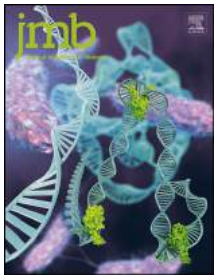


Ruggedness in the Free Energy Landscape Dictates Misfolding of the Prion Protein

Roumita Moulick[1], **Rama Reddy Goluguri**[1] **and Jayant B. Udgaonkar**[1,2]

*1 - **National Centre for Biological Sciences,** Tata Institute of Fundamental Research, Bengaluru 560065, India*
*2 - **Indian Institute of Science Education and Research,** Pashan, Pune 411008, India*

***Correspondence to Jayant B. Udgaonkar:** Indian Institute of Science Education and Research, Pashan, Pune 411008, India. jayant@iiserpune.ac.in*
https://doi.org/10.1016/j.jmb.2018.12.009
Edited by Sheena Radford

Abstract

Experimental determination of the key features of the free energy landscapes of proteins, which dictate their adeptness to fold correctly, or propensity to misfold and aggregate and which are modulated upon a change from physiological to aggregation-prone conditions, is a difficult challenge. In this study, sub-millisecond kinetic measurements of the folding and unfolding of the mouse prion protein reveal how the free energy landscape becomes more complex upon a shift from physiological (pH 7) to aggregation-prone (pH 4) conditions. Folding and unfolding utilize the same single pathway at pH 7, but at pH 4, folding occurs on a pathway distinct from the unfolding pathway. Moreover, the kinetics of both folding and unfolding at pH 4 depend not only on the final conditions but also on the conditions under which the processes are initiated. Unfolding can be made to switch to occur on the folding pathway by varying the initial conditions. Folding and unfolding pathways appear to occupy different regions of the free energy landscape, which are separated by large free energy barriers that change with a change in the initial conditions. These barriers direct unfolding of the native protein to proceed *via* an aggregation-prone intermediate previously identified to initiate the misfolding of the mouse prion protein at low pH, thus identifying a plausible mechanism by which the ruggedness of the free energy landscape of a protein may modulate its aggregation propensity.

Introduction

The free energy landscape of a protein describes all possible conformations of the protein: the monomeric unfolded (U), native (N), and intermediate (I) states, as well as aggregated states. The free energy landscape is encoded in the protein sequence [1,2] and is modulated by external factors such as pH, temperature, and solvent, which are also important in the maintenance of protein solubility and function. Under physiological conditions, aggregated states of proteins may be more stable than soluble, monomeric states [3]. Nevertheless, proteins conformationally convert into aggregated forms with surprising rarity *in vivo*, which suggests that large free energy barriers and other kinetic factors obviate their aggregation propensities [4,5]. However, the differences in the nature of intrinsic free energy barriers which separate the different monomeric soluble states of proteins, and which separate these states from aggregated states, remains unclear.

Energy barriers on a protein free energy surface originate from incomplete enthalpy-entropy compensation [6–9], which may be due to conformational frustration [10]. The foldability of the protein is determined by the heights of these energy barriers [11]. Large barriers on folding pathways can lead to the transient population of apparently discrete intermediate states. The partially formed structures of folding intermediates may make them prone to misfolding and aggregation, especially if the intermediates are long-lived. On the other hand, small energy barriers, which would be the result of near-perfect enthalpy-entropy compensation, lead to an apparently barrier-less folding process [12–15]. It has been suggested that very fast folding proteins, which fold in a down-hill, barrier-less manner, could be more aggregation-prone [16]. It has become clear

that the cooperativity of folding reactions has an important role in determining whether the protein will misfold or not.

The presence of a single dominant energy barrier, as against a multitude of small barriers, to folding, is often considered to be an evolutionary strategy to minimize the aggregation of proteins. Nevertheless, if only a single folding pathway were to be present, deleterious mutations might make the protein either unable to fold or more susceptible to misfolding [17]. Multiple folding routes not only make proteins more robust folders and less susceptible to the effects of mutations [18], but they also enhance the adaptability of proteins to variations in cellular demands and facilitates the acquisition of new traits during evolution [19]. A comprehensive understanding of how barrier heights or folding cooperativity, as well as multiplicity of folding routes or folding heterogeneity, modulate the aggregation propensities of proteins demands a comparative analysis of the free energy surface of a protein in soluble and in aggregation-prone conditions.

In the case of the prion protein, aggregation-prone intermediates have been identified in both folding [20–23] and unfolding [24] studies, as well as by pressure perturbation studies [25]. In the case of the mouse prion protein (moPrP), an early study could not detect any intermediate at physiological pH [26], but recent studies carried out at low pH have detected a molten globule-like, aggregation-prone intermediate, obligatory for the folding of the protein [27,28]. These studies indicated that the folding of moPrP is a barrier-limited process, with predominantly one folding pathway being operative. In contrast, single-molecule force spectroscopy studies have provided evidence for multiple misfolding pathways accessed from the unfolded state of the protein [29] and have indicated that the energy landscape is rugged.

The current study reports the results of kinetic folding studies that were carried out on the C-terminal globular domain of the mouse prion protein (CTD moPrP), to gain insights into the complexities of the free energy landscape under both physiological (pH 7) and aggregation-prone (pH 4) conditions. A fluorescence-sensitive variant of CTD moPrP (W144F/F174W CTD), dm CTD (Fig. 1) was used. At pH 7, both the folding and unfolding reactions were found to be describable, quantitatively, by a mechanism that incorporates a sequential, linear pathway on which at least two intermediate states are transiently populated. In marked contrast, at pH 4, at least two pathways were found to be utilized, with the unfolding pathway being distinct from the refolding pathway. Even more remarkably, the folding/unfolding pathway utilized depended on the initial conditions, that is, where on the energy landscape folding/unfolding was commenced from. Denaturant modulated the energy landscape, so that that the unfolding pathway could be made to become accessible during refolding, and the refolding pathway could be made to become accessible during unfolding. Importantly, the intermediate identified on the unfolding pathway appeared to be

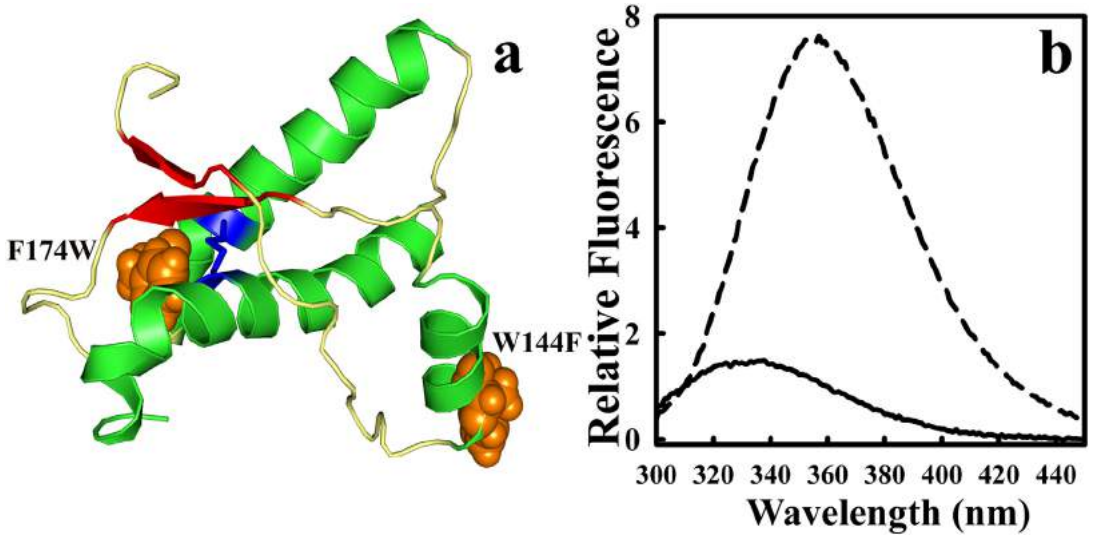


Fig. 1. Structural and spectroscopic characterization of dm CTD. (a) Residues W144 and F174 (orange) mutated to W144F/F174W moPrP (121-231) referred to as dm CTD. The α-helices are shown in green, β-strands in red, loops in pale yellow and the disulfide linking residues C178 and C213 in blue. The PDB ID is 1AG2. (b) Fluorescence spectrum of native (–) and unfolded (– -) dm CTD upon excitation at 295 nm at pH 4, 25 °C, normalized to the fluorescence signal of the native protein at 360 nm.

similar to the intermediate previously identified to be the monomeric precursor conformation from which misfolding initiates [30]. In summary, the current study experimentally validated several theoretically predicted features of energy landscapes of proteins, which have been rarely observed in earlier studies.

Results

Kinetic studies of dm CTD at pH 7

Figure 2a shows representative kinetic traces of the unfolding of native dm CTD, starting from 3 M urea, to final urea concentrations in the range of 5 to 6.8 M. Figure 2b shows a kinetic unfolding trace at 6.8 M urea with single (red) and double exponential (black) fits through the data. It is seen that this kinetic trace and all the other kinetic unfolding traces fit satisfactorily only to a double exponential equation. An initial lag phase was observed in all the kinetic unfolding traces, which suggested the existence of an intermediate with native-like fluorescence, which populates within ~50 μs of the initiation of unfolding.

dm CTD, unfolded in 7.9 M urea, was refolded at final urea concentrations ranging from 2.3 to 5 M. Figure 2c shows representative kinetic refolding traces, which, along with all other kinetic refolding traces, fit satisfactorily to a double exponential equation as evident from the residuals of a representative trace in Fig. 2d. At any urea concentration, the fluorescence value obtained by extrapolating the double exponential fit to zero time, was found to be higher than the fluorescence value obtained by linearly extrapolating the unfolded protein baseline of the equilibrium unfolding transition, to the same urea concentration.

Figure 3a is a plot of the dependence of the observed rate constants of the fast and slow phases obtained from double exponential fits to the unfolding and refolding kinetic traces, on urea concentration. It is important to note here that at the intermediate concentration of 5 M urea, the observed rate constant of unfolding of native dm CTD was identical to that obtained from refolding to 5 from 7.9 M urea. The initial and final fluorescence signals of all kinetic traces of unfolding were identical to those expected from equilibrium unfolding of the protein (Fig. 3b), indicating that all unfolding reactions were monitored till completion. In contrast, the initial signals of the kinetic traces of refolding were higher than those expected from the linear extrapolation of the unfolded protein baseline of the equilibrium unfolding study (Fig. 3b). This indicated that a refolding intermediate with a fluorescence signal higher than expected for the unfolded state in refolding conditions (which is given by the linearly extrapolated unfolded protein baseline) was populated during the burst phase of refolding. The final fluorescence signals of kinetic traces of refolding in concentrations of urea in the range of 2.3 to 4 M were in agreement with the equilibrium unfolding curve of the protein (Fig. 3b). However, the final fluorescence signal in refolding to intermediate urea concentrations in the range of 4 to 5 M was higher than expected (Fig. S1 in SI), suggesting that dm CTD refolds to native-like N* state in the sub-millisecond timescale from where it folds to N-state on a much slower timescale. This slow phase was not observed during the unfolding of native dm CTD to identical final urea concentrations.

