ELSEVIER

Biophysical Chemistry 111 (2004) 159–171

Biophysical Chemistry

www.elsevier.com/locate/bpc

The folding pathway of ubiquitin from all-atom molecular dynamics simulations

Neelan J. Marianayagam*, Sophie E. Jackson

Centre for Protein Engineering, Department of Chemistry, University of Cambridge, Cambridge CB2 1EW, UK

Received 24 March 2004; received in revised form 14 May 2004; accepted 17 May 2004
Available online 20 July 2004

Abstract

The folding (unfolding) pathway of ubiquitin is probed using all-atom molecular dynamics simulations. We dissect the folding pathway using two techniques: first, we probe the folding pathway of ubiquitin by calculating the evolution of structural properties over time and second, we identify the rate determining transition state for folding. The structural properties that we look at are hydrophobic solvent accessible surface area (SASA) and C_α-root-mean-square deviation (rmsd). These properties on their own tell us relatively little about the folding pathway of ubiquitin; however, when plotted against each other, they become powerful tools for dissecting ubiquitin's folding mechanism. Plots of C_α-rmsd against SASA serve as a phase space trajectories for the folding of ubiquitin. In this study, these plots show that ubiquitin folds to the native state via the population of an intermediate state. This is shown by an initial hydrophobic collapse phase followed by a second phase of secondary structure arrangement. Analysis of the structure of the intermediate state shows that it is a collapsed species with very little secondary structure. In reconciling these observations with recent experimental data, the transition that we observe in our simulations from the unfolded state (U) to the intermediate state (I) most likely occurs in the dead-time of the stopped flow instrument. The folding pathway of ubiquitin is probed further by identification of the rate-determining transition state for folding. The method used for this is essential dynamics, which utilizes a principal component analysis (PCA) on the atomic fluctuations throughout the simulation. The five transition state structures identified in silico are in good agreement with the experimentally determined transition state. The calculation of ϕ-values from the structures generated in the simulations is also carried out and it shows a good correlation with the experimentally measured values. An initial analysis of the denatured state shows that it is compact with fluctuating regions of nonnative secondary structure. It is found that the compactness in the denatured state is due to the burial of some hydrophobic residues. We conclude by looking at a correlation between folding kinetics and residual structure in the denatured state. A hierarchical folding mechanism is then proposed for ubiquitin.

Keywords: Transition state; Intermediate; Essential dynamics; Kinetics; β-grasp

1. Introduction

The use of computer simulations for the analysis of protein folding pathways has been the subject of many studies over the past 15 years [1]. Both simplified and all-atom representations of proteins in conjunction with Monte Carlo or molecular dynamics algorithms have given much insight into the folding process [2]. Issues such as kinetically foldable sequences, the shape of the energy landscape and the actual folding mechanism have been addressed in these studies.

In the investigation of which sequences will actually fold to a unique native state, lattice models have been an invaluable tool. Studies on a14 mer on a 2-D lattice with two types of amino acids (hydrophobic and hydrophilic) show that a large number of these sequences are able to fold to compact conformations [3]. From an analysis of fully enumerated Monte Carlo simulations, it was concluded that the ability for a sequence to form a unique compact structure is directly proportional to the chain length of the polypeptide [3].

Another factor that has been used to explain kinetic foldability is the random energy model (REM). The REM

* Corresponding author. Present Address: School of Molecular and Microbial Biosciences, University of Sydney, Building G08, Sydney, NSW 2006, Australia. Tel.: +61-2-9351-6023; fax: +61-2-9351-4726.
E-mail address: neelanm@mmb.usyd.edu.au (N.J. Marianayagam).

0301-4622/$ - see front matter © 2004 Elsevier B.V. All rights reserved.
doi:10.1016/j.bpc.2004.05.009

was first derived for the description of the kinetics of spin glasses [4]. When applied to protein folding, this model states that in order for a sequence to fold, there must be a substantial energy gap between native and denatured states [5,6]. If one knows this energy gap, then the probability of folding for a specific sequence can be estimated. The REM gives rise to a spectrum of states, with the denatured state containing a large number of conformations and the native state with a significantly smaller conformational space.

The investigation of folding kinetics and actual folding mechanisms has also been probed using simplified models. It has been proposed by Thirumalai and Klimov [7] that protein folding kinetics are determined by two characteristic temperatures, T_c and T_f, which are the collapse and folding temperatures, respectively, of the polypeptide chain. These parameters can be combined to give a σ factor. The σ factor is a measure of how close the T_c and T_f are to each other. Thirumalai and Klimov [7] have found a linear correspondence between the folding time and σ. As σ increases, the folding time increases. The σ factor also gives insight into the folding mechanism itself. For $\sigma = 0-0.3$, folding is said to be fast and takes place by a nucleation–condensation mechanism [7]. When $0.3 < \sigma < 0.65$, the folding time is said to be moderate and the native state is achieved via a kinetic partitioning mechanism (KPM) [7]. The KPM showed a fraction of the molecules folding through a nucleation-collapse mechanism while the rest populated a stable intermediate on the way to the native state [7]. When values of $\sigma > 0.65$, the folding is extremely slow, and it is suggested that these sequences represent those that will need chaperones to fold [7].

