# PROTEIN FOLDING IN THE ABSENCE OF A CLEAR FREE-ENERGY BARRIER*

Anders Irbäck

Complex Systems Division, Department of Theoretical Physics, Lund University
Sölvegatan 14A, 223 62 Lund, Sweden

*(Received August 5, 2003)*

Many small proteins fold in a two-state manner, the rate-limiting step being the passage of the free-energy barrier separating the unfolded state from the native one. The free-energy barrier is, however, weak or absent for the fastest-folding proteins. Here a simple diffusion picture for such proteins is discussed. It is tested on a model protein that makes a three-helix bundle. Assuming the motion along individual reaction coordinates to be diffusive on timescales beyond the reconfiguration time for a single helix, it is found that the relaxation time can be predicted within a factor of two. It is also shown that melting curves for this protein to a good approximation can be described in terms of a simple two-state system, despite the absence of a clear free-energy barrier.

PACS numbers: 87.15.Aa

## 1. Introduction

The folding of proteins to their functional states is a remarkable process [1]. In the cell, the folding process may require the assistance from helper molecules. However, as shown by refolding experiments, many proteins have the ability to fold spontaneously to their native states. This implies that the amino acid sequence contains all the information needed for the formation of the functional state [2]. The questions of how the folding process takes place and how the structure is encoded in the sequence are fascinating and in the focus of both experimental and theoretical research.

Many small single-domain proteins share the common property of folding in a two-state manner, without significantly populating any meta-stable intermediate state [3]. It is tempting to interpret the apparent two-state behaviour of these proteins in terms of a simple free-energy landscape with two

* Presented at the Workshop on Random Geometry, Kraków, Poland, May 15–17, 2003.

minima separated by a single barrier, where the minima represent the native and unfolded states, respectively. If the barrier is high, this picture provides an explanation of why the folding kinetics tend to be single exponential, and why the folding thermodynamics show two-state character.

However, it is well-known that the free-energy barrier, $\Delta F$, is not high for all these proteins. In fact, assuming the folding time $\tau_f$ to be given by $\tau_f = \tau_0 \exp(\Delta F/kT)$ with $\tau_0 \sim 1\,\mu$s [4], it is easy to find examples of proteins with $\Delta F$ values of a few $kT$ [3] ($k$ is Boltzmann's constant and $T$ the temperature).

Suppose the native and unfolded states coexist at the folding temperature and that there is no well-defined intermediate state, but that a clear free-energy barrier between the native and unfolded states is missing. What type of relaxation behaviour should one then expect? Furthermore, would such a protein show easily detectable deviations from thermodynamic two-state behaviour? To gain insights into these questions, our group recently performed a Monte Carlo (MC) study of a designed three-helix-bundle protein [5]. Inspired by energy-landscape theory (for a recent review, see [6,7]), we compared the calculated relaxation time for this protein with predictions from a simple one-dimensional diffusion picture.

The paper is organised as follows. In Sec. 2 the diffusion analysis is discussed. In Sec. 3 our MC study of the thermodynamics and kinetics of the three-helix-bundle protein is presented. A brief summary can be found in Sec. 4.

## 2. Diffusion picture

In the energy-landscape approach [6, 7], the high-dimensional folding process is projected onto one or a few coordinates; in its simplest form, the folding process is modelled as one-dimensional Brownian motion in an external potential $F(r) = -kT \ln P_{\rm eq}(r)$, where $r$ is the reaction coordinate studied and $P_{\rm eq}(r)$ denotes the equilibrium distribution of $r$. The probability distribution of $r$ at time $t$, $P(r,t)$, then obeys Smoluchowski's diffusion equation

$$\frac{\partial P(r,t)}{\partial t} = \frac{\partial}{\partial r}\left[D(r)\left(\frac{\partial P(r,t)}{\partial r} + \frac{P(r,t)}{kT}\frac{\partial F(r)}{\partial r}\right)\right], \qquad (1)$$

where $D(r)$ is the diffusion coefficient.