Global analysis of kinetic data at pH 7

The kinetic traces of unfolding in urea at concentrations in the range of 5 to 6.8 M and that of refolding in the range of 2.3 to 4 M urea, along with the dependences of the observed rate constants of the fast and slow kinetic phases on urea concentration, were globally fit to a reduced three-state model (Scheme 2 in Materials and Methods). The solid lines in Fig. 3a represent the rate constants obtained from a numerical solution of the differential equations describing Scheme 2 (see Materials and Methods). The global fits of representative refolding and unfolding kinetic traces are shown in Fig. S2 (SI). The fluorescence values obtained for the intermediate states I_1 and I_2 were 0.9 ± 0.02 and 8.5 ± 0.6, indicating that I_1 had native-like fluorescence properties, whereas I_2 had unfolded-like fluorescence properties. The burst-phase increase in signal during refolding was accounted for by the very fast accumulation of I_2, as observed from the kinetic traces simulated at these urea concentrations using a four-state kinetic model (Fig. S2 in SI). The global analysis yielded the values of the microscopic rate constants defining the reduced three-state mechanism, as well as of the parameters, which define their denaturant dependences (Fig. 3). These parameters are listed in Table 1. The equilibrium constants obtained for each elementary step were used to calculate the fractional populations of N, I_1, I_2, and U at equilibrium at different denaturant concentrations, as shown in Fig. 3c. The fluorescence signal arising from the populations of the N, I_1, I_2, and U at any urea concentration was in good agreement with the fluorescent signal expected at that urea concentration from equilibrium unfolding studies (data not shown). Figure 3c indicates that the kinetic intermediates were populated substantially at equilibrium at concentrations of urea greater than 5 M. I_1 and I_2 are maximally populated to 20% and 17% at high urea concentrations. Interestingly, in a previous study, coincident equilibrium unfolding transitions monitored by circular dichroism and by fluorescence at 280-nm excitation had suggested that the unfolding transition

is two-state [30]. It seems that I_1, in addition to having N-like fluorescence properties, also has secondary structure similar to that of N. I_2, in addition to having U-like fluorescence properties, has secondary structural content similar to that of U. Hence, these intermediates could not be distinguished in the previous equilibrium unfolding study of dm CTD at pH 7 [30].

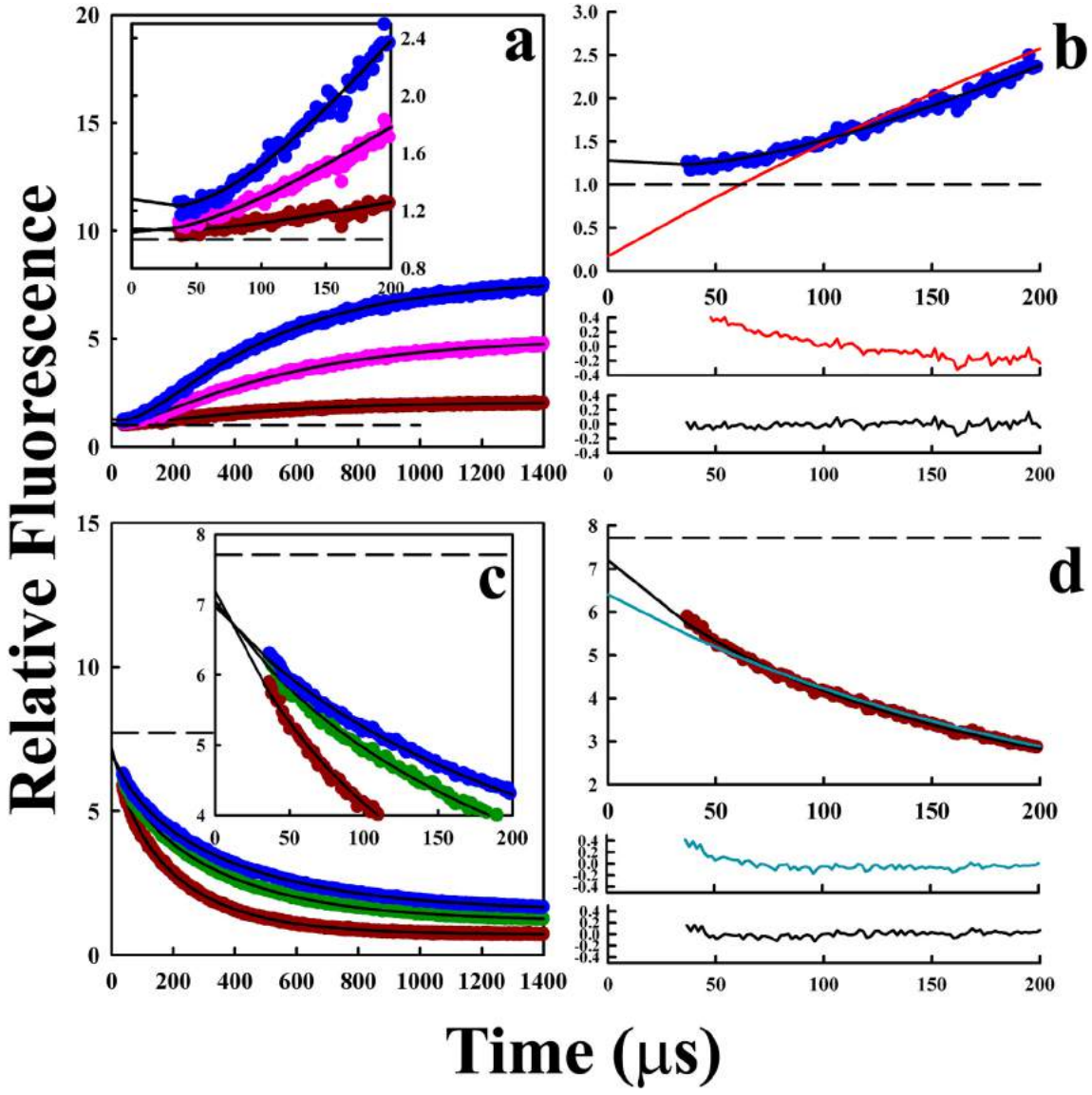


Fig. 2. Folding kinetics of dm CTD at pH 7. Effect of urea on the folding kinetics of dm CTD monitored by a change in intrinsic tryptophan fluorescence at 360 nm. (a) Unfolding kinetic traces: 3 → 5 M (●), 3 → 6 M (●), and 3 → 6.8 (●) urea (bottom to top); non-linear least square fits of the data to a double exponential equation (–) shown. Inset shows the initial parts of the reactions. (b) Representative unfolding kinetic trace in 6.8 M urea (●). Fits of the kinetic unfolding trace to a double (–) and a single exponential equation (—), respectively, with corresponding residuals (black: double and red: single exponential fit) shown. Fluorescence signal of native dm CTD in 3 M urea (- -) shown in panels a and b. (c) Refolding kinetic traces of dm CTD unfolded in 7.9 M urea: 7.9 → 4 M (●), 7.9 → 3.5 M (●), and 7.9 → 2.3 M (●) (top to bottom); non-linear squares fits of the data to a double exponential equation (–) shown. Inset shows the initial parts of the reactions. (d) Representative refolding kinetic trace in 2.3 M urea (●). Double (–) and single (—) exponential fits to the data with the corresponding residuals shown below (cyan: single and black: double exponential fit). Fluorescence signal of dm CTD unfolded in 7.9 M urea (- -) shown in panels c and d. All the kinetic refolding and unfolding traces have been normalized to the fluorescence of the native protein in 3 M urea.

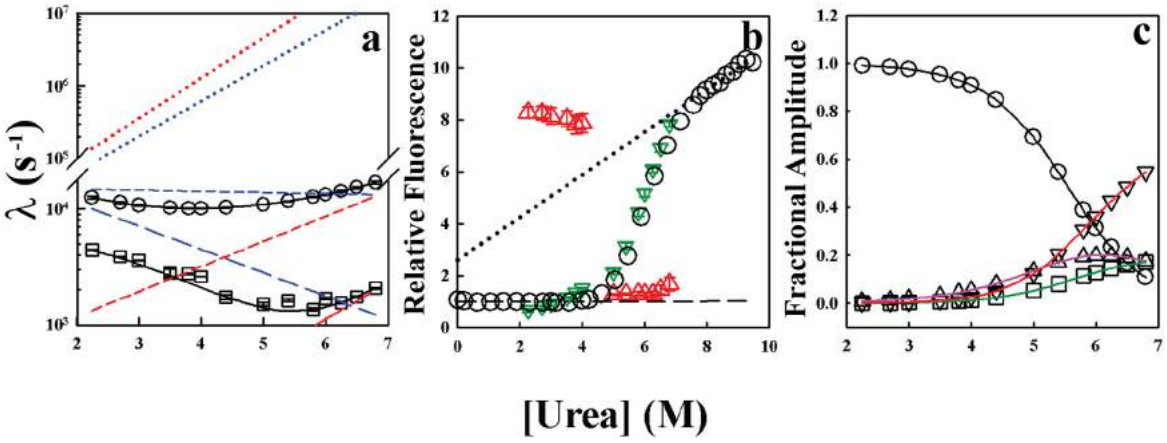


Fig. 3. Dependence of the rates and amplitudes of the fast and slow kinetic phases of folding of dm CTD at pH 7 on urea concentration. (a) Denaturant dependences of observed rate constants of the fast (○) and slow phases (□) of refolding and unfolding; Apparent rate constants of refolding and unfolding (–) predicted by global analysis of the kinetic data using Scheme 2 (Materials and Methods); global analysis also yielded rate constants of inter-conversion of I_2 and U in the denatured state ensemble; forward (- -) and backward (- -) microscopic rate constants of the N ↔ I_1 step; forward (---) and backward (---) microscopic rate constants of the I_1 ↔ I_2 step; forward (·····) and backward (·····) microscopic rate constants of the I_2 ↔ U step. (b) ○, Equilibrium unfolding transition relative to the N-state fluorescence in 3 M urea; △, $t = 0$ and ▽, $t = \infty$ fluorescence signals, in 2.3–4 M urea obtained from double exponential fits to the refolding kinetic traces; △, $t = 0$ and ▽, $t = \infty$ fluorescence signals, in 5–6.8 M urea obtained from double exponential fits to unfolding kinetic traces; fluorescence signals of N (– -) and U (·····) relative to the fluorescence signal at 3 M urea, respectively. (c) Dependence of fractional populations of N (○), intermediates I_1 (△) and I_2 (□), and U (▽) on urea concentration, determined from the equilibrium constants obtained from global analysis of the kinetic data.

Kinetic folding studies on dm CTD at pH 4

Destabilization of the native state of dm CTD by as much as 2.6 kcal mol^{-1} and an increase in the aggregation propensity of the protein at pH 4 are indirect indications of the potential influence pH might have on tuning the free energy landscape of the prion protein. At pH 4, an aggregation-prone, partially unfolded intermediate [30] is known to be selectively populated; hence, folding studies in such solvent conditions were expected to provide insights into whether this intermediate is obligatory for folding.

Kinetic unfolding and refolding experiments of native dm CTD were carried out at pH 4 at protein concentrations ranging from 6 to 24 μM. As shown in Fig. S3, folding kinetics at pH 4 was observed to be independent of protein concentration. Kinetic unfolding studies were carried out by unfolding the native protein at high urea concentrations in the range of 3 to 6.2 M. Since the rate constants obtained from kinetic unfolding data for native dm CTD starting from 0 to 1.9 M urea were identical (data not shown), all unfolding kinetic data reported here were obtained starting with native dm CTD in 1.9 M urea, which enabled the study of unfolding of the protein at higher urea concentrations. Figure 4a shows representative kinetic unfolding traces along with the fits to a double exponential equation. Figure 4b shows a representative kinetic trace of unfolding in 6.2 M

Table 1. Kinetic parameters obtained from a global analysis of the folding and unfolding kinetics of dm CTD at pH 7, using Scheme 2 (reduced from Scheme 1)

Unfolding				Refolding				Free energy (kcal mol^{-1})
Rate constants (s^{-1})		*m* value (kcal mol^{-1} M^{-1})		Rate constants (s^{-1})		*m* value (kcal mol^{-1} M^{-1})		ΔG^0_{NU} = 6.5
k^0_1	15.1 ± 3	m_1	0.43 ± 0.02	k^0_{-1}	28,790 ± 2000	m_{-1}	−0.28 ± 0.005	ΔG^0_1 4.5
k^0_2	435 ± 55	m_2	0.3 ± 0.01	k^0_{-2}	15,582 ± 1000	m_{-2}	−0.01 ± 0.002	ΔG^0_2 2.1
k^0_3	7820 ± 750	m_3	0.8 ± 0.05	k^0_{-3}	6967±680	m_{-3}	−0.7 ± 0.04	ΔG^0_3 −0.7

The global fits indicated that the fluorescence values of I_1 and I_2 relative to the fluorescence value of N at pH 7 are 0.9 ± 0.02 and 8.5 ± 0.6, respectively. The denaturant dependences, $s_{I_1}$and $s_{I_2}$of the fluorescence of I_1 and I_2, are 0.001 and 1.5 ± 0.2, respectively. ΔG^0_{NU}, ΔG^0_1, ΔG^0_2,and ΔG^0_3 represent the free energy change of unfolding from N to U, N to I_1, I_1 to I_2, and I_2 to U, respectively, in the absence of urea.