All-atom molecular dynamics simulations have also provided insights into the folding process. Much of the work in this area has shown reasonable agreement to experimental data. The simulations of Li and Daggett have been able to reproduce experimental results for CI2 [8], barnase [9], and for FKBP12 [10]. Due to the complexity of all atom models, the folding (unfolding) is probed by using high temperatures to speed up the kinetics. Li and Daggett [8,11] used high-temperature MD simulations of CI2 to probe the unfolding of this protein in silico. In this work, they generated four transition state structures and found that they agreed well with experimental results. Theoretical ϕ-values were calculated by looking at the loss of contacts upon mutation, and these also agreed with the experimentally observed ϕ-values [8]. In a later study, ϕ-values measured by calculation of free-energy differences between wild type and mutant in silico; utilizing free-energy perturbation techniques also seemed to agree with experimental results [12]. Ladurner et al. [13] were also able to identify unfavorable interactions in the transition state. When these interactions were removed by mutation, the experimentally determined refolding rates were actually faster than for the wild-type protein.

So what are the in silico techniques that we can use to identify the transition state from all-atom protein unfolding simulations? It is difficult to identify the transition state energetically in high-temperature simulations, so structural methods of identification must be employed. Kazmirski et al. [14] have used several techniques, based on the dynamics of the polypeptide chain, to identify the transition state.

Another method to identify the transition state was used by Pande and Rokhsar [15]. In their thermal denaturation simulations of a β-hairpin fragment, they identified candidate transition state structures by looking at major structural changes and at corresponding spikes in the heat capacity profile [15]. They take these putative transition states and simulate at lower temperatures. In theory, a transition state structure should fold to the native state with a probability $p = 0.5$ [16]. They calculate this p value by carrying out hundreds of simulations on the candidate structures [15]. This method is feasible for small systems; however, for larger proteins carrying out hundreds of simulations on robust, all-atom models becomes computationally expensive.

While the work of Daggett and co-workers reproduces experimental data well, it gives very little insight into the shape of the potential energy surface (PES). The structure and orientation of a molecule is determined by the topology of its PES. For a simple system with well-defined interactions, the calculation of the PES, though intellectually challenging, can be carried out to a high degree of accuracy [17,18]. However, for systems with many degrees of freedom, transient interactions and complicated reaction kinetics, the determination of an energy surface is nontrivial. Such is the case for proteins.

A statistical description of the energy surface has been developed through the use of analytical theories of phase transitions and disordered systems. This view has been further investigated by the simulation of simplified homopolymers and random heteropolymers [19]. Based on these studies, the shape of a protein's PES is said to be a rugged funnel biased toward its basin, which represents the native state. While these theories are abstract in nature, proponents of the folding funnel have made an effort to relate their ideas to experimental data on fast folding proteins [20].

Methods to generate representations of the energy surface for all-atom models have been developed by Shea and Brooks III [1]. They utilize molecular dynamics algorithms to carry out thermal denaturation of all-atom models. From these thermal denaturations, they generate various nonnative conformations of the protein in question. These conformations are binned into coordinate space to generate initial conditions. These initial conditions are used to carry out further biased sampling of these nonnative conformations at 298 K. From these room temperature simulations, the free energy of the conformation is calculated as a function of the spatial coordinates. The calculated free energies of each conformation are then projected onto coordinate space to

give elegant free-energy representations of the protein. The coordinates onto which the free energy is projected represent progress variables that can be used to monitor the folding reaction. The coordinates used in these studies are the number of native contacts (N) and the radius of gyration (R_g). The folding mechanism can then be determined for this free energy surface. Proteins with mainly local contacts such as all helical proteins are expected to have diagonal profiles in N and R_g space which represent simultaneous collapse and formation of native structure [1]. However, proteins with more long-range interactions such as β-sheet containing proteins show more of an L-shaped free-energy profile which represents two distinct kinetic phases, one for collapse and the other for rearrangement to the native state [1].

This work describes thermal denaturation studies of ubiquitin, a 76 residue mixed α/β protein which has a β-grasp fold. The structure of ubiquitin is shown in Fig. 1A [21]. The folding of ubiquitin has been under heated debate by many experimentalists in the folding community. The initial experiments by Khorasanizadeh et al. [22] showed that ubiquitin folded through a burst-phase intermediate. However, this claim has been refuted by the recent work of Krantz and Sosnick [23] who have attributed the extra phase to instrumental limitations. It was concluded that the folding of this protein was two state. However, recent experiments in this laboratory have shown that an intermediate state of ubiquitin can be populated under certain conditions [56]. It is possible that this intermediate is a metastable burst-phase intermediate that is populated during the dead time of the stopped-flow instrument.

In this paper, further evidence is provided that ubiquitin folds via the population of an intermediate state. Our simulations show folding with an initial collapse of the polypeptide chain and then rearrangement to the final folded structure. This is shown by plots of the hydrophobic solvent accessible surface area (SASA) against the C_α-root-mean-square deviation (rmsd) over the time course of the simulation.

