomolecular recognition, Proc. Natl. Acad. Sci. (2013) 201220699. http://dx.doi.org/10.1073/pnas.1220699110.
[73] P.C. Whitford, J.K. Noel, S. Gosavi, A. Schug, K.Y. Sanbonmatsu, J.N. Onuchic, An all-atom structure-based potential for proteins: Bridging minimal models with all-atom empirical forcefields, Proteins Struct. Funct. Bioinform. 75 (2009) 430–441. http://dx.doi.org/10.1002/prot.22253.
[74] B. Schuler, E.A. Lipman, W.A. Eaton, Probing the free-energy surface for protein folding with single-molecule fluorescence spectroscopy, Nature 419 (2002) 743–747. http://dx.doi.org/10.1038/nature01060.
[75] B. Schuler, W. Kremer, H.R. Kalbitzer, R. Jaenicke, Role of entropy in protein thermostability: Folding kinetics of a hyperthermophilic cold shock protein at high temperatures using 19F NMR†, Biochem. (Mosc.) 41 (2002) 11670–11680. http://dx.doi.org/10.1021/bi026293l.
[76] D. Perl, C. Welker, T. Schindler, K. Schroder, M.A. Marahiel, R. Jaenicke, F.X. Schmid, Conservation of rapid two-state folding in mesophilic, thermophilic and hyperthermophilic cold shock proteins, Nat. Struct. Mol. Biol. 5 (1998) 229–235. http://dx.doi.org/10.1038/nsb0398-229.
[77] D. Wassenberg, C. Welker, R. Jaenicke, Thermodynamics of the unfolding of the cold-shock protein from Thermotoga maritima, J. Mol. Biol. 289 (1999) 187–193. http://dx.doi.org/10.1006/jmbi.1999.2772.
[78] F. Cooper, A. Khare, U.P. Sukhatme, Supersymmetry in Quantum Mechanics, World Scientific, 2001.
[79] E. Drigo Filho, Supersimetria Aplicada à MecâNica QuâNtica, UNESP, 2009.
[80] E. Hille, Ordinary Differential Equations in the Complex Domain, Dover Publications, Mineola N.Y., 1997.
[81] R.B. Best, G. Hummer, Coordinate-dependent diffusion in protein folding, Proc. Natl. Acad. Sci. 107 (2010) 1088–1093. http://dx.doi.org/10.1073/pnas.0910390107.
[82] S. Yang, J.N. Onuchic, H. Levine, Effective stochastic dynamics on a protein folding energy landscape, J. Chem. Phys. 125 (2006) 054910. http://dx.doi.org/10.1063/1.2229206.
[83] R. Coifman, I. Kevrekidis, S. Lafon, M. Maggioni, B. Nadler, Diffusion maps, reduction coordinates, and low dimensional representation of stochastic systems, Multiscale Model. Simul. 7 (2008) 842–864. http://dx.doi.org/10.1137/070696325.
[84] A.K. Sangha, T. Keyes, Proteins fold by subdiffusion of the order parameter, J. Phys. Chem. B 113 (2009) 15886–15894. http://dx.doi.org/10.1021/jp907009r.
[85] W. Kremer, B. Schuler, S. Harrieder, M. Geyer, W. Gronwald, C. Welker, R. Jaenicke, H.R. Kalbitzer, Solution NMR structure of the cold-shock protein from the hyperthermophilic bacterium Thermotoga maritima, Eur. J. Biochem. 268 (2001) 2527–2539. http://dx.doi.org/10.1046/j.1432-1327.2001.02127.x.
[86] V.G. Contessoto, D.T. Lima, R.J. Oliveira, A.T. Bruni, J. Chahine, V.B.P. Leite, Analyzing the effect of homogeneous frustration in protein folding, Proteins Struct, Funct. Bioinforma 81 (2013) 1727–1737. http://dx.doi.org/10.1002/prot.24309.
[87] J.K. Noel, P.C. Whitford, K.Y. Sanbonmatsu, J.N. Onuchic, SMOG@ctbp: simplified deployment of structure-based models in GROMACS, Nucleic Acids Res. (2010) 1–5. http://dx.doi.org/10.1093/nar/gkq498.
[88] J.K. Noel, P.C. Whitford, J.N. Onuchic, The shadow map: A general contact definition for capturing the dynamics of biomolecular folding and function, J. Phys. Chem. B 116 (2012) 8692–8702. http://dx.doi.org/10.1021/jp300852d.
[89] S.S. Cho, Y. Levy, P.G. Wolynes, P versus Q: Structural reaction coordinates capture protein folding on smooth landscapes, Proc. Natl. Acad. Sci. USA 103 (2006) 586–591. http://dx.doi.org/10.1073/pnas.0509768103.
[90] S. Kumar, J. Rosenberg, D. Bouzida, R. Swendsen, P. Kollman, The weighted histogram analysis method for free-energy calculations on biomolecules, I. The method, J. Comput. Chem. 13 (1992) 1011–1021.
[91] A. Szabo, K. Schulten, Z. Schulten, First passage time approach to diffusion controlled reactions, J. Chem. Phys. 72 (1980) 4350. http://dx.doi.org/10.1063/1.439715.
[92] B. Schuler, Single-molecule fluorescence spectroscopy of protein folding, ChemPhysChem 6 (2005) 1206–1220. http://dx.doi.org/10.1002/cphc.200400609.
[93] E. Rhoades, M. Cohen, B. Schuler, G. Haran, Two-state folding observed in individual protein molecules, J. Am. Chem. Soc. 126 (2004) 14686–14687. http://dx.doi.org/10.1021/ja046209k.
[94] M. Oliveberg, P.G. Wolynes, The experimental survey of protein-folding energy landscapes, Q. Rev. Biophys. 38 (2005) 245–288. http://dx.doi.org/10.1017/S0033583506004185.
[95] H.S. Chan, K.A. Dill, Protein folding in the landscape perspective: Chevron plots and non-arrhenius kinetics, Proteins Struct. Funct. Bioinform. 30 (1998) 2–33. http://dx.doi.org/10.1002/(SICI)1097-0134(19980101)30:1<2::AID-PROT2>3.0.CO;2-R.
[96] M.J. Lindberg, J. Normark, A. Holmgren, M. Oliveberg, Folding of human superoxide dismutase: Disulfide reduction prevents dimerization and produces marginally stable monomers, Proc. Natl. Acad. Sci. USA 101 (2004) 15893–15898. http://dx.doi.org/10.1073/pnas.0403979101.
[97] F. Liu, Y.G. Gao, M. Gruebele, A survey of [lambda] repressor fragments from two-state to downhill folding, J. Mol. Biol. 397 (2010) 789–798. http://dx.doi.org/10.1016/j.jmb.2010.01.071.