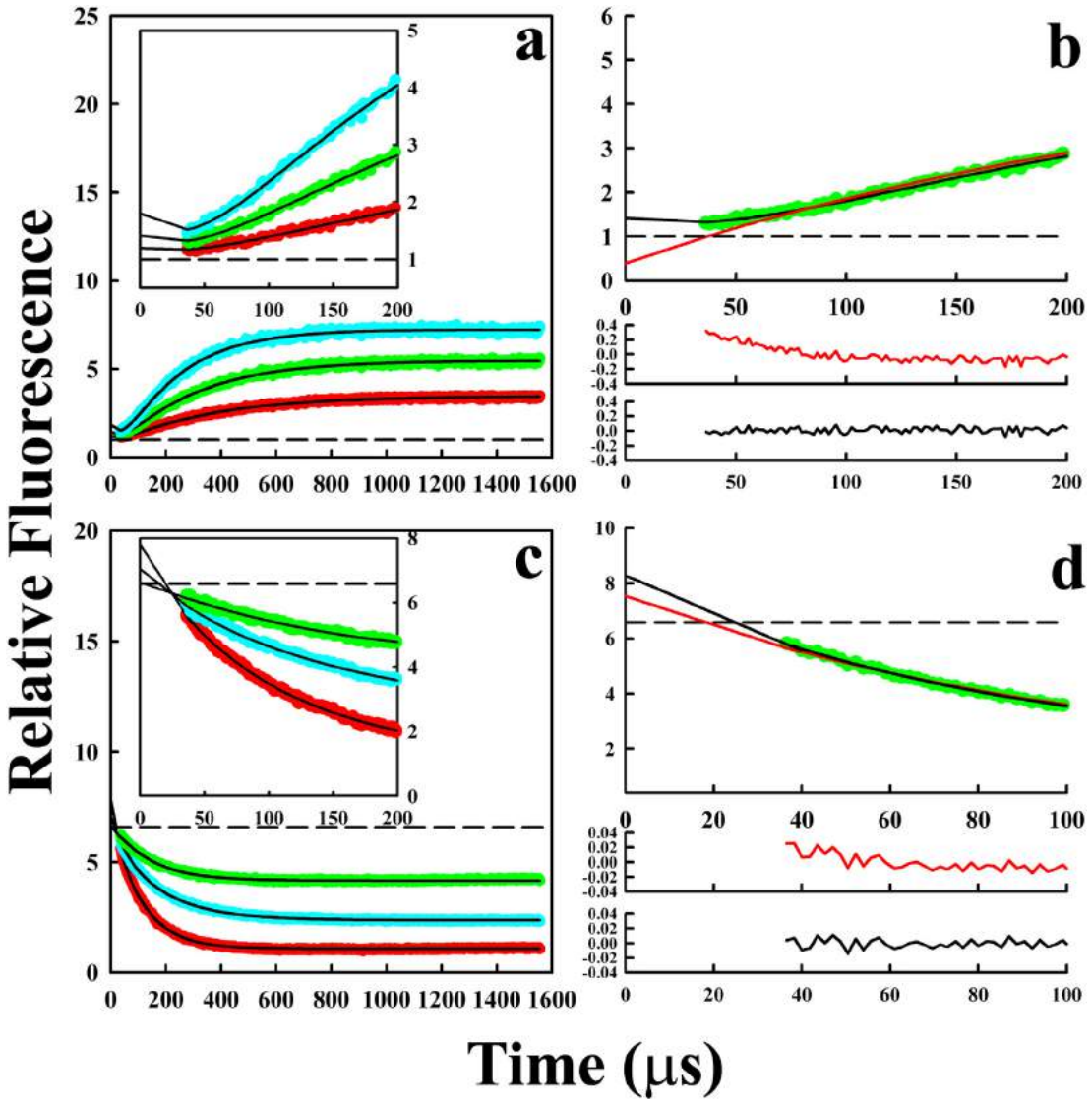


Fig. 4. Folding kinetics of dm CTD at pH 4. Effect of urea on the folding kinetics of dm CTD monitored by the change in intrinsic tryptophan fluorescence at 360 nm. (a) Kinetic unfolding traces: 1.9 → 4 M (●), 1.9 → 5 M (●), and 1.9 → 6.2 M (●) urea (bottom to top); non-linear least squares fits of the data (–) to a double exponential equation shown. Inset shows the initial parts of the reactions. (b) Representative unfolding kinetic trace in 5 M urea (●); fits of the kinetic unfolding trace to a double (–) and a single (—) exponential equation shown with the corresponding residuals (black: double and red: single exponential fit). (c) Kinetic refolding traces: 6 → 4.2 M (●), 6 → 3.7 (●) M and 6 → 1.7 M (●) urea (top to bottom); non-linear least-squares fit of the data (–) to a double exponential equation. Inset shows the initial parts of the reactions. (d) Representative kinetic refolding trace in 1.7 M urea (●); double (–) and single (—) exponential fits to the data with the corresponding residuals shown (upper: single, lower: double exponential fit). All kinetic refolding and unfolding traces have been normalized to the signal of the native protein in 1.9 M urea. - - shows relative fluorescence signals for native protein in 1.9 M urea in panels a and b and for unfolded protein in 6 M urea in panels c and d.

urea, where the presence of a lag phase is evident. In all kinetic unfolding traces, a lag phase that completes in ~50 μs was observed. From the single and double exponential fits to the data in Fig. 4b and from the residuals of both the fits, it is quite evident that the kinetic trace fits best to a double exponential equation, which accounted for the lag phase. This suggested that an intermediate with N-like fluorescence accumulates within 50 μs, which unfolded across a rate-limiting energy barrier to the denatured state ensemble. All the kinetic unfolding traces, as well as the dependences of the observed rate constants on urea concentration, obtained from fits to these traces, were fit globally to a reduced three-state unfolding mechanism (Scheme 2). Figure S4 (SI) shows representative simulated unfolding kinetic traces, which fit to the data. The microscopic rate constants of interconversion between the

Table 2. Kinetic parameters obtained from a global analysis of the unfolding kinetics of dm CTD at pH 4, using Scheme 2 (reduced from Scheme 1)

Unfolding				Refolding				Free energy (kcal mol^{-1})
Rate constants (s^{-1})		*m* value (kcal mol^{-1} M^{-1})		Rate constants (s^{-1})		*m* value (kcal mol^{-1} M^{-1})		ΔG^0_{NU} 4.7
k^0_1	$2 \times 10^3 \pm 66$	m_1	0.1 ± 0.04	k^0_{-1}	$1.9 \pm 0.06 \times 10^5$	m_{-1}	−0.5 ± 0.012	ΔG^0_1 2.7 ± 0.1
k^0_2	$1.4 \times 10^3 \pm 290$	m_2	0.232 ± 16	k^0_{-2}	$6.1 \pm 2.4 \times 10^4$	m_{-2}	−0.26 ± 0.01	ΔG^0_2 2.3 ± 0.1
k^0_3	42 ± 4	m_3	0.042 ± 0.013	k^0_{-3}	22 ± 2	m_{-3}	−0.043 ± 0.01	ΔG^0_3 −0.3 ± 0.1

The global fits indicated that the fluorescence values of I_U (I_1 in Scheme 2) and I_2 at pH 4 relative to the fluorescence value of the native state at pH 4 are 1.1 ± 0.07 and 2.8 ± 0.9, respectively. The denaturant dependences, s_{I_U} and $s_{I_2}$of the fluorescence of I_U and I_2 at pH 4, are 0.04 ± 0.01 and 1.0 ± 0.3, respectively. ΔG^0_{NU}, ΔG^0_1, ΔG^0_2,and ΔG^0_3 represent the free energy change of unfolding from N to U, N to I_U, I_U to I_2, and I_2 to U, respectively, in the absence of urea.

conformations populated on the unfolding pathway and their dependences on urea concentration are reported in Table 2. The equilibrium populations of N, unfolding intermediate I_U (I_1 in Scheme 2) and denatured state ensemble (I_2' which is D′ at pH 4) determined from the global analysis of the unfolding data (Table 2), predicted the experimentally observed equilibrium unfolding curve [30].

Kinetic refolding studies of dm CTD at pH 4 were carried out at urea concentrations in the range of 1.7 to 5 M. Representative kinetic refolding traces are shown in Fig. 4c. All refolding traces fit well to a double exponential equation, as seen for a representative kinetic trace of refolding to 1.7 M urea shown in Fig. 4d. Figure 4c and d shows that a burst-phase increase in fluorescence occurred during the 37-μs dead time of mixing, indicating that a hyperfluorescent intermediate (I_2') must be populated during the dead time of mixing. The hyperfluorescence phase is most prominent at lower urea concentrations where the extrapolated fluorescence signal at the initiation of the refolding reaction is maximum (Fig. S5, SI). Surprisingly, the hyperfluorescence signal observed at the initiation of refolding also increased with an increase in the concentration of urea in which dm CTD was denatured, as shown for refolding in 4 M urea of dm CTD that had been unfolded in 6.5, 7, and 7.5 M urea (Fig. S5, SI).

An important result was that when refolding and unfolding reactions were carried out in 4 M urea, starting from 6 and 1.9 M urea, respectively, the kinetics of the approach to equilibrium were not identical (Fig. 5a). Refolding was faster than unfolding. Figure 5b confirms that the rate constants of relaxation at 4 M urea were indeed different when the relaxation was started from U or from N. Figure 5c shows that the observed rate constant of the fast phase, which defines chevron λ_1, did not depend on the starting concentration of urea, but that the observed rate constant of the slow phase did depend on the starting concentration of urea. The relaxation rate constants observed for refolding starting from protein unfolded in 6 M urea defined chevron λ_2, and those observed for unfolding starting from native protein in 1.9 M urea defined chevron λ_3 (Fig. 5c). Figure 5d shows that the fractional amplitudes of the fast and slow phases of folding and unfolding did not depend on the starting concentration of urea. The unexpected observation that the relaxation rate constants at the same urea concentration depended on whether the relaxation was commenced from U or from N suggests that the folding and unfolding pathways are different at least in the range of urea concentrations where both can be studied. Consequently, it was not possible to fit both the folding and unfolding data to a single pathway of folding and unfolding (Fig. S4, SI).