Due to the projection onto a single reaction coordinate $r$, Eq. (1) is not expected to hold on short timescales. It may still be useful if the motion in $r$ is diffusive beyond some timescale that is small compared with the relaxation time. It is then possible to predict the relaxation time from this equation. In [8] such an analysis was successfully carried through for a lattice model protein.

If the free energy $F(r)$ has the shape of a double well with a clear barrier in between, Eq. (1) predicts single-exponential relaxation, with a rate constant given by Kramers' well-known formula [9, 10]. However, this result cannot be applied to systems that lack a clear free-energy barrier. To compare the behaviour of our fast-folding model protein with that predicted by Eq. (1), we therefore solved this equation numerically [5], by a finite-difference scheme of Crank–Nicolson type. The analysis was carried out using the full $F(r)$ and $D(r)$, as obtained from simulations (see below).

In the idealised situation when $F(r)$ has the shape of a square well and $D(r)$ is a constant, Eq. (1) can be solved in a closed form. It is instructive to take a look at this solution. Suppose, for convenience, that the reaction coordinate $r$ is the energy $E$. The equilibrium distribution is then given by $P_{\mathrm{eq}}(E) \propto \exp(-\delta\beta E)$ if $E$ is in the square well and $P_{\mathrm{eq}}(E) = 0$ otherwise, where $\delta\beta = 1/kT - 1/kT_{\mathrm{f}}$, $T_{\mathrm{f}}$ being the folding temperature. With this $P_{\mathrm{eq}}(E)$, Eq. (1) takes the form

$$\frac{\partial P(E,t)}{\partial t} = \frac{\partial}{\partial E}\left[D\left(\frac{\partial P(E,t)}{\partial E} + \delta\beta P(E,t)\right)\right], \tag{2}$$

where the diffusion coefficient is assumed constant, $D(E) = D$. The initial distribution $P(E, t = 0)$ is taken to be the equilibrium distribution at an arbitrary temperature $T_0$.

By separation of variables, it is straightforward to solve Eq. (2) with this initial condition for $P(E,t)$, the energy distribution at time $t$. The average energy at time $t$, $E(t)$, is found to be

$$E(t) = \langle E\rangle + \sum_{k=1}^{\infty} A_k \mathrm{e}^{-t/\tau_k}, \tag{3}$$

where the time constants $\tau_k$ are given by

$$\frac{1}{\tau_k} = \frac{D}{\Delta E_{\mathrm{sw}}^2}\left(\pi^2 k^2 + \tfrac{1}{4}\delta\beta^2 \Delta E_{\mathrm{sw}}^2\right), \tag{4}$$

$\Delta E_{\mathrm{sw}}$ being the width of the square-well potential. Expressions for the equilibrium average $\langle E\rangle$ and the expansion coefficients $A_k$ can be found in [5].

For a general reaction coordinate $r$, Eq. (4) remains valid at $T = T_{\mathrm{f}}$; that is, $\tau_k = \Delta r_{\mathrm{sw}}^2/D\pi^2 k^2$, where $\Delta r_{\mathrm{sw}}$ is the width of the assumed square-well potential in $r$. However, for a general $r$, the temperature dependence is not as simple as that in Eq. (4).

Two properties of the expansion coefficients $A_k$ are worth mentioning. First, $A_k$ scales as $k^2$ if $k \ll \frac{1}{2\pi}|\delta\beta|\Delta E_{\mathrm{sw}}$, and as $1/k^4$ if $k \gg \frac{1}{2\pi}|\delta\beta|\Delta E_{\mathrm{sw}}$.

Second, all $A_k$ with even $k$ are suppressed if $T$ is close to $T_\mathrm{f}$; in fact, they vanish if $T = T_\mathrm{f}$. From these two facts it follows that $|A_1|$ is much larger than the other $|A_k|$ if $T$ is near $T_\mathrm{f}$. This tends to make the deviation from single-exponential behaviour smaller than one might expect from Eq. (4). At the same time, it should be pointed out that the exact vanishing of $A_k$ for even $k$ at $T = T_\mathrm{f}$ is accidental and related to the perfect symmetry in this particular case (square-well potential and constant diffusion coefficient).