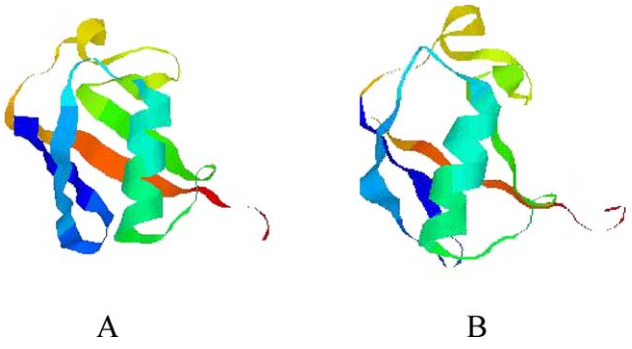

Fig. 1. (A) Crystal structure of ubiquitin. (B) Structure of ubiquitin after 500 ps of native state simulation.

The folding pathway is then probed further by the identification of the rate-determining transition state. In this study, we adopt a new strategy for identifying the transition state. Essential dynamics is a method that was developed to analyze native state simulations of proteins [24]. This method utilizes principal component analysis (PCA) on the actual coordinates of the system and thus gives the essential motion of the protein in phase space. The identification of the transition state from essential dynamics is more accurate than just looking at the variation of several structural properties. This is due to the fact that only the important motion of the protein through phase space is considered. The transition states, which we will call the transition state ensemble (TSE), that we identify for the 76-residue protein ubiquitin through this method, are in good agreement with the experimentally mapped transition state [57]. This study shows that essential dynamics is not just useful for the analysis of native state simulations but can be an invaluable tool for the identification of transition state structures from thermal denaturation reactions.

2. Materials and methods

All simulations were carried out using the GROMOS 96 force field [52] within the GROMACS software package [53]. The native state simulations were carried out at 298 K while the thermal denaturations were carried out at 500 K. Both were done at a constant pressure of 1 atm. Electrostatics were dealt with using a 8-Å coulombic cut-off.

The native-state simulations were run for 2 ns and the thermal denaturations were run for 2.5 ns, except in one case, where it was run for 3.0 ns. Six of the denaturations were run from a structure, generated from the native state ensemble. While one was run from the crystal structure (1UBQ) [21], each structure was fully solvated with SPC water [54] in a cube with an edge length of 20 Å. The structures were equilibrated in two steps: first, the structures were energy-minimized using steepest descents for 1000 steps for each of the native state simulations, and second, a position-restrained MD run was carried out, which holds the protein atoms fixed and allows the solvent to equilibrate around the solute. With the thermal denaturation simulations, energy minimizations were varied between 200 and 1000 steps to ensure sampling of the native state ensemble.

The atoms in the system were given initial velocities according to a Maxwellian distribution. The system was allowed to evolve according to Newton's equations of motion, with the equations being integrated every 2 fs using the Verlet algorithm. The progress of the simulations was monitored by calculating several structural parameters over time: Cα-rmsd (root-mean-square deviation), radius of gyration (R_g) and hydrophobic solvent accessible surface area (SASA).

Each of these structural properties was calculated within the software package GROMACS. The Cα-rmsd is calculated by using a least-squares fitting routine which calculates the fluctuation in this structural property over time. This can be expressed in the following equation:

$$\mathrm{RMSD}(t_1, t_2) = \left[\frac{1}{N}\sum_{i=1}^{N}(r_i(t_1) - r_i(t_2))^2\right]^{\frac{1}{2}} \quad (1)$$

where $t_2 = 0$ is the time of the reference structure (usually the starting structure for the simulation), t_1 is the particular time point in the simulation, r_i is the position of atom i at the specified time and N is the number of atoms being considered. For example, in a 76-residue protein, there are 76 α carbons, so $N = 76$.

The radius of gyration R_g is defined in polymer physics as the root-mean-square distance of the collection of atoms from their common center of mass. The radius of gyration is calculated using a center of mass weighted equation:

$$R_g = \left(\frac{\sum_i r_i^2 m_i}{\sum_i m_i r_i}\right) \quad (2)$$

where m_i is the mass of atom i and r_i is the position of atom i with respect to the center of mass of the molecule.

SASA is calculated by using a circular probe that is 1.4 Å in diameter. The probe is rolled around the surface of the molecule and the accessible surface area is calculated.

2.1. Calculation of principal components

The calculation of the eigenvectors and eigenvalues, and their projection along the first two principal components, was carried out according to protocol of Amadei et al. [24] within the GROMACS software package. This method divides the conformational space of the protein into two subspaces, an essential subspace and a nonessential, physically constrained subspace. The essential subspace is described by the unconstrained, anharmonic motion of the positional fluctuation of the atoms. The motion in the remaining subspace is defined by a narrow Gaussian distribution.