Kinetic unfolding studies on dm CTD pre-equilibrated in 3 or 4 M urea at pH 4

To further validate the observation that the relaxation kinetics at intermediate urea concentrations depended on the initial conditions from which the jump in urea concentration was affected, the kinetics of unfolding in 5.6 M urea was studied when unfolding was commenced from protein that had been equilibrated in 1.9 M urea, and from protein that had been equilibrated in 3 or 4 M urea. Figure 6a and b shows that the approach to equilibrium in 5.6 M urea was significantly faster when the protein was initially in 3 or 4 M urea than when it was initially in 1.9 M urea. Surprisingly, the observed rate constants of the slow unfolding phase for the 3 to 5.6 M urea and from 4 to 5.6 M urea jumps did not fall on the λ_3 chevron (Fig. 5c) as did the rate constant observed for the 1.9 to 5.6 M urea jump but fell instead on the λ_2 chevron defined by the rate constants observed for refolding.

Previously, it had been shown that in 4 M urea, dm CTD exists as 33% N, 44% U, and 23% I_U, where I_U is an aggregation-prone intermediate [30]. The near absence of a lag phase for unfolding from 3 or 4 M urea to higher concentrations of urea (Fig. 6a) may be attributed to the faster unfolding of I_R, which has higher fluorescence compared to the unfolding of N. In contrast, the protein exists as N in 1.9 M urea, and the lag phase was clearly apparent.

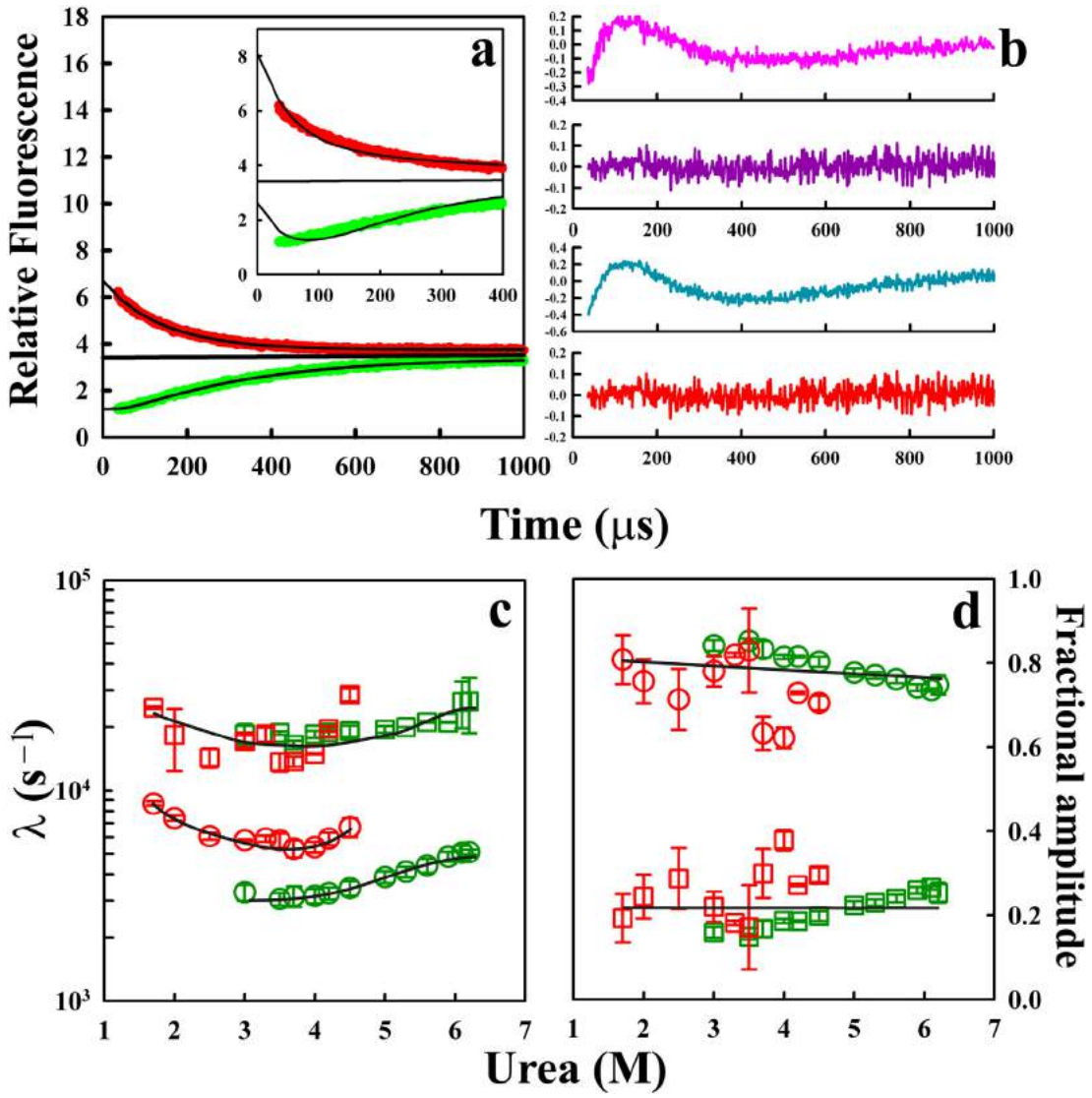


Fig. 5. Evidence for different pathways of refolding and unfolding of dm CTD at pH 4. (a) Representative kinetic trace of refolding: 6 → 4 M urea, (●), and unfolding: 1.9 → 4 M urea, (●). — shows free fits of the kinetic traces to double exponential equations; – shows fluorescence signal at equilibrium in 4 M urea. Inset: Solid black line through the refolding kinetic trace ●) is the fit to a double exponential equation with the rate constants constrained to those from a free fit of the unfolding trace obtained in 4 M urea. — through the unfolding kinetic trace (●) shows the fit to a double exponential equation by constraining the rate constants obtained from a free fit of the refolding kinetic trace at 4 M urea. (b) Residuals of the free fit (purple) and constrained fit (pink) to the kinetic trace of refolding and that of the free fit (red) and constrained fit (cyan) to the kinetic trace of unfolding in 4 M urea. (c) Denaturant dependences of the observed rate constants of the fast (squares) and slow (circles) phases of refolding (red) and unfolding (green) obtained from double exponential fits to refolding kinetic traces starting from 6 M urea, and unfolding kinetic traces starting from 1.9 M urea, to various concentrations of urea. The lines through the data are drawn to guide the eye. (d) Dependence of the fractional amplitudes of the fast (squares) and slow (circles) kinetic phases of unfolding from 1.9 M urea to higher urea concentrations (green) and refolding from 6 M urea, to lower urea concentrations (red). In all cases, the amplitudes are obtained from double exponential fits of the kinetic traces of unfolding and refolding.

When refolding from 6 M urea was carried out at lower concentrations of urea, the fast and slow relaxation rate constants defined the chevrons λ_1 and λ_2 (Fig. 5c), respectively. Refolding from 4 M to lower urea concentrations occurred, however, in a single kinetic phase (Fig. 6b) at rate constants defining the λ_2 chevron. Surprisingly, when refolding was commenced from 3 M urea to identical lower concentrations of urea as refolding from 6 or 4 M urea, the entire fluorescence change occurred again in a single phase but at a rate constant falling on the λ_3 chevron and as predicted from global analysis of kinetic unfolding data for those urea concentrations [Figs. 6b, c and S4c (SI)]. Hence, refolding was