In Sec. 3.3 the relaxation time of our three-helix-bundle protein is compared both to the numerical solution of Eq. (1) and to the analytical solution of the simplified Eq. (2).

## 3. Monte Carlo study of a designed three-helix bundle

### *3.1. Model and methods*

The three-helix-bundle protein was studied using a reduced off-lattice model, introduced in [11]. In this model, each amino acid is represented by five or six atoms, three of which are the backbone atoms N, $C_\alpha$ and $C'$. Also included are the H and O atoms of the peptide units, which are used to define hydrogen bonds. The sixth atom is a large $C_\beta$ which represents the side chain. The $C_\beta$ atom is taken to be either hydrophobic, polar or absent, which gives us three types of amino acids: H with hydrophobic $C_\beta$, P with polar $C_\beta$, and G (glycine) without $C_\beta$. All bond lengths, bond angles and peptide torsion angles (180°) are held fixed, which means that the model contains two degrees of freedom per amino acid, the Ramachandran torsion angles $\phi$ and $\psi$.

The potential function

$$E = E_\mathrm{loc} + E_\mathrm{ev} + E_\mathrm{hb} + E_\mathrm{hp} \tag{5}$$

is composed of four terms. The local potential $E_\mathrm{loc}$ has a standard form with threefold symmetry,

$$E_\mathrm{loc} = \frac{\varepsilon_\phi}{2} \sum_i (1 + \cos 3\phi_i) + \frac{\varepsilon_\psi}{2} \sum_i (1 + \cos 3\psi_i)\,. \tag{6}$$

The excluded-volume term $E_\mathrm{ev}$ is given by a hard-sphere potential of the form

$$E_\mathrm{ev} = \varepsilon_\mathrm{ev} {\sum_{i<j}}' \left(\frac{\sigma_{ij}}{r_{ij}}\right)^{12}, \tag{7}$$

where the sum runs over all possible atom pairs except those consisting of two hydrophobic $C_\beta$. The parameter $\sigma_{ij}$ is given by $\sigma_{ij} = \sigma_i + \sigma_j + \Delta\sigma_{ij}$, where $\Delta\sigma_{ij} = 0.625\,\text{Å}$ for $C_\beta C'$, $C_\beta N$ and $C_\beta O$ pairs that are connected by

three covalent bonds, and $\Delta\sigma_{ij} = 0\,\text{Å}$ otherwise. The introduction of the parameter $\Delta\sigma_{ij}$ can be thought of as a change of the local potential.

The hydrogen-bond term $E_{\text{hb}}$ has the form

$$E_{\text{hb}} = \varepsilon_{\text{hb}} \sum_{ij} u(r_{ij}) v(\alpha_{ij}, \beta_{ij})\,, \tag{8}$$

where the functions $u(r)$ and $v(\alpha, \beta)$ are given by

$$u(r) \;=\; 5\left(\frac{\sigma_{\text{hb}}}{r}\right)^{12} - 6\left(\frac{\sigma_{\text{hb}}}{r}\right)^{10}, \tag{9}$$

$$v(\alpha, \beta) \;=\; \begin{cases} \cos^2\alpha \cos^2\beta & \alpha, \beta > 90^\circ \\ 0 & \text{otherwise}\,. \end{cases} \tag{10}$$

The sum in Eq. (8) runs over all possible HO pairs, and $r_{ij}$ denotes the HO distance, $\alpha_{ij}$ the NHO angle, and $\beta_{ij}$ the HOC′ angle.

The last term of the potential, $E_{\text{hp}}$, is an effective hydrophobic attraction given by

$$E_{\text{hp}} = \varepsilon_{\text{hp}} \sum_{i<j} \left[ \left(\frac{\sigma_{\text{hp}}}{r_{ij}}\right)^{12} - 2\left(\frac{\sigma_{\text{hp}}}{r_{ij}}\right)^{6} \right], \tag{11}$$

where the sum runs over all possible pairs of hydrophobic $C_\beta$.