The first step in essential dynamics is the generation of nonmass weighted coordinate matrix. For an N atom system, this will be $3N \times 3N$ matrix which we will call **A**. The covariance matrix of **A**, which we will call **C**, is defined by the following equation:

$$\mathbf{C} = \mathbf{A}^T\mathbf{A} \quad (3)$$

where T is the transpose of the matrix. The transpose is found by exchanging the rows and columns of a matrix. The eigenvectors of the covariance matrix are the principal components. This then turns into a simple eigenvalue problem:

$$\mathbf{Cx} = \lambda\mathbf{x} \quad (4)$$

where λ is the eigenvalue associated with the eigenvector **x**. For an N atom system, there are $3N$ eigenvectors and associated eigenvalues. Eq. (4) can be simplified to the following:

$$(\mathbf{C} - \lambda\mathbf{I})\mathbf{x} = 0 \quad (5)$$

where **I** is the identity matrix. A solution to Eq. (5) can be obtained by taking the determinant of the term in parenthesis and setting it to equal to zero. This will involve the expansion of a large polynomial. While this technique might be straightforward, it can be computationally expensive for larger systems. A much more desirable technique is to use matrix diagonalization. The diagonal matrix, **D**, of the covariance matrix is defined by the following:

$$\mathbf{D} = \mathbf{U}^{-1}\mathbf{CU} \quad (6)$$

The matrix **U** contains the eigenvectors and **D** is a matrix of the corresponding eigenvalues. The eigenvector with the highest eigenvalue is considered the first principal component, the eigenvector with the second highest eigenvalue is considered the second principal component and so on. It has been shown that the majority of the motion for proteins can be accounted for in the first two principal components [24]. The dynamics of a protein can thus be analyzed by projecting its atomic motion during a MD simulation onto its first two principal components. Essential dynamics is a powerful tool for monitoring protein dynamics in phase space because the observed motion is unconstrained and represents the atomic fluctuations of the protein.

2.2. Identification of the transition state ensemble from projections

Identification and analysis of the transition state and denatured state was carried out on five out of the seven simulations. The transition state ensemble was identified using the reasoning of Kazmirski et al. [14]. Kazmirski et al. [14] define the region just before the first major structural change as the transition state. The reason behind this is that, for a process where the enthalpy increases and the entropy changes very little, a high free-energy barrier is produced [14]. As such, we choose three structures from each simulation that occur in the transiently populated region just before the native state cluster (see plot in Fig. 5). It is interesting to note that

this transiently populated region occurs in all five simulations where the transition state was analysed. Three putative transition state structures were generated for each of the simulations for a total of 15 structures which comprise the transition state ensemble. Five representative structures from the transition state ensemble are shown in Fig. 6.

2.3. Identification of the denatured state ensemble

The denatured state ensemble was identified from plots in Fig. 3. Clustering occurs at high values of SASA and rmsd (see Fig. 3). These clusters occur at the end points of each of the simulations and represent the denatured state. We pick 6 structures from this denatured state cluster for each simulation, for a total of 30 structures that comprise the denatured state ensemble. Five representative structures of the denatured state ensemble are shown in Fig. 7. The quantification of the transition state ensemble and the denatured state ensemble, described below, represents average calculations over 15 structures and 30 structures, respectively.

2.4. Calculation of SASA and z-values

The percentage of solvent accessible surface area (%SASA) of the various structures was calculated using the program GETAREA 1.1 [55]. We define z-values as a measure of the difference in SASA between the folded and the unfolded state (z_{U-F}) and the folded state and the transition state ($z_{\ddagger-F}$). These were calculated per residue using the following equations:

$$z_{U-F} = \%SASA^{U} - \%SASA^{F} \quad (7)$$

$$z_{\ddagger-F} = \%SASA^{\ddagger} - \%SASA^{F} \quad (8)$$

where $\%SASA^{U}$ and $\%SASA^{F}$ are the %SASA in the unfolded and folded state, respectively, and $\%SASA^{\ddagger}$ is this value in the transition state ensemble. The final values obtained for z_{U-F} and $z_{\ddagger-F}$ are obtained by averaging of 30 unfolded state and 15 transition state structures, respectively. The value for $\%SASA^{F}$ is obtained by averaging the values for the crystal structure and the structure generated at 500 ps of native state simulation.

2.5. Calculation of ϕ-values from simulations

The calculation of ϕ-values from the simulation was carried out using the method of Li and Daggett [8]. This method involves taking the ratio of the number of contacts in the transition state to the number of contacts in the native state. This is done per residue (with the exception of glycine residues). This procedure was carried out for all 15 structures in the transition state ensemble, and then averaged.

3. Results

Seven thermal denaturation simulations were carried out at 500 K. One simulation was carried out from the crystal structure of ubiquitin (1UBQ) [21], see Fig. 1A, while the remaining six were run from a structure generated after 500 ps of native state simulation (i.e., at 298 K), see Fig. 1B. The average rmsd for the native state simulations does not deviate greater than 2.5 Å from the crystal structure (data not shown). This is slightly greater than the value obtained by Alonso and Daggett [25]; however, they omit some of the C-terminal residues in their calculation, which tend to have large rmsd fluctuations. Fig. 2A illustrates the changes rmsd over time for four of the thermal denaturations. In all cases, a significant change in rmsd ranging form 6–10 Å