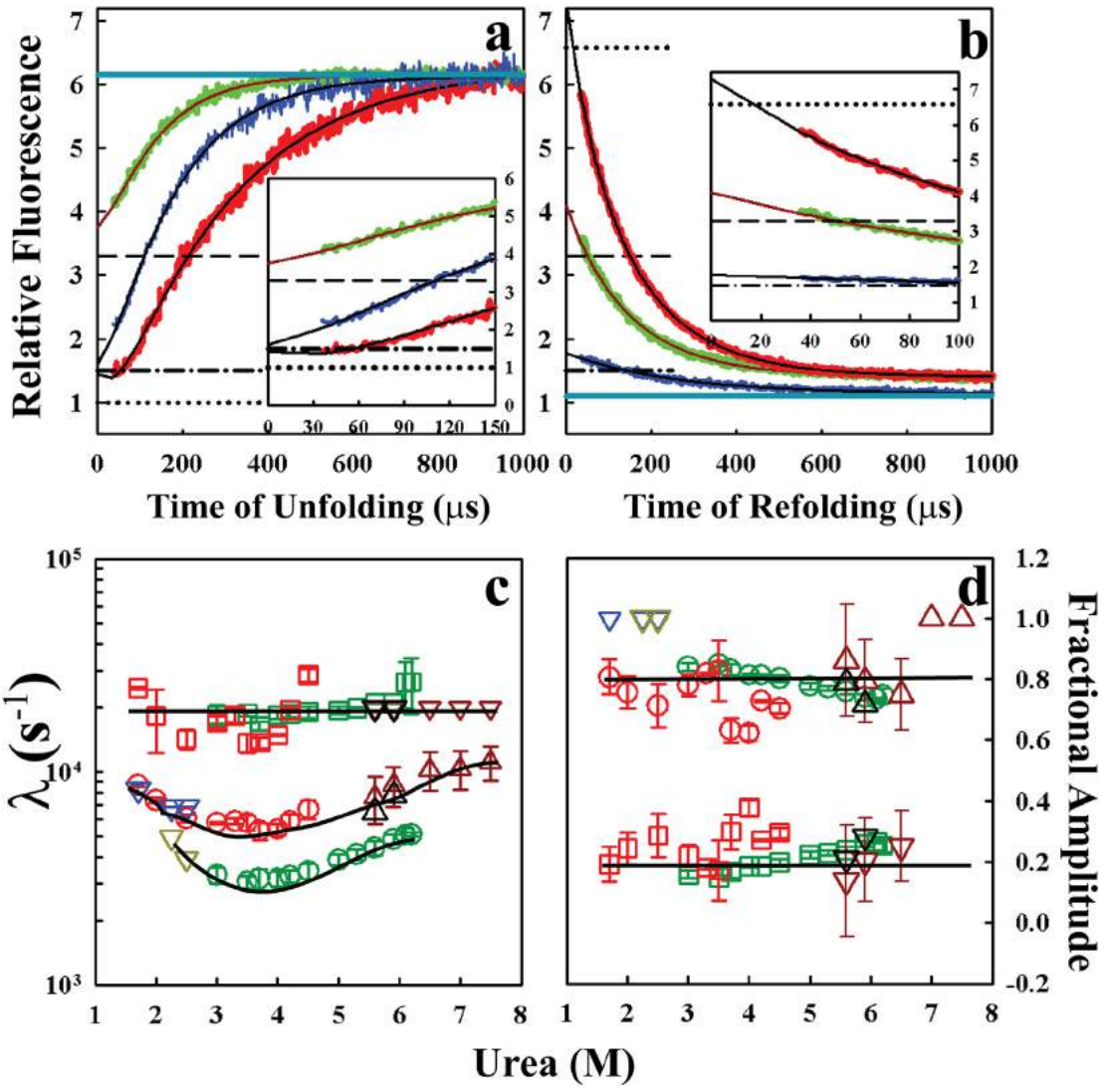


Fig. 6. Comparison of folding kinetics of protein in 1.9, 3, and 4 M urea, at pH 4. (a) Kinetic traces of unfolding of native dm CTD: from 1.9 → 5.6 M (—); from 3 → 5.6 M (—) and 4 → 5.6 M (—) urea, respectively; non-linear least squares fit of 1.9 → 5.6 M kinetic trace to a double exponential equation shown (—). Non-linear least squares fit of the kinetic traces: 3 → 5.6 M (—) and 4 → 5.6 M (—) urea to a single exponential equation shown. Inset shows the initial parts of the reactions; fluorescence signals in 4 M (– –), 3 M (– ·), and 1.9 M (·····) urea, shown relative to fluorescence signal in 1.9 M urea. (b) Kinetic trace of refolding from 6 → 2.5 M urea (—); non-linear least squares fit of the data to a double exponential equation (—); refolding kinetic traces from 4 → 2.5 M (—) and 3 → 2.5 M (—) urea, respectively. (—) and (—) are non-linear least-squares fit of kinetic traces to a single exponential equation; fluorescence signals in 4 M (– –), 3 M (– ·), and 6 M (·····) urea, relative to fluorescence signal in 1.9 M urea shown. Inset shows the initial parts of the reactions. — in panels a and b is the fluorescence signal at 5.6 and 2.5 M urea, respectively, obtained from equilibrium unfolding transition of dm CTD at pH 4. (c) Denaturant dependences of the observed rate constants of the fast (squares) and slow (circles) phases of refolding (red) and unfolding (green) obtained from double exponential fits to refolding kinetic traces starting from 6 M urea and unfolding starting from 1.9 M urea to various concentrations of urea; observed rate constants of the fast (inverted) and slow (upright) unfolding phases from 4 M (dark red) and from 3 M (black) to higher concentrations of urea obtained from double exponential fits to the data. Kinetic traces of unfolding from 4 → 7 M and 4 → 7.5 M were fit to single exponential equations as the lag phase was not temporally resolvable. Observed rate constants of refolding from 3 M (▽) and from 4 M (▽) to lower concentrations of urea obtained from single exponential fits to the kinetic traces of refolding. (d) Dependence of the fractional amplitudes of the fast (squares) and slow (circles) kinetic phases of unfolding from 1.9 M urea to higher urea concentrations (green) and refolding from 6 M urea, to lower urea concentrations (red) obtained from double exponenial fits of the kinetic traces of unfolding and refolding. The dark red and black triangles are the dependence of fractional amplitudes of the fast (inverted) and slow (upright) phases of unfolding from 4 and 3 M urea, respectively, to higher concentrations of urea; fractional amplitudes of refolding from 4 (▽) and 3 M (▽) urea to lower concentrations of urea. The lines through the data in panels c and d are to guide the eye.

slower when commenced from 3 M than from 6 or 4 M urea.

Discussion

The free energy landscape is relatively smooth at pH 7

The equilibrium unfolding study at pH 7, which suggested two-state unfolding of dm CTD [30], was unable to capture the complexities observed in the kinetic folding studies reported here. The kinetic studies indicate that unfolding/folding at pH 7 occurs predominantly *via* a linear, sequential folding pathway, on which at least two intermediates are populated. Folding becomes more complex at intermediate urea concentrations (4–5 M), where the relative stabilities of N and U differ by only ~2.5 kcal mol^{-1}. The final fluorescence of the protein at 1 ms of refolding was higher than that expected from the equilibrium unfolding transition and reached the equilibrium fluorescence signal on a slower time scale (Fig. S1, SI). This suggests that an additional slow kinetic phase of refolding occurs from a N* state of higher fluorescence to the N state, which is not observed during unfolding to identical urea concentrations. The slower time scale of conversion from N* to N suggests that a large energy barrier is present. The data suggest that in 4 to 5 M urea, folding dm CTD molecules, after crossing the major folding energy barriers, get trapped in the N* state while searching for the N state.

Distinct unfolding and refolding pathways are utilized at pH 4

When protein folding and unfolding reactions occur *via* the same pathway in opposite directions, the observed kinetics of relaxation following a jump or drop in denaturant concentration will depend only on the final denaturant concentration and be independent of the initial conditions [31–35]. On the other hand, when the folding or unfolding kinetics are found to depend on the initial conditions, the utilization of multiple parallel pathways, accompanied by shifts in the positions of the native, intermediate or unfolded conformations along a multidimensional free energy surface, is indicated [36–39]. In this study, the observation that when dm CTD was unfolded and refolded at the same final urea concentration, starting from different initial conditions, the rate constant of the major (80%) phase of unfolding was half that of the major (80%) phase of refolding, indicates that folding does not follow the same kinetic pathway as unfolding. It was not possible to fit both the folding and unfolding data to any model involving reversible folding pathways, which are linear or parallel. When the unfolding data were globally fit to Scheme 2 to obtain the microscopic rate constants of folding and unfolding for each of the individual steps, it was found that the predicted folding rate constants did not predict the experimentally determined λ values (Fig. 5c). Moreover, during folding, an initial burst of hyperfluorescence was observed, which represents the formation of the denatured ensemble D′. Global analysis of the kinetic data on unfolding does not, however, predict the occurrence of an initial burst of hyperfluorescence during refolding. Very importantly, the results suggest that folding protein molecules traverse a different region of the free energy landscape than do unfolding protein molecules. The thermodynamic states accessible from the unfolding pathway are not accessible to folding molecules, and *vice versa*; if they were, then the rate constants observed for folding and unfolding would have depended only on the final urea concentration and not on the initial urea concentration from which the folding and unfolding reactions were commenced.

A switch in the unfolding pathway is observed upon change in the initial conditions

In a previous equilibrium unfolding study [30], it had been shown that the unfolding of dm CTD was describable by a N ↔ I ↔ U mechanism, where I is an intermediate with N-like fluorescence. In 3 M urea, the three-state analysis suggested that the population of protein molecules consists of 67% N, 18% I, and 15% U, while in 4 M urea, it consists of 33% N, 23% I, and 44% U. Surprisingly, however, the total fluorescence change that occurred upon unfolding protein pre-equilibrated in either 3 or 4 M urea was found to occur in only one kinetic phase. This suggests that the earlier minimalistic three-state analysis of the equilibrium unfolding data was an over-simplification and that unfolding is more complex: more intermediates must be present. What, however, is more remarkable is that the kinetics of unfolding starting from protein pre-equilibrated in either 3 or 4 M urea is very different from that starting from protein pre-equilibrated in 1.9 M urea (in which the protein is native). In fact, the observed rate constants of unfolding from 3 or 4 M urea fall on the λ_2 chevron defined by the observed rate constants of refolding and not on the λ_3 chevron defined by the observed rate constant of unfolding starting from 1.9 M urea. This result indicates that in 3 or 4 M urea, but not in 1.9 M urea, protein molecules can access the refolding pathway on the free energy landscape. In addition, the data in Fig. 6 indicate that while the protein in 4 M urea can refold *via* the refolding pathway (with an apparent rate constant defined by the λ_2 chevron), protein in 3 M urea refolds by the unfolding pathway used by native protein molecules (starting from 1.9 M urea) to unfold. It appears therefore that at intermediate

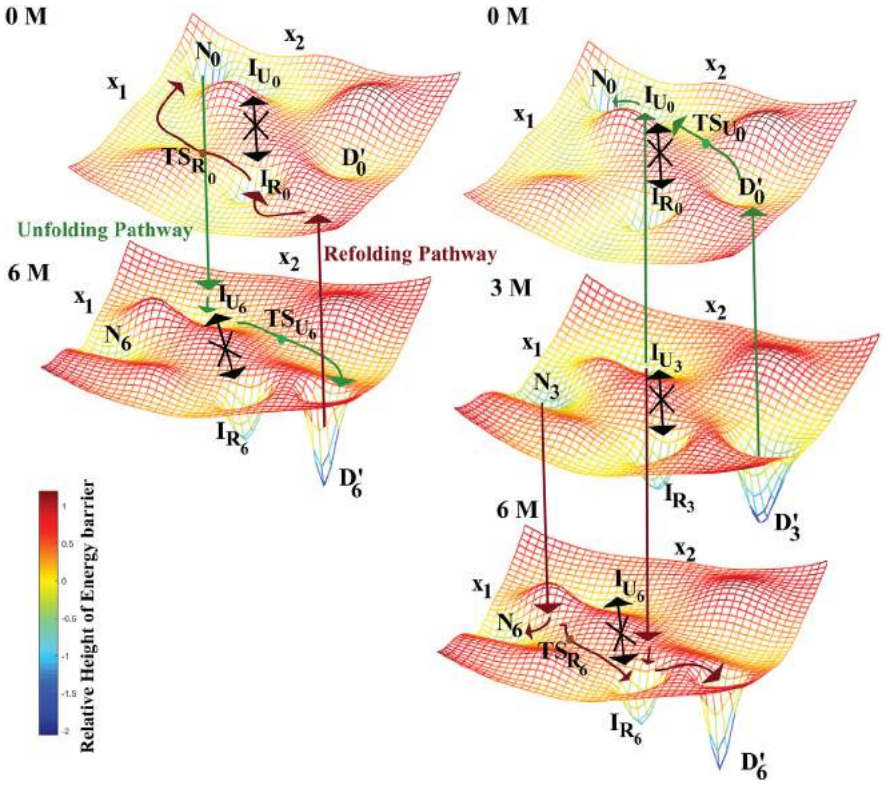


Fig. 7. Schematic representation of the free energy surfaces at varying urea concentrations. x_1 and x_2 represent generalized unfolding reaction coordinates. The symbols N, I_U, and TS_U represent the native state, unfolding intermediate and transition state on the unfolding pathway, respectively. The symbols I_R and TS_R are the refolding intermediate, and transition state of the refolding pathway, respectively. D′ refers to the denatured state ensemble. The subscripts denote the urea concentration. The arrows represent jumps from an initial to a final urea concentration. The arrows on the final surface indicate possible pathways. The refolding pathway is indicated by red arrows and unfolding pathway by green arrows. The black arrows with a cross indicate that the thermodynamic states are not accessible from each other. The roughness of the landscape and a large energy barrier between TS_U and TS_R, which make these two pathways inaccessible, have been indicated by an arrow with a cross for simplification.

urea concentrations, protein molecules are partitioned into two different regions of the free energy landscape, which are separated by large free energy barriers. Protein molecules in one region cannot access the other region in the sub-millisecond time domain over which folding occurs. In future studies, it would be interesting to examine the effect of varying the time of equilibration in 3 M or 4 M urea, on the folding and unfolding kinetics of the protein. Surprisingly, the rate constant corresponding to the fastest phase of refolding (~20,000 s^{-1}) is not observed in the fit to the kinetic refolding traces of dm CTD, obtained upon jumps from either 3 or 4 M urea, to lower urea concentrations.

Experimental evidence of folding of dm CTD on a multi-dimensional energy landscape

The observation that the major pathway utilized for (un)folding depends on the initial conditions can be explained only if un(folding) is considered to commence from different points on a multi-dimensional free energy landscape, with the positions of the N, I, and denatured states being dependent on the urea concentration. Figure 7 illustrates how the complex free energy surfaces of dm CTD at pH 4 might vary at different concentrations of urea. Two scenarios of un (folding) have been illustrated: (a) When the urea concentration is jumped from 0-1.9 to 6 M urea, unfolding proceeds *via* the intermediate I_U and transition state TS_U to the denatured state ensemble. When the urea concentration is jumped from 6 M to lower urea concentrations, refolding proceeds *via* the refolding intermediate I_R and transition state TS_R. (b) Starting from 3 M urea, unfolding at higher urea concentrations proceeds *via* I_R and TS_R, while refolding at lower urea concentrations proceeds *via* I_U and TS_U.

When equilibrated in 3 M urea, 67% of dm CTD molecules are present as N [30], which should unfold at rates identical to those seen for the unfolding of N starting from 0 to 1.9 M urea. However, 80% of the molecules were observed to unfold with rate

constants falling on the refolding chevron arm (λ_2), suggesting that although N has identical fluorescence properties in 0–1.9 M urea and in 3 M urea, the N state in 3 M urea must have shifted on the free energy surface along another reaction coordinate where the lowest energy pathway for unfolding is the refolding pathway. Similarly, the 15% U molecules present in 3 M urea should refold with rate constants identical to those observed for the refolding of U starting from 6 M urea. But this was not observed (Fig. 6). All U molecules, starting from 3 M urea, refolded with rate constants falling on the unfolding chevron arm (λ_3). These observations suggest that both N and U, and likely the I states as well, shift along a reaction coordinate different from the fluorescence coordinate. Such movements of N and U on a multi-dimensional free energy landscape results in to the folding pathway utilized being dependent on initial conditions, as has been observed earlier for small proteins, which are two-state or downhill folders [36,37,39].