To speed up the calculations, a cutoff radius $r_{\text{c}}$ is used, which is taken to be 4.5 Å for $E_{\text{ev}}$ and $E_{\text{hb}}$, and 8 Å for $E_{\text{hp}}$. Numerical values of all energy and geometry parameters can be found in [11].

A slightly extended version of this model, with five amino acids rather than three, has been applied to the three-helix-bundle B domain of staphylococcal protein A [12] and two related sequences [13].

The idealised three-helix-bundle protein studied here contains 54 amino acids and is a truncated three-letter version [14, 15] of a four-helix-bundle protein *de novo* designed by Regan and DeGrado [16]. It consists of three identical stretches of H and P amino acids, connected by two GGG segments. The HP segment is given by PPHPPHHPPHPPHHPP and is such that it can make an $\alpha$-helix with all H on the same side.

The thermodynamic behaviour of this sequence was studied by using simulated tempering [17–19], in which the temperature is a dynamic variable. This method was used in order to speed up the calculations at low temperatures. For a review of simulated tempering and other generalised-ensemble techniques for protein folding, see [20]. Our simulations were started from random configurations. Two kinds of conformation moves were used: first, the pivot move in which a single torsion angle is turned; and second, a semi-local method [21] that works with seven or eight adjacent torsion angles,

which are turned in a coordinated manner. The non-local pivot move was included in order to accelerate the evolution of the system at high temperatures.

An MC-based kinetic study was performed, too. These simulations are only meant to mimic the time evolution of the system in a qualitative sense. The kinetic simulations differ from the thermodynamic ones in two ways: first, the temperature was held constant; and second, the non-local pivot update was not used, but only the semi-local method [21]. This restriction is needed in order to avoid large unphysical deformations of the chain.

### *3.2. Thermodynamics*

It turns out that this designed sequence does make a stable three-helix bundle in this model, except for a twofold topological degeneracy. Fig. 1 is a schematic illustration of representative structures for the two topologies, as obtained by energy minimisation. The difference between the two topologies is that if one lets the first two helices form a U, then the third helix is either in front of (FU) or behind (BU) that U. In order for the model to be able to discriminate between these states, it would probably be necessary to change the hydrophobicity potential $E_{\rm hp}$. A simple pairwise additive potential like that in Eq. (11) has problems with this task because the contact patterns are very similar in the two topologies [23]. A measure of structural similarity with the (degenerate) native state is provided by the parameter

$$Q = \max\left[\exp\left(-\frac{\delta_{\rm FU}^2}{(10\text{Å})^2}\right), \exp\left(-\frac{\delta_{\rm BU}^2}{(10\text{Å})^2}\right)\right], \tag{12}$$

where $\delta_{\rm FU}$ and $\delta_{\rm BU}$ are the root-mean-square deviations from the ideal FU and BU conformations in Fig. 1 (calculated over all backbone atoms).

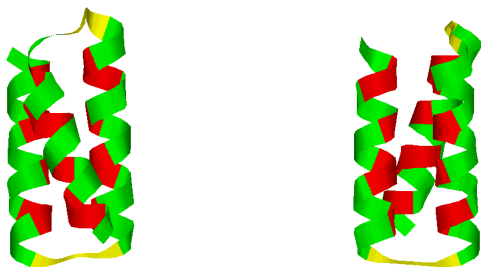


Fig. 1. Representative structures for the two topologies, FU and BU. Drawn with RasMol [22].

The thermodynamic behaviour of this model protein was studied in detail in [5, 11]. In particular, it was found to have the following properties:

- Its helices are more stable than those of the corresponding one- and two-helix segments, which is in agreement with the well-known fact that secondary-structure elements in general are less stable in isolation than as parts of a full protein.

- It undergoes an abrupt folding transition from an expanded state to the three-helix-bundle state, without forming any well-defined intermediate state. The temperature dependence of quantities such as the hydrogen-bond energy and the radius of gyration can be quite well described in terms of a simple two-state system, as illustrated in Fig. 2. This figure also shows that helix formation and chain collapse occur in parallel for this protein.