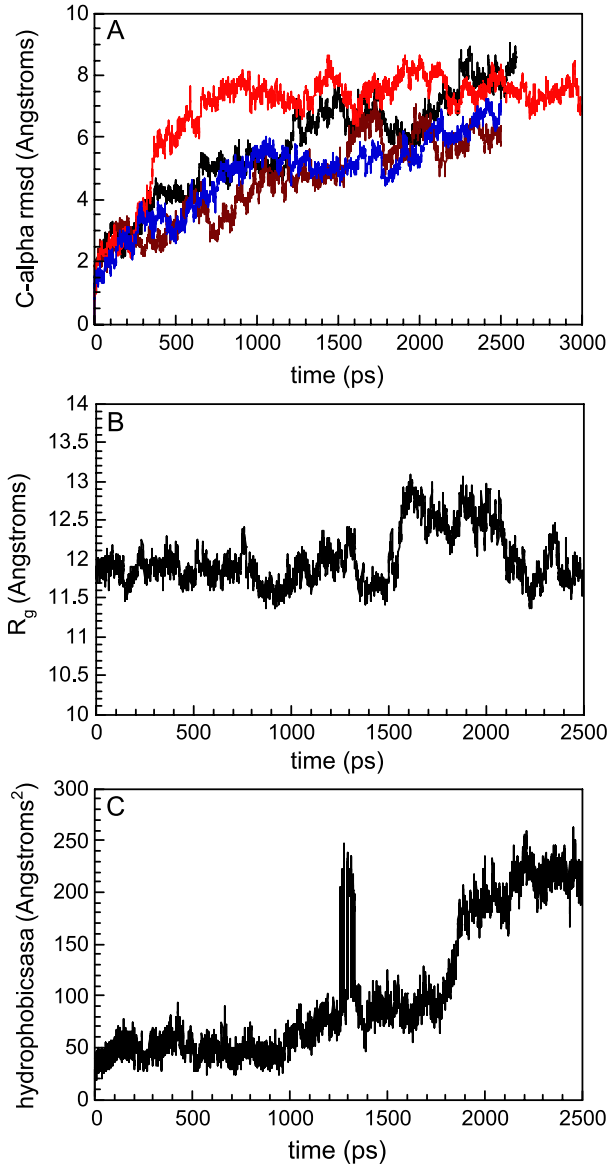

Fig. 2. Evolution of structural properties over time during the course of the simulation: (A) C_α-rmsd; (B) Radius of gyration; (C) Hydrophobic Solvent accessible surface area.

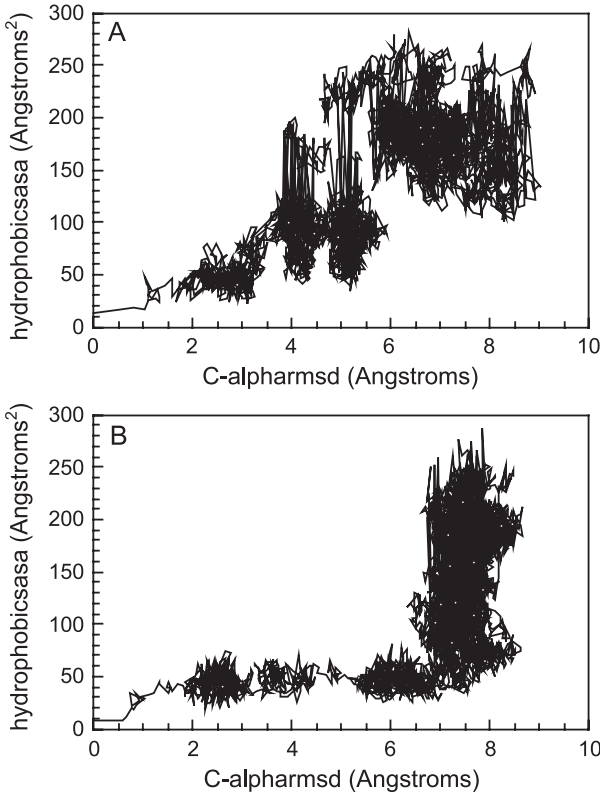

Fig. 3. Phase space trajectories for the folding of ubiquitin. (A) Diagonal profile representing two-state kinetics. (B) L-shaped profile, which is observed for six out of the seven simulations, indicates the population of an intermediate state.

occurs by the end of the simulation. The native state simulation stays within 2 Å of the crystal structure (data not shown), which is in agreement with the native state simulations carried out by Daggett et al. [25]

Fig. 2B shows a plot of R_g against time for one simulation. The plots for the other simulations are not shown because they are similar and they fluctuate around the native value of 12 Å. Based on rmsd, the native structure has been lost; however, the unfolded state is still compact as measured by R_g. A simulation carried out for twice as long (i.e., 5000 ps) also shows that the R_g fluctuates around 12 Å (data not shown). The implications of these results are discussed later.

In Fig. 2C, a representative plot of SASA against time is shown; this profile shows a cooperative transition, from the native state, in which most of the hydrophobic surface area is buried, to an unfolded ensemble which has significant amount of hydrophobic surface area exposed to solvent. This cooperative transition is observed for the six other simulations.

The plots generated in the previous section are a good indicator of the progress of the actual simulation. However, they tell us very little about the global folding (unfolding). It is difficult to gain knowledge about a possible folding mechanism from these plots. Monitoring how structural elements vary concurrently over time, however, will give us an idea of the kinetics of the system, as long as the properties monitored probe different structural features. In simulations of CI2, Lazaridis and Karplus [26] plot R_g against rmsd. This works well for analyzing the folding of CI2; however, for a system like ubiquitin, these plots would not be effective because R_g does not change throughout the course of the denaturations.