Low pH-induced energetic frustration results in folding *via* multiple pathways

For dm CTD, it seems that the free energy landscape at pH 7 is sufficiently smooth that the folding and unfolding routes utilized are separated by very small energy barriers. Consequently, the folding and unfolding routes all merge into a single, sequential, linear pathway. It seems that a change in pH to 4 induces barriers on the free energy landscape, which are large enough to prevent exchange between the populations of folding and unfolding molecules on the millisecond time scale. Consequently, unfolding and folding protein molecules would utilize different, distinct pathways separated by the energy barriers, and traverse different regions of the free energy landscape. The barriers and free energy landscape themselves appear to change with a change in denaturant concentration in such a way that a previously inaccessible region defining a folding pathway becomes accessible to unfolding protein molecules, or a previously inaccessible region defining an unfolding pathway becomes accessible to folding protein molecules.

At pH 4, energy barriers large enough to separate the populations of folding and unfolding molecules could arise from structural and energetic constraints in the conformations sampled, which would be the result of competing native and non-native interactions, cavity formation from imperfect fits of amino acid side chains, or structural rearrangements between locally favorable secondary and tertiary structures [8,10,40,41]. In the case of moPrP, it might even be possible that its single disulfide bond is present in alternative conformations at pH 4 but not at pH 7. The transitions between the alternative conformations could be slow (in the millisecond time domain) [42,43] compared to the folding transitions, and this may lead to the folding and unfolding kinetics being dependent on the initial conditions. This explanation for pH-induced energy barriers keeping apart folding and unfolding conformations on the free energy landscape at pH 4 appears, however, unlikely because an NMR study of prion proteins in the pH range of 3.5 to 7 did not observe any line broadening of the resonances arising from the disulfide-bonding residues [44], and because another NMR study did not detect any structural perturbation of the disulfide-bonded residues in the native state of moPrP at intermediate urea concentrations [45].

The most plausible reason for the heights of barriers separating the populations of molecules on the folding and unfolding pathways on the free energy landscape of moPrP, increasing dramatically upon a reduction in pH from 7 to 4, is that electrostatic interactions become significantly perturbed. NMR studies have shown that structure at the C-terminus of helix 2 of moPrP, is drastically affected when the pH is lowered to below pH 4.8, because of the protonation of H186. Molecular dynamics simulations have indicated that the protonation of H186 in turn disrupts the network of electrostatic interactions in the protein [46–48]. Theoretical studies have also indicated that electrostatic networks, including salt bridges, play a critical role in maintaining the integrity of PrP structure [49,50], and it is well known that the perturbation of a salt bridge can drastically alter the height of a free energy barrier in protein folding [51]. In the case of moPrP itself, it is the perturbation of the K193–E195 salt bridge that initiates misfolding and oligomerization [52]. Moreover, mutational studies on other proteins have brought out how sub-optimal electrostatic interactions can lead to significant frustration on the free energy landscape [53].

Ruggedness on protein free energy landscapes make the energy basins of thermodynamic states rough [19,41,54–56]. This slows down diffusion of the folding protein chain on the free energy landscape and increases the barrier crossing time (the pre-exponential factor in Kramer's theory). On a one-dimensional free energy landscape, the diffusion constant has been shown to increase exponentially with the mean of squared ruggedness [54]. On the multi-dimensional free energy landscape of protein folding, ruggedness could be expected to have an even larger effect on the diffusion constant and hence on the barrier crossing time. In this context, it should be noted that dm CTD folds and unfolds in the sub-millisecond time domain, indicating that the barrier heights on the folding and unfolding pathways, which slow down the reactions, are about 5–6 $k_B T$. It might therefore be expected

that the ruggedness observed at pH 4 would make the folding of dm CTD slower at pH 4 than at pH 7, but that is not observed (Figs. 3–6). It seems that the large free energy barrier dominates over the ruggedness of the free energy landscape.

Ruggedness of energy landscape and its role in protein aggregation

A free energy landscape defined by many small and large barriers separating different states is expected to be populated by intermediate states to different extents, depending on their Boltzmann distribution. Since the free energy landscape appears to reflect a trade-off between folding and function [11,57], the presence of multiple minima could be crucial for function [58–60]. Multiple states of similar energies but marginally different structures may be selectively populated in the presence of specific ligands, or by changing the pH or solvent conditions, which are crucial for its function. Although protein sequences have evolved to decrease their propensity to aggregate [61], protein motions may lead to the population of functionally relevant metastable states which are nevertheless aggregation-prone, such as those observed in the case of apomyoglobin [5], iLBPs [62], serpins [63], and ataxins [64].

In the case of moPrP, previous studies have indicated that the N state is structurally dynamic and can access high energy folding intermediate states referred to as partially unfolded forms (PUFs) [65]. One of these PUFs (PUF2) was seen to initiate misfolding and oligomerization at low pH [30]. In the current study, it was observed that PUF2 is similar in energy to I_U (I_1 in Scheme 2), the unfolding intermediate, which is populated on the unfolding pathway of native protein. The ruggedness on the free-energy landscape of dm CTD at pH 4 is such that it results in the protein unfolding *via* the obligatory aggregation-prone intermediate. This is most likely the reason behind the high oligomerization propensity of the protein at low pH. At physiological pH, the protein aggregates only upon destabilization of the native state by denaturant, or upon exposure to harsh conditions.

A very early stopped-flow kinetic study, with millisecond time resolution, of the folding of the CTD of the F174W variant of moPrP, at pH 7 and at 4 °C, had failed to detect any folding intermediates [26]. In contrast, the current study of the folding of dm CTD, using continuous flow mixing with microsecond time resolution, has identified two folding intermediates. Folding intermediates have also been identified in folding studies of the human prion protein [20–23] and the sheep prion protein [24]. The effect of pathogenic mutations on the extent to which these intermediates were populated suggested the involvement of the intermediates in the aggregation of the proteins [21,24].

Earlier studies of the folding and unfolding of the Y217W mutant variant of moPrP at low pH (pH 2 and 4.8) had revealed four-state reversible unfolding and the accumulation of a molten globule-like intermediate [27,28]. Although the same studies reported single exponential folding kinetics, the curvature seen in the folding arm of the chevron was indicative of at least one folding intermediate. The current study has revealed an even greater complexity in the folding and unfolding of the dm CTD variant of moPrP at low pH (pH 4), with the folding and unfolding kinetics being observed to be dependent on the initial conditions. This dependence on initial conditions, which was not investigated in the earlier studies of the folding/unfolding of the human PrP and Syrian hamster PrP [20–24,27,28], makes it difficult to compare the intermediates observed in the earlier studies to those observed in the current study. In the earlier studies, aggregation has been implicated to initiate from early refolding intermediates [20,21,23] or from the unfolded state [29,66]. The aggregation propensities of the intermediates identified in the current study will be studied in the future.

Conclusion

The free energy surface of dm CTD has been shown to be altered drastically by a change in pH. At pH 7, where the protein misfolds only upon drastic destabilization, folding and unfolding occur *via* one principal pathway populated by two intermediates. At pH 4, where the protein is prone to aggregation, folding and unfolding occur *via* independent pathways, both of which are populated by intermediate states accessible either for folding or for unfolding. The two pathways on the free energy landscape appear to be kept separate by large and small barriers that attest to a ruggedness of the landscape at pH 4, which is not seen at pH 7. The utilization of the pathways and hence the rate constants observed for folding and unfolding depend surprisingly on not only the final conditions but also on the starting conditions. This observation provides strong evidence for the modulation of the free energy landscape by the physico-chemical conditions in which the protein is present. Misfolding and aggregation appear to initiate from the intermediate populated on the unfolding pathway of the protein starting from native-like conditions.

Materials and Methods

Protein expression and purification

dm CTD [W144F/F174W moPrP (121–231)] was purified as described earlier [30].

Chemicals and buffers

The chemicals used in this study were obtained from Sigma, unless mentioned otherwise. For experiments that were carried out at pH 4, the native buffer (NB) was 20 mM sodium acetate, and the unfolding buffer (UB) was NB containing variable concentrations of urea (obtained from US Biochemicals) in the range 0–10 M. The pH in both buffers was adjusted to 4. For experiments that were carried out at physiological pH, the NB was 50 mM sodium phosphate and the UB was NB containing variable concentrations of urea in the range 0–10 M. The pH in both buffers was adjusted to 7. All solutions were filtered through 0.22 μm filters from Whatman.

Microsecond mixing kinetic studies of the folding and unfolding of dm CTD

All folding kinetic studies were carried out at 25 °C. The intrinsic fluorescence of W174 of dm CTD was used to monitor the urea-induced unfolding/refolding of dm CTD using a custom-made continuous flow set up described elsewhere [67]. The excitation wavelength used was 295 nm and emission was monitored at 360 nm. The mixing dead time was determined to be 37 ± 5 μs by monitoring the kinetics of the quenching of the fluorescence of NATA (*N*-acetyl-L-tryptophanamide) by different concentrations of *N*-bromo succinimide (NBS), under pseudo first-order reaction conditions.

Non-uniform illumination along the length of the channel was accounted for by expressing the signal from the continuous flow as relative fluorescence [68] which was calculated according to the following equation:

$$S_{\mathrm{Rel}} = \frac{S_{\mathrm{R}} - S_{\mathrm{B}}}{S_{\mathrm{N}} - S_{\mathrm{NB}}} \qquad (1)$$

where S_{Rel} is the relative fluorescence signal, S_{R} is the signal from the reaction, and S_{B} is the signal from the corresponding buffer. S_{N} refers to the signal of the sample used for normalization which is constant with time, and S_{NB} is the signal of the corresponding buffer. The unfolding traces were normalized to the fluorescence of native dm CTD at identical protein concentration, in 1.9 M urea at pH 4, or in 3 M urea at pH 7. The refolding kinetic traces were normalized to the signal of unfolded protein of identical concentration, in 6 M urea at pH 4, or in 7.9 M urea at pH 7.

In unfolding studies carried out at pH 4, native dm CTD in 1.9 M urea was mixed in a 3:4 ratio with UB to obtain final urea concentrations in the range 3 to 6.2 M. For refolding studies carried out at pH 4, dm CTD was unfolded in 6 M urea and equilibrated for 15 to 30 min. The refolding reaction was initiated by mixing the unfolded protein with NB containing variable concentrations of urea either in the ratio 2:5 or in the ratio 3:4 such that the final urea concentration was in the range of 1.7 to 4.5 M.

In unfolding studies carried out at pH 7, native dm CTD (in 3 M urea) was mixed in a 3:4 ratio with UB to obtain final urea concentrations in the range of 5 to 6.8 M. For refolding studies at pH 7, dm CTD was unfolded in 7.9 M urea and equilibrated for 30 min. Refolding was initiated by mixing the unfolded protein with NB in the ratio 2:5 or 3:4 such that the final concentration of urea was in the range of 2.3 to 5 M.

The relative fluorescence values defining each kinetic trace at both pH 4 and at pH 7 was calculated using Eq. (1). The protein concentrations used for the kinetic folding experiments were in the range of 6 to 24 μM, and the combined flow rate used in mixing was 7 mL/min.

Folding kinetics at pH 4 from different initial conditions

dm CTD, pre-equilibrated in either 3 or 4 M urea, was either unfolded by mixing with UB, to obtain final urea concentrations in the range of 5.6 to 7.5 M, or refolded by mixing with NB, to obtain final urea concentrations in the range of 1.7 to 2.5 M.

dm CTD pre-equilibrated in 6.5 to 7.5 M urea was refolded by mixing with NB to a final urea concentration of 4 M.