- It has no clear free-energy barrier between the unfolded and native states. Fig. 3 shows the free-energy profiles $F(E)$ and $F(Q)$ at $T = T_{\mathrm{f}}$. $F(Q)$ exhibits a very weak barrier, $< 1\,kT$, whereas $F(E)$ shows no barrier at all. This clearly demonstrates that at a simple two-state description is an oversimplification, despite that the melting curves show approximate two-state character (see Fig. 2).

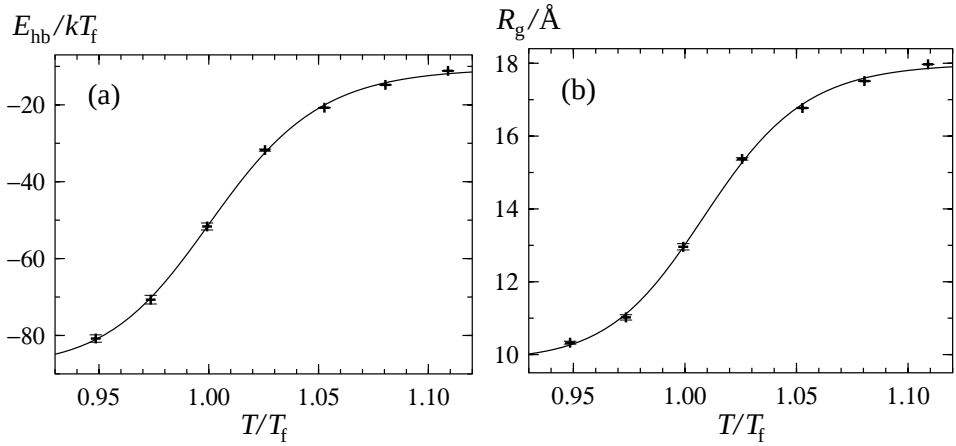



Fig. 2. Temperature dependence of (a) the hydrogen-bond energy $E_{\mathrm{hb}}$ and (b) the radius of gyration $R_{\mathrm{g}}$. The lines are fits to the two-state expression $X(T) = [X_{\mathrm{u}} + X_{\mathrm{n}}K(T)]/[1 + K(T)]$, where $K(T) = \exp[(1/kT - 1/kT_{\mathrm{m}})\Delta E]$ is the effective equilibrium constant and $X_{\mathrm{n}}$ and $X_{\mathrm{u}}$ denote the native and unfolded values of $X$, respectively. Such a fit has four parameters: the energy difference $\Delta E$, the melting temperature $T_{\mathrm{m}}$, and the two baselines $X_{\mathrm{u}}$ and $X_{\mathrm{n}}$.

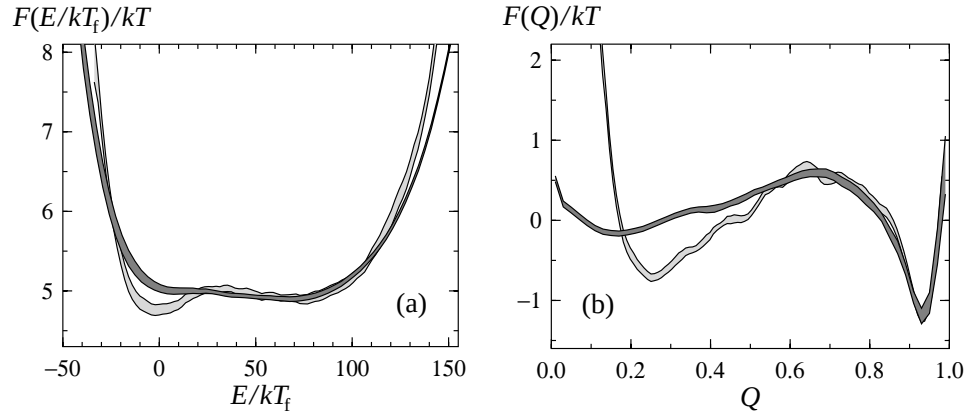



Fig. 3. Free-energy profiles at $T = T_{\rm f}$ for (a) the energy $E$ and (b) the similarity parameter $Q$ (dark bands). The light-grey bands show free energies for block averages (see Eq. 13), using a block size of $\tau_b = 10^6$ MC steps. Each band is centered around the expected value and shows statistical $1\sigma$ errors.