Instead we choose to monitor the change in SASA with rmsd. These two properties probe different structural features of the polypeptide. SASA looks at how the hydrophobic core evolves with time while rmsd looks at backbone dynamics. Essentially, we have probes for tertiary and secondary structure. Two plots of SASA against rmsd are shown in Fig. 3, which were calculated from two different simulations. The plot in Fig. 3B has an L-shaped profile which is observed in six out of the seven simulations. However, the plot in Fig. 3A seems to have more of a diagonal movement in SASA and Cα-rmsd space. This type of diagonal profile is observed in only one of the simulations.

Another interesting feature of these plots are the clusters that appear. The clusters represent stable states that are populated on the folding pathway. The first cluster in each plot represents the native state, while the final cluster populated at high values of SASA and rmsd is attributed to the denatured state ensemble. The first cluster populated at the bottom of the L shape represents an intermediate state, see Fig. 3B. Structures corresponding to this cluster were generated in order to investigate the properties of this intermediate state. Four representative structures from the intermediate state ensemble are shown in Fig. 4.

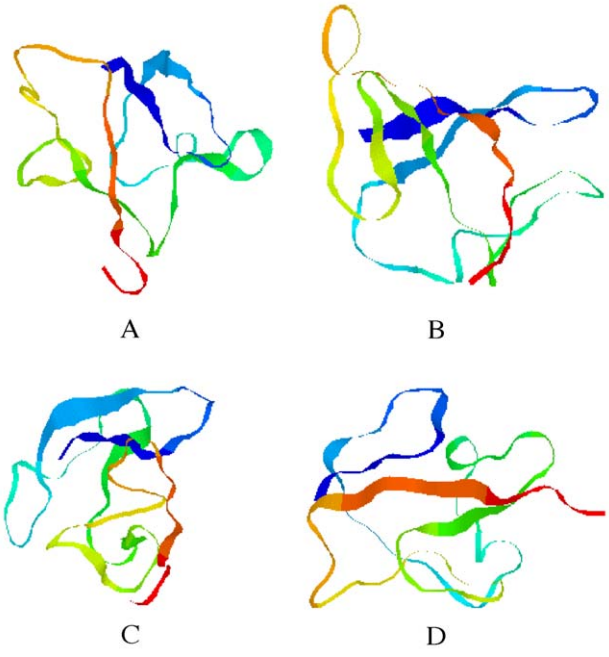

Fig. 4. Representative structures of the intermediate state populated on ubiquitin's folding pathway.

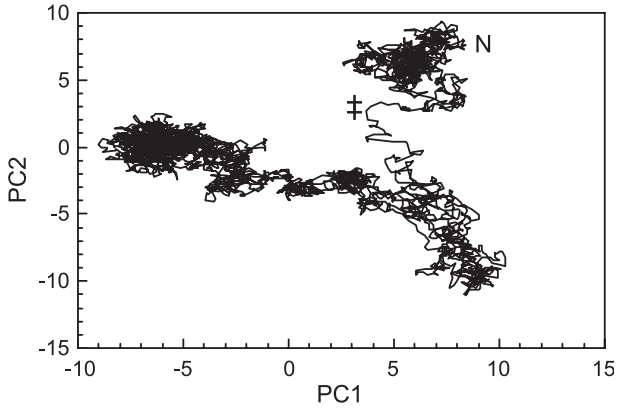

Fig. 5. Projection of the motion of ubiquitin during a thermal denaturation onto its first two principal components for atomic motion. The cluster corresponding to the native state is marked with the letter N. The region corresponding to the transition state is represented by the ‡ symbol.

The plot in Fig. 5 shows the motion of ubiquitin in phase space during a thermal denaturation simulation, projected along its first two principal components for atomic motion. Plots like these were generated for four other simulations (data not shown). All plots have a cluster corresponding to the native state, with a transiently populated area, which corresponds to the transition state. The transition states are identified as points just before population of the native state cluster.

Fig. 6 shows the structures of five representative transition states from the transition state ensemble. The five transition state structures are all slightly different; however, the one unifying feature is that β-1 and β-2 are structured in each. This is in good agreement with the experimental ϕ-values. The helix is partially formed in TS1, 3, 4 and 5, which is also in agreement with experimental results (TS1 represents a transition state structure generated from simulation 1, etc.).