Data analysis

The normalized kinetic traces of the unfolding and refolding of dm CTD at pH 4 and pH 7 were fitted to a single or a double exponential equation using Sigma plot. In every case, the residuals were carefully compared to determine the best fit to the data. From the double exponential fits to the kinetic traces, the rate constant and amplitude of each kinetic phase of refolding or unfolding at any particular denaturant concentration were obtained.

Global analysis of kinetic data at pH 7

For global analysis of the kinetic refolding and unfolding data, all kinetic traces were normalized to the signal of the native protein at 3 M urea. The normalized kinetic traces were fit to a kinetic model, after making the following assumptions: all folding and unfolding reactions were considered to be unimolecular. Indeed, the rate constants obtained from the kinetic studies were found to be independent of the concentration of the protein used. The dependence of the log of each microscopic rate constant of conversion between two species was considered to have a linear dependence on urea concentration, as

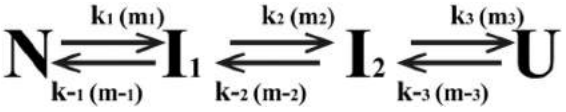
shown in Eq. (2).

$$\ln k_i = \ln k_i^0 + \left(\frac{m_i^{\ddagger}}{RT}\right) \times [\text{urea}] \qquad (2)$$

Here, k_i^0 and k_i are the microscopic rate constants in the absence and presence of the denaturant urea, respectively. i is an elementary reaction step between two conformations. $m_i^{\#}$ is a measure of the difference in accessible surface area between the starting conformation and the transition state (TS) defining the step, R is the universal gas constant, and T is the temperature. The protein was considered to be completely in the unfolded state (U) in strongly denaturing solvent conditions and completely in the native (N) state in native solvent conditions. The global fitting of the kinetic data was done using a matrix-based program [32,69] written in MATLAB® [70].

Solution for the four-state sequential scheme

Global fits of the kinetic data to a three-state linear or parallel pathway model were unable to account for the experimental kinetic data. However, a simple linear model with four states (Scheme 1) could fit the normalized refolding and unfolding kinetic traces of dm CTD at pH 7. The four-state model when further simplified and reduced to a three-state linear pathway model was also able to explain the kinetic data. This is described as below:



Scheme 1.

Here, k_1, k_2, and k_3 are the microscopic rate constants of conversion of N to I_1, I_1 to I_2, and I_2 to U, respectively, and the microscopic rate constants for the backward reactions I_1 to N, I_2 to I_1, and U to I_2 are k_{-1}, k_{-2}, and k_{-3}, respectively. The kinetic m value for each step is m_i, where i is any elementary reaction step between two conformations as indicated in the scheme. The net equilibrium constant (K) of the overall unfolding reaction shown in Scheme 1 is equal to the product of the ratios of the forward to backward rate constants of each step involved, as shown in Eq. (4).

The assumption that under all denaturing conditions, the I_2 to U inter-conversion is much faster than the inter-conversion rates of the other elementary reaction steps would mean that a pre-equilibrium is established between I_2 and U. Hence, the four-state scheme may be reduced to the following three state scheme in which I_2' represents the pre-equilibrium mixture of I_2 and U.

Here,

$$k'_{-2} = \frac{k_{-2}}{(1+K_3)} \qquad (3)$$

and

$$K_3 = \frac{k_3}{k_{-3}} = \frac{[\text{U}]}{[\text{I}_2]} \qquad (4)$$

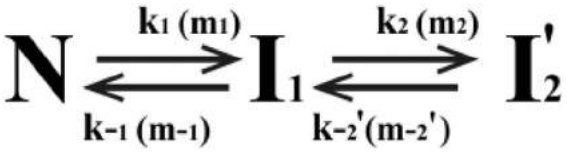
where $k_{-3} = \frac{k_1.k_2.k_3}{k_{-1}.k_{-2}.K}$

K is the net equilibrium constant of unfolding from N to U. Scheme 2 may be described by the following system of first order linear differential equations:

$$\frac{dN}{dt} = -k_1.[\text{N}] + k_{-1}.[\text{I}_1] \qquad (5a)$$

$$\frac{dI_1}{dt} = k_1.[\text{N}] - (k_{-1}+k_2).[\text{I}_1] + k'_{-2}.[\text{I}'_2] \qquad (5b)$$

$$\frac{dI'_2}{dt} = k_2.[\text{I}_1] - k'_{-2}.[\text{I}'_2] \qquad (5c)$$



Scheme 2.

The elementary rate constants at any given denaturant concentration calculated using Eq. (2) were used to define a rate matrix based on the above differential equations [Eqs. (5a)–(5c)] describing a reduced three-state sequential scheme. The eigenvalues of the rate matrices at each denaturant concentration were calculated using the MATLAB function eig. To generate fluorescence monitored kinetic refolding and unfolding traces, MATLAB function odes23s was used to solve the differential equations for Scheme 2 [Eqs. (5a)–(5c)] for a particular set of initial conditions, and the fractional concentrations of N, I_1, and I_2' were obtained at different times, t, at any particular denaturant condition. The fluorescence monitored kinetic unfolding and refolding traces were then generated using Eqs. (6)–(8).

$$F_T(t) = F_{\text{N}}.[\text{N}](t) + F_{\text{I}_1}.[\text{I}_1](t) + F_{\text{I}'_2}.[\text{I}'_2](t) \qquad (6)$$

where,

$$F_{\text{I}'_2} = \frac{(F_{\text{I}_2} + K_3.F_{\text{U}})}{(1+K_3)} \qquad (7)$$

Here, F_T is the normalized fluorescence signal at time t; F_N, F_{I_1}, F_{I_2}, and F_U are the relative fluorescence signals of the N, I_1, I_2, and U states relative to that of the N state in 3 M urea; and the [N (t)], [I_1(t)],and [I_2'(t)] are the concentrations of N, I_1, and I_2' at any time t. The fluorescence of I_2' was determined using Eq. (7). The fluorescence values of the I_1 and I_2 states were assumed to be linearly dependent on urea as shown in Eq. (8), where F_i^0 is the fluorescence of state i in the absence of urea, and s_i is its dependence on urea concentration. The fluorescence values of the native (F_N) and unfolded (F_U) states were obtained from linear extrapolations of the native and unfolded protein baselines to high and low urea concentrations, respectively, as shown by the dashed and dotted lines in Fig. 3b.

$$F_i = F_i^0 + s_i.[\text{urea}] \quad (8)$$

For fitting, the initial set of values of the fitting parameters (k_i, m_i, $F^0_{N/I_1/I2/U,}$ and $s_{N/I_1/I2/U}$) were simultaneously varied using the MATLAB function "fminsearchbnd." At any urea concentration, the fluorescence value obtained by extrapolating the double exponential fit to zero time was found to be higher than the fluorescence value obtained by linearly extrapolating the unfolded protein baseline of the equilibrium unfolding transition, to the same urea concentration to achieve the lowest least-squares difference between the experimental and simulated kinetic data by MATLAB.

The parameters thus obtained from the best fit of the global analysis of folding kinetics at pH 7 obtained the dependence of the forward and backward microscopic rate constants of individual steps upon unfolding from N to U on urea, the resultant observed kinetic rate constant and the relative fluorescence change upon unfolding or refolding.

Global analysis of unfolding kinetic data at pH 4

Since the rate constants observed for unfolding and refolding at identical final urea concentrations were two-fold different, a combined global analysis of both unfolding and refolding data was not possible. Kinetic unfolding traces of dm CTD at pH 4 normalized to the fluorescence signal of native dm CTD at 1.9 M urea were globally fit to Scheme 2. In the subsequent text and figures for pH 4, I_1 in Scheme 2 is referred to as I_U and I_2' represents the denatured state ensemble, D′.

Acknowledgment

We thank Dr. Sandeep Krishna, Dr. Madan Rao, and Dr. Shachi Gosavi for their intellectual inputs. We thank Dr. Shreyas Gokhale, Dr. Nilesh Aghera, and members of our laboratory for discussions. J.B.U. is a recipient of a JC Bose National Fellowship from the Government of India. This work was funded by the Tata Institute of Fundamental Research and by the Department of Biotechnology, Government of India.

Author Contributions: J.U. acquired funding. R. M. and J.U. conceptualized the study. R.M. and R.G. carried out the investigation and data curation. R.M. provided formal analysis with supervision from J.U. R.M. and J.U. wrote the original draft with editing from R.G. All authors approved the final version of the manuscript.

Notes

The authors declare no competing financial interests.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jmb.2018.12.009.

Received 25 July 2018;
Received in revised form 15 November 2018;
Accepted 16 December 2018
Available online 4 January 2019

Keywords:
prion;
cooperativity;
energy barrier;
ruggedness;
aggregation

Abbreviations used:
moPrP, mouse prion protein; dm CTD, double-mutant variant of the C-terminal domain; PUF, partially unfolded form.

References

[1] P.E. Leopold, M. Montal, J.N. Onuchic, Protein folding funnels: a kinetic approach to the sequence–structure relationship, Proc. Natl. Acad. Sci. 89 (1992) 8721–8725.
[2] J.O. Onuchic, Z. Luthey-Schulten, P.G. Wolynes, Theory of protein folding: the energy landscape perspective, Annu. Rev. Phys. Chem. 48 (1997) 545–600.
[3] A.J. Baldwin, T.P.J. Knowles, G.G. Tartaglia, A.W. Fitzpatrick, G.L. Devlin, S.L. Shammas, C.A. Waudby, M.F. Mossuto, S. Meehan, S.L. Gras, J. Christodoulou, S.J. Anthony-Cahill, P.D. Barker, M. Vendruscolo, C.M. Dobson, Metastability of native proteins and the phenomenon of amyloid formation, J. Am. Chem. Soc. 133 (2011) 14160–14163.
[4] C.M. Dobson, Protein folding and misfolding, Nature 426 (2003) 884–890.