It should be stressed that the behaviour of the model depends strongly on the parameters $\varepsilon_{\rm hb}$ and $\varepsilon_{\rm hp}$ that sets the strengths of the hydrogen bonds and the hydrophobic attraction, respectively. These parameters must be carefully chosen in order for the folding transition to be first-order-like [24].

### *3.3. Kinetics*

In our MC-based kinetic study, the relaxation of ensemble averages of various quantities was studied at $T = T_{\rm f}$. For this purpose, a set of 3 000 folding simulations was performed, starting from equilibrium conformations at the temperature $T_0 \approx 1.06\,T_{\rm f}$. At this temperature, the chain is extended with a relatively low helix content (see Fig. 2).

Fig. 4 shows the relaxation behaviour of the energy $E$ and the similarity parameter $Q$ [see Eq. (12)]. Fits of the large-time data to an exponential give relaxation times of $\tau \approx 1.7 \times 10^7$ and $\tau \approx 1.8 \times 10^7$ for $E$ and $Q$, respectively, in units of elementary MC steps.

These calculated relaxation times were compared with predictions from the diffusion picture discussed in Sec. 2. For this purpose, it is necessary to perform a coarse-graining in time, since the behaviour is not expected to be diffusive on short time scales. A convenient way to implement that is to consider block averages $b(t)$ defined by

$$b(t) = \frac{1}{\tau_b} \sum_{t \le s < t+\tau_b} r(s)\,, \qquad t = 0, \tau_b, 2\tau_b, \ldots\,, \tag{13}$$

where $\tau_b$ is the block size and $r$ is the reaction coordinate considered. The block size was taken to be $\tau_b = 10^6$ MC step, corresponding to the reconfiguration time for an individual helix [5]. Using the block variables, the diffusion coefficient was estimated by using $D_{\rm b}(r) = \langle(\delta b)^2\rangle_r/2\tau_b$. The free energy $F_{\rm b}(r)$ for the block averages was calculated too, and was found to be similar to that for the unblocked variables, as can be seen from Fig. 3. Having obtained $D_{\rm b}(r)$ and $F_{\rm b}(r)$, the relaxation time was calculated in two ways. The first estimate, $\tau_{\rm pred,0}$, was obtained by using a square-well approximation of $F_{\rm b}(r)$ and a constant $D_{\rm b}(r) = D_{\rm b}$. The second estimate, $\tau_{\rm pred}$, was obtained by numerical solution of Eq. (1), using the full $F_{\rm b}(r)$ and $D_{\rm b}(r)$.

The results of this analysis are summarised in Table I. From this table it can be seen that the simple estimate $\tau_{\rm pred,0}$ is correct within a factor of two for both $E$ and $Q$. This is encouraging, but should not be taken to suggest that the underlying diffusion picture is perfect. If it had been perfect, the more elaborate estimate, $\tau_{\rm pred}$, would have agreed with the observed value $\tau$, which is not the case. In fact, $\tau_{\rm pred}$ is not better than $\tau_{\rm pred,0}$, at least not in $Q$, despite that there is a weak barrier in this coordinate (see Fig. 3(b)). That this one-dimensional description of the folding process is not perfect is no surprise, given that it completely ignores non-Markovian effects. How non-Markovian effects may affect folding times has been discussed in [7,25]. Another way to refine the analysis would be to use a set of two or more reaction coordinates rather than a single one [7,26,27]. With a multidimensional representation of the folding process, non-Markovian effects could become smaller.