Fig. 7 shows five representations of the denatured state. The structures are all relatively compact in agreement with calculated radius of gyration data. However, the native structure is lost as shown by high values of rmsd and SASA. The denatured state ensemble has varying degrees of residual structure. In D1 and D2, β-3 and β-4 are almost native-like and are interacting with each other. In D3,

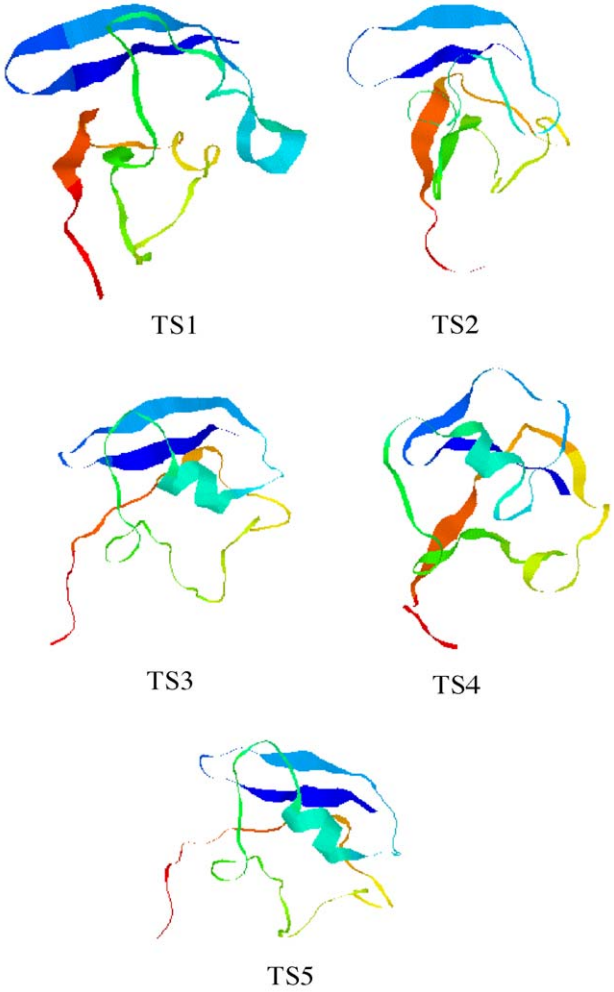

Fig. 6. The rate determining transition state ensemble for the folding of ubiquitin.

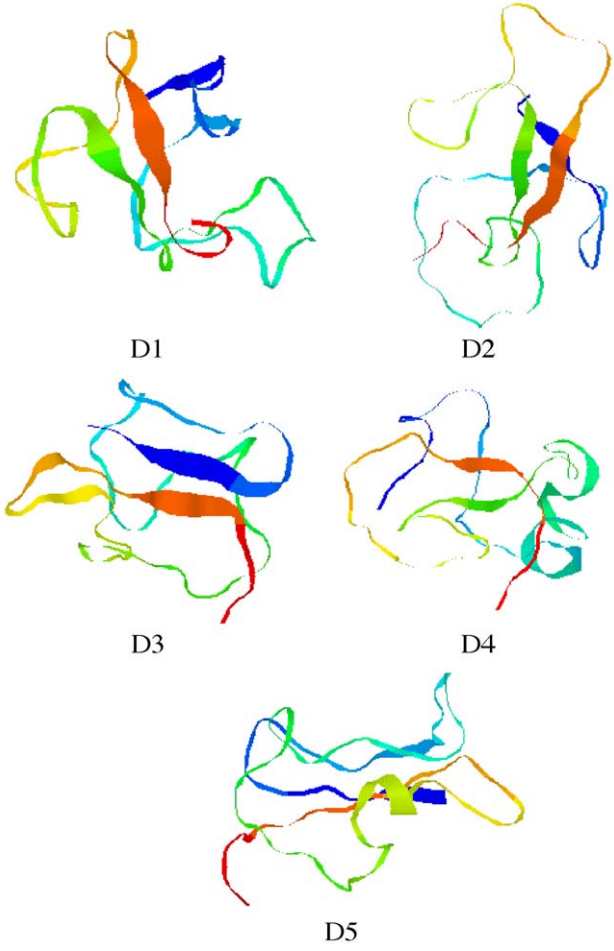

Fig. 7. The denatured state ensemble of ubiquitin.

nonnative interactions seem to be present between β-1 and β-3. In D4, there is nonnative helical structure and a small amount of structure for β-3 and β-4. In D5, there is very little β structure, but there is some nonnative helical structure.

Fig. 8 shows plots of calculated z_{U-F} and $z_{\ddagger-F}$ for each residue averaged over the five simulations. These values represent the change in percent solvent accessible surface area between the folded and unfolded state (z_{U-F}) and between the folded state and the transition state ($z_{\ddagger-F}$). These values are calculated using Eqs. (7) and (8) in the methods section. From the plots in Fig. 8B, one can identify nonnative interactions in the transition state. These can be defined as residues with negative z-values, which would indicate that they are actually more buried in the transition state than they are in the native state. Note that this is the case with Asn 60 and with Gln 62. Another interesting point about the TSE is that residues 5 and 7 (Val and Leu) are nearly as buried in the TSE as they are in the native state.

The plots in Fig. 8 give us a profile of the types of interactions present in both the denatured state and the transition state ensembles. In the denatured state ensemble, nearly half the residues are as buried or more buried in the denatured state as they are in the native state. Most of these amino acids are hydrophobic, indicating that the denatured state is stabilized by nonspecific hydrophobic interactions.