[5] A. De Simone, A. Dhulesia, G. Soldi, M. Vendruscolo, S.-T.D. Hsu, F. Chiti, C.M. Dobson, Experimental free energy surfaces reveal the mechanisms of maintenance of protein solubility, Proc. Natl. Acad. Sci. 108 (2011) 21057–21062.
[6] H. Kaya, H.S. Chan, Energetic components of cooperative protein folding, Phys. Rev. Lett. 85 (2000) 4823–4826.
[7] H. Kaya, H.S. Chan, Contact order dependent protein folding rates: kinetic consequences of a cooperative interplay between favorable nonlocal interactions and local conformational preferences, Proteins Struct. Funct. Genet. 52 (2003) 524–533.
[8] J.N. Onuchic, P.G. Wolynes, Theory of protein folding, Curr. Opin. Struct. Biol. 14 (2004) 70–75.
[9] H.S. Chan, Z. Zhang, S. Wallin, Z. Liu, Cooperativity, local–nonlocal coupling, and nonnative interactions: principles of protein folding from coarse-grained models, Annu. Rev. Phys. Chem. 62 (2011) 301–326.
[10] L.L. Chavez, 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.
[11] M. Gruebele, Protein folding: the free energy surface, Curr. Opin. Struct. Biol. 12 (2002) 161–168.
[12] V. Muñoz, Thermodynamics and kinetics of downhill protein folding investigated with a simple statistical mechanical model, Int. J. Quantum Chem. 90 (2002) 1522–1528.
[13] M. Gruebele, Downhill protein folding: evolution meets physics, C. R. Biol. 328 (2005) 701–712.
[14] P. Malhotra, J.B. Udgaonkar, Secondary structural change can occur diffusely and not modularly during protein folding and unfolding reactions, J. Am. Chem. Soc. 138 (2016) 5866–5878.
[15] P. Malhotra, J.B. Udgaonkar, How cooperative are protein folding and unfolding transitions? Protein Sci. 25 (2016) 1924–1941.
[16] H. Gelman, M. Gruebele, Fast protein folding kinetics, Q. Rev. Biophys. 47 (2014) 95–142.
[17] S.C. Harrison, R. Durbin, Is there a single pathway for the folding of a polypeptide chain? Proc. Natl. Acad. Sci. 82 (1985) 4028–4030.
[18] J.B. Udgaonkar, Multiple routes and structural heterogeneity in protein folding, Annu. Rev. Biophys. 37 (2008) 489–510.
[19] R. Nevo, V. Brumfeld, R. Kapon, P. Hinterdorfer, Z. Reich, Direct measurement of protein energy landscape roughness, Sci. Rep. 6 (2005) 6–10.
[20] A.C. Apetri, W.K. Surewicz, Kinetic intermediate in the folding of human prion protein, J. Biol. Chem. 277 (2002) 44589–44592.
[21] A.C. Apetri, K. Surewicz, W.K. Surewicz, The effect of disease-associated mutations on the folding pathway of human prion protein, J. Biol. Chem. 23 (2004) 18008–18014.
[22] T. Hart, L.L.P. Hosszu, C.R. Trevitt, G.S. Jackson, J.P. Waltho, J. Collinge, A.R. Clarke, Folding kinetics of the human prion protein probed by temperature jump, Proc. Natl. Acad. Sci. U. S. A. 106 (2009) 5651–5656.
[23] A.C. Apetri, K. Maki, H. Roder, W.K. Surewicz, Early intermediate in human prion protein folding as evidenced by ultrarapid mixing experiments, J. Am. Chem. Soc. 128 (2006) 11673–11678.
[24] K.C. Chen, M. Xu, W.J. Wedemeyer, H. Roder, Microsecond unfolding kinetics of sheep prion protein reveals an intermediate that correlates with susceptibility to classical scrapie, Biophys. J. 101 (2011) 1221–1230.
[25] D.C. Jenkins, D.S. Pearson, A. Harvey, I.D. Sylvester, M.A. Geeves, T.J.T. Pinheiro, Rapid folding of the prion protein captured by pressure-jump, Eur. Biophys. J. 38 (2009) 625–635.
[26] G. Wildegger, S. Liemann, R. Glockshuber, Extremely rapid folding of the C-terminal domain of the prion protein without kinetic intermediates, Nat. Struct. Biol. 6 (1999) 550–553.
[27] R.P. Honda, K.I. Yamaguchi, K. Kuwata, Acid-induced molten globule state of a prion protein: crucial role of strand 1–helix 1–strand 2 segment, J. Biol. Chem. 289 (2014) 30355–30363.
[28] R.P. Honda, M. Xu, K. Yamaguchi, H. Roder, K. Kuwata, A native-like intermediate serves as a branching point between the folding and aggregation pathways of the mouse prion protein, Structure 23 (2015) 1–8.
[29] H. Yu, X. Liu, K. Neupane, A. Nath, A.M. Brigley, A. Solanki, A.N. Gupta, A.M. Brigley, A. Solanki, I. Sosova, M.T. Woodside, Direct observation of multiple misfolding pathways in a single prion protein molecule, Proc. Natl. Acad. Sci. 109 (2012) 5283–5288.
[30] R. Moulick, J.B. Udgaonkar, Identification and structural characterization of the precursor conformation of the prion protein which directly initiates misfolding and oligomerization, J. Mol. Biol. 429 (2017) 886–899.
[31] A.R. Fersht, Linear free energy relationships are valid! Protein Eng. Des. Sel. 1 (1987) 442–445.
[32] M.N. Berberan-Santos, J.M.G. Martinho, The integration of kinetic rate equations by matrix methods, J. Chem. Educ. 67 (1990) 375.
[33] I.E. Sánchez, T. Kiefhaber, Non-linear rate-equilibrium free energy relationships and Hammond behavior in protein folding, Biophys. Chem. 100 (2003) 397–407.
[34] F. Khan, J.I. Chuang, S. Gianni, A.R. Fersht, The kinetic pathway of folding of barnase, J. Mol. Biol. 333 (2003) 169–186.
[35] H. Taskent, J. Cho, D.P. Raleigh, Temperature-dependent Hammond behavior in a protein-folding reaction: analysis of transition-state movement and ground-state effects, J. Mol. Biol. 378 (2008) 699–706.
[36] J. Sabelko, J. Ervin, M. Gruebele, Observation of strange kinetics in protein folding, Proc. Natl. Acad. Sci. U. S. A. 96 (1999) 6031–6036.
[37] D.T. Leeson, F. Gai, H.M. Rodriguez, L.M. Gregoret, R.B. Dyer, Protein folding and unfolding on a complex energy landscape, Proc. Natl. Acad. Sci. 97 (2000) 2527–2532.
[38] R.D. Hutton, J. Wilkinson, M. Faccin, E.M. Sivertsson, A. Pelizzola, A.R. Lowe, P. Bruscolini, L.S. Itzhaki, Mapping the topography of a protein energy landscape, J. Am. Chem. Soc. 137 (2015) 14610–14625.
[39] L. Zhu, K. Ghosh, M. King, T. Cellmer, O. Bakajin, L.J. Lapidus, Evidence of multiple folding pathways for the villin headpiece subdomain, J. Phys. Chem. B 115 (2011) 12632–12637.
[40] H. Frauenfelder, S. Sligar, P. Wolynes, The energy landscapes and motions of proteins, Science 254 (1991) 1598–1603.
[41] P.G. Wolynes, Z.A. Luthey-Schulten, J.N. Onuchic, Fast-folding experiments and the topography of protein folding energy landscapes, Chem. Biol. 3 (1996) 425–432.
[42] G. Otting, E. Liepinsh, K. Wuethrich, Disulfide bond isomerization in BPTI and BPTI(G36S): an NMR study of correlated mobility in proteins, Biochemistry 32 (1993) 3571–3582.
[43] M.J. Grey, C. Wang, A.G. Palmer, Disulfide bond isomerization in basic pancreatic trypsin inhibitor: multisite chemical exchange quantified by CPMG relaxation dispersion and

chemical shift modeling, J. Am. Chem. Soc. 125 (2003) 14324–14335.
[44] S.H. Bae, G. Legname, A. Serban, S.B. Prusiner, P.E. Wright, H.J. Dyson, Prion proteins with pathogenic and protective mutations show similar structure and dynamics, Biochemistry 48 (2009) 8120–8128.
[45] O. Julien, S. Chatterjee, T.C. Bjorndahl, B. Sweeting, S. Acharya, V. Semenchenko, A. Chakrabartty, E.F. Pai, D.S. Wishart, B.D. Sykes, N.R. Cashman, Relative and regional stabilities of the hamster, mouse, rabbit, and bovine prion proteins toward urea unfolding assessed by nuclear magnetic resonance and circular dichroism spectroscopies, Biochemistry 50 (2011) 7536–7545.
[46] E. Langella, R. Improta, V. Barone, Checking the pH-induced conformational transition of prion protein by molecular dynamics simulations: effect of protonation of histidine residues, Biophys. J. 87 (2004) 3623–3632.
[47] E. Langella, R. Improta, O. Crescenzi, V. Barone, Assessing the acid–base and conformational properties of histidine residues in human prion protein (125–228) by means of pKa calculations and molecular dynamics simulations, Proteins Struct. Funct. Bioinforma. 70 (2008) 311–319.
[48] J. Garrec, I. Tavernelli, U. Rothlisberger, Two misfolding routes for the prion protein around pH 4.5, PLoS Comp. Biol. 9 (2013), e1003057.
[49] J. Zuegg, J.E. Gready, Molecular dynamics simulations of human prion protein: importance of correct treatment of electrostatic interactions, Biochemistry 38 (1999) 13862–13876.
[50] W.C. Guest, N.R. Cashman, S.S. Plotkin, Electrostatics in the stability and misfolding of the prion protein: salt bridges, self energy, and solvation, Biochem. Cell Biol. 381 (2010) 371–381.
[51] C.D. Waldburger, T. Jonsson, R.T. Sauer, Barriers to protein folding: formation of buried polar interactions is a slow step in acquisition of structure, Proc. Natl. Acad. Sci. 93 (1996) 2629–2634.
[52] I. Sengupta, S.H. Bhate, R. Das, J.B. Udgaonkar, Salt-mediated oligomerization of the mouse prion protein monitored by real-time NMR, J. Mol. Biol. 429 (2017) 1852–1872.
[53] S.S. Strickler, A.V. Gribenko, A.V. Gribenko, T.R. Keiffer, J. Tomlinson, T. Reihle, V.V. Loladze, G.I. Makhatadze, Protein stability and surface electrostatics: a charged relationship, Biochemistry 45 (2006) 2761–2766.
[54] R. Zwanzig, Diffusion in a rough potential, Proc. Natl. Acad. Sci. 85 (1988) 2029–2030.
[55] 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.
[56] C. Hyeon, D. Thirumalai, Can energy landscape roughness of proteins and RNA be measured by using mechanical unfolding experiments? Proc. Natl. Acad. Sci. 100 (2003) 10249–10253.
[57] A. Gershenson, L.M. Gierasch, A. Pastore, S.E. Radford, Energy landscapes of functional proteins are inherently risky, Nat. Chem. Biol. 10 (2014) 884–891.
[58] K. Henzler-Wildman, D. Kern, Dynamic personalities of proteins, Nature 450 (2007) 964–972.
[59] H.N. Motlagh, J.O. Wrabl, J. Li, V.J. Hilser, The ensemble nature of allostery, Nature 508 (2014) 331–339.
[60] Á. Tóth-Petróczy, D.S. Tawfik, The robustness and innovability of protein folds, Curr. Opin. Struct. Biol. 26 (2014) 131–138.
[61] E. Monsellier, F. Chiti, Prevention of amyloid-like aggregation as a driving force of protein evolution, EMBO Rep. 8 (2007) 737–742.
[62] L. Banaszak, N. Winter, Z. Xu, D.A. Bernlohr, S. Cowan, T.A. Jones, Lipid-binding proteins: a family of fatty acid and retinoid transport proteins, Adv. Protein Chem. 45 (1994) 89–151.
[63] P.G.W. Gettins, Serpin structure, mechanism, and function, Chem. Rev. 102 (2002) 4751–4804.
[64] L. Masino, G. Nicastro, A. De Simone, L. Calder, J. Molloy, A. Pastore, The Josephin domain determines the morphological and mechanical properties of ataxin-3 fibrils, Biophys. J. 100 (2011) 2033–2042.
[65] R. Moulick, R. Das, J.B. Udgaonkar, Partially unfolded forms of the prion protein populated under misfolding-promoting conditions: characterization by hydrogen exchange mass spectrometry and NMR, J. Biol. Chem. 290 (2015) 25227–25240.
[66] K.R. Srivastava, L.J. Lapidus, Prion protein dynamics before aggregation, Proc. Natl. Acad. Sci. 114 (2017) 3572–3577.
[67] R.R. Goluguri, J.B. Udgaonkar, Microsecond rearrangements of hydrophobic clusters in an initially collapsed globule prime structure formation during the folding of a small protein, J. Mol. Biol. 428 (2016) 3102–3117.
[68] M.C.R. Shastry, S.D. Luck, H. Roder, A continuous-flow capillary mixing method to monitor reactions on the microsecond time scale, Biophys. J. 74 (1998) 2714–2721.
[69] H. Roder, K. Maki, H. Cheng, Early events in protein folding explored by rapid mixing methods, Chem. Rev. 106 (2006) 1836–1861.
[70] U.S. The Mathworks Inc., Natick, MA, MATLAB and Statistics Toolbox Release 2012a, (n.d.).