TABLE I

The predictions $\tau_{\rm pred,0}$ and $\tau_{\rm pred}$ (see text) along with the observed relaxation time $\tau$, for the energy $E$ and the similarity parameter $Q$. $\Delta r_{\rm sw}$ is the width of the square-well potential and $D_{\rm b}$ is the average diffusion coefficient.

| | $\Delta r_{\rm sw}$ | $D_{\rm b}$ | $\tau_{\rm pred,0}$ | $\tau_{\rm pred}$ | $\tau$ |
|---|---|---|---|---|---|
| $E$ | $140\,kT_{\rm f}$ | $(9.3 \pm 0.2) \times 10^{-5}(kT_{\rm f})^2$ | $2.1 \times 10^7$ | $1.9 \times 10^7$ | $1.7 \times 10^7$ |
| $Q$ | 1.0 | $(1.00 \pm 0.02) \times 10^{-8}$ | $1.0 \times 10^7$ | $0.8 \times 10^7$ | $1.8 \times 10^7$ |

The relaxation time analysis indicates that the dynamics are approximately diffusive on timescales beyond $10^6$ MC steps $\sim \tau/20$. If the potential is close to a square well and the diffusion coefficient approximately constant, Eqs. (3) and (4) suggest that the leading correction term to the asymptotic exponential behaviour should be a second exponential with a time constant of $\tau_2 = \tau/4$ at $T = T_{\rm f}$ (unless $A_2$ is very small). A look at the data in

Fig. 4 shows that $Q(t)$ is approximately single exponential down to very small $t$, whereas there are non-negligible deviations from this behaviour in $E(t)$ below $t \sim \tau/3$. The data for $E(t)$ can be well described by a double exponential with time constants that differ by a factor of 4, but drawing any firm conclusion about the value of the second time constant is impossible, due to limited statistics. There is, however, a recent experimental study [28] of fast-folding mutants of the five-helix-bundle protein $\lambda_{6-85}$, in which double-exponential fits were performed near $T = T_{\mathrm{f}}$ (for the mutant $\lambda$Q33Y). Although it could be accidental, it is interesting to note that the two fitted time constants differed by a factor close to 4, as predicted by Eq. (4).

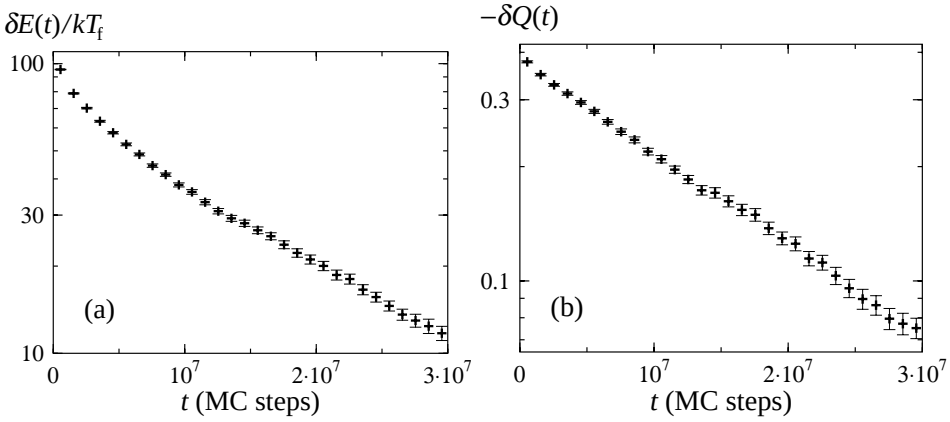



Fig. 4. Relaxation behaviour at the folding temperature $T_{\mathrm{f}}$, starting from $T_0 \approx 1.06\,T_{\mathrm{f}}$. (a) $\delta E(t) = E(t) - \langle E \rangle$ against simulation time $t$, where $E(t)$ is the average $E$ after $t$ MC steps (3 000 runs) and $\langle E \rangle$ denotes the equilibrium average. (b) Same plot for the similarity parameter $Q$.

## 4. Summary

In this paper, a simple diffusion-based theory for fast-folding proteins was discussed. It was tested against MC results for a three-helix-bundle protein, which were obtained using a reduced off-lattice model with a relatively detailed chain representation. The main findings were as follows.

- Assuming the motion in individual reaction coordinates to be diffusive on timescales beyond the reconfiguration time for a single helix, the relaxation time could be predicted within a factor of two.

- The closed-form solution for a square-well potential and a constant diffusion coefficient predicts that the leading corrections to the asymptotic exponential behaviour for large $t$ are exponentials with time constants of $\tau_2 = \tau/4$ and $\tau_3 = \tau/9$ at $T = T_\mathrm{f}$. Due to statistical limitations, this could not be tested on our model protein, but a double-exponential fit of recent experimental data for a fast-folding protein actually gave time constants differing by a factor close to 4 [28]. Whether that was accidental or not remains to be seen.

- Although the relaxation time could be predicted quite well, it is clear that the one-dimensional diffusion description leaves room for improvement, as could be seen from our numerical solution of the diffusion equation.

This work was in part supported by the Swedish Research Council.

## REFERENCES

[1] C. Branden, J. Tooze, *Introduction to Protein Structure*, Garland Publishing, New York 1991.

[2] C.B. Anfinsen, *Science* **181**, 223 (1973).

[3] S.E. Jackson, *Fold. Des.* **3**, R81 (1998).

[4] S.J. Hagen, J. Hofrichter, A. Szabo, W.A. Eaton, *Proc. Natl. Acad. Sci. USA* **93**, 11615 (1996).

[5] G. Favrin, A. Irbäck, B. Samuelsson, S. Wallin, *Biophys. J.* **85**, 1457 (2003).

[6] S.S. Plotkin, J.N. Onuchic, *Q. Rev. Biophys.* **35**, 111 (2002).

[7] S.S. Plotkin, J.N. Onuchic, *Q. Rev. Biophys.* **35**, 205 (2002).

[8] N.D. Socci, J.N. Onuchic, P.G. Wolynes, *J. Chem. Phys.* **104**, 5860 (1996).

[9] H.A. Kramers, *Physica* **7**, 284 (1940).

[10] J.D Bryngelson, J.N. Onuchic, N.D. Socci, P.G. Wolynes, *Proteins* **21**, 167 (1995).

[11] A. Irbäck, F. Sjunnesson, S. Wallin, *Proc. Natl. Acad. Sci. USA* **97**, 13614 (2000).

[12] G. Favrin, A. Irbäck, S. Wallin, *Proteins* **47**, 99 (2002).

[13] G. Favrin, A. Irbäck, S. Wallin, *Proteins* (2003), in press.

[14] Z. Guo, D. Thirumalai, *J. Mol. Biol.* **263**, 323 (1996).

[15] S. Takada, Z. Luthey-Schulten, P.G. Wolynes, *J. Chem. Phys.* **110**, 11616 (1999).

[16] L. Regan, W.F. DeGrado, *Science* **241**, 976 (1988).

[17] A.P. Lyubartsev, A.A. Martsinovski, S.V. Shevkunov, P.N. Vorontsov-Velyaminov, *J. Chem. Phys.* **96**, 1776 (1992).

[18] E. Marinari, G. Parisi, *Europhys. Lett.* **19**, 451 (1992).

[19] A. Irbäck, F. Potthast, *J. Chem. Phys.* **103**, 10298 (1995).

[20] U.H.E. Hansmann, Y. Okamoto, *Curr. Opin. Struct. Biol.* **9**, 177 (1999).

[21] G. Favrin, A. Irbäck, F. Sjunnesson, *J. Chem. Phys.* **114**, 8154 (2001).

[22] R. Sayle, E.J. Milner-White, *Trends Biochem. Sci.* **20**, 374 (1995).

[23] S. Wallin, J. Farwer, U. Bastolla, *Proteins* **50**, 144 (2003).

[24] A. Irbäck, F. Sjunnesson, S. Wallin, *J. Biol. Phys.* **27**, 169 (2001).

[25] S.S. Plotkin, P.G. Wolynes, *Phys. Rev. Lett.* **80**, 5015 (1998).

[26] R. Du, V.S. Pande, A.Y. Grosberg, T. Tanaka, E.I. Shakhnovich, *J. Chem. Phys.* **108**, 334 (1997).

[27] N.D. Socci, J.N. Onuchic, P.G. Wolynes, *Proteins* **32**, 136 (1998).

[28] W.Y. Yang, M. Gruebele, *Nature* **423**, 193 (2003).