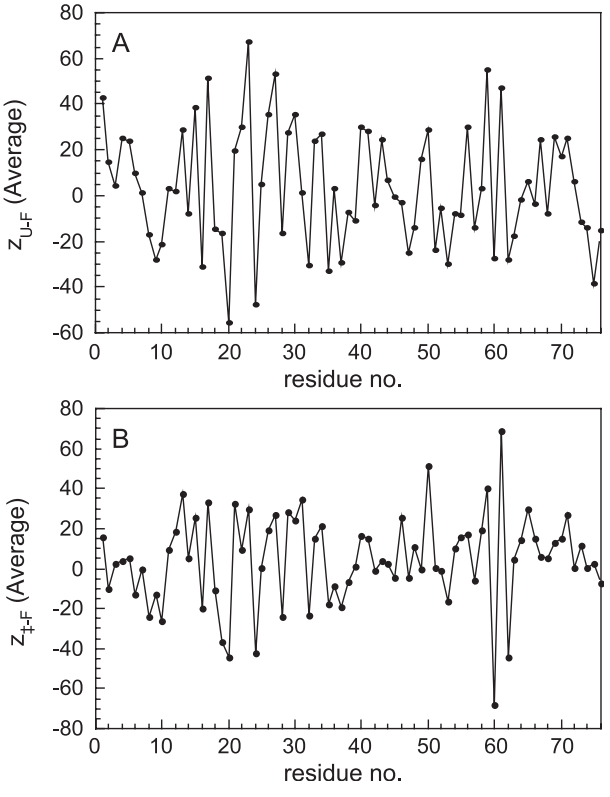

Fig. 8. Plot of z_{U-F} (A) and $z_{\ddagger-F}$ (B) values against residue number averaged over five simulations.

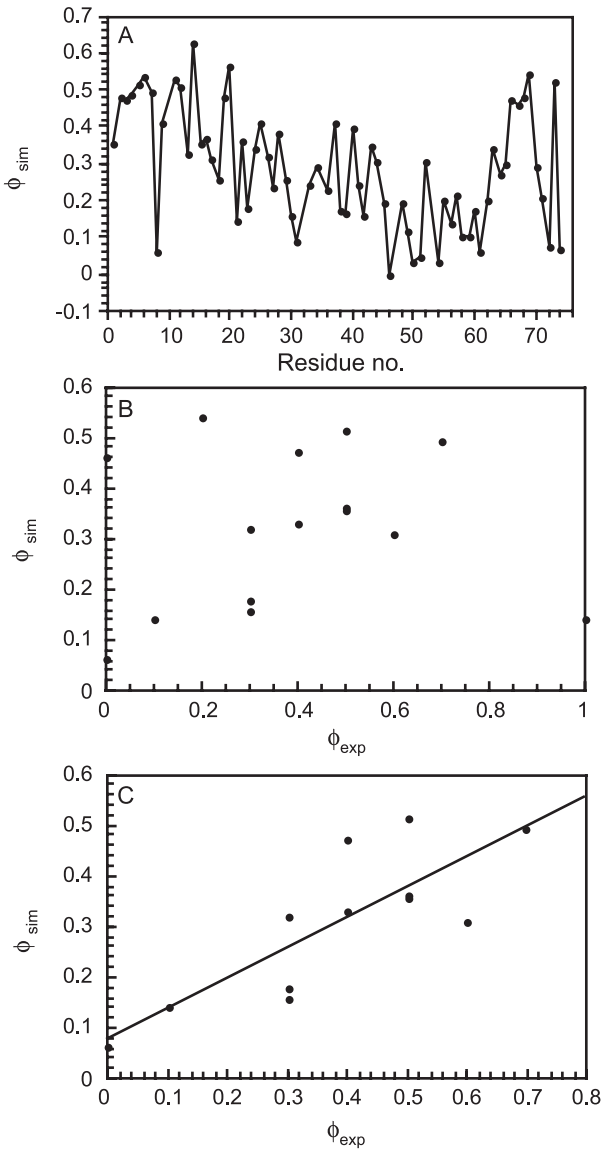

Fig. 9. (A) ϕ-values calculated from the simulations. (B) Correlation between ϕ_{exp} and ϕ_{sim}. (C) Correlation between ϕ_{exp} and ϕ_{sim} with three of the ϕ-values removed.

To make more direct comparisons between experiment and theory, ϕ-values were calculated from simulations (ϕ_{sim}) as described in the methods section. These results are presented in Fig. 9. In Fig. 9A, one can see the ϕ_{sim}-values are high in the N-terminal region of the protein and generally low in the C-terminal region, which is in good agreement with the experimental data [57]. However, when plotted against the experimental ϕ-values (ϕ_{exp}) in Fig. 9B, a poor correlation is observed ($R = 0.09$). When three of the ϕ-values are removed, there is a much stronger correlation between experiment and simulation ($R = 0.81$). The significance of these results is mentioned in the discussion.

4. Discussion

It is best to describe the plots in Fig. 3 as phase space trajectories for the folding of ubiquitin. They give insight into the folding mechanism while at the same time, they show stable states populated on the pathway. These stable states are represented by clusters in the plots of rmsd vs. SASA. How do these plots differ from those presented by Shea and Brooks III [1]? Those plots are projections of the potential of mean force onto structural space, while our plots are direct probes of structural space; that is, how does one structural quantity (rmsd) vary with another (SASA)? Shea and Brooks [1] have been able to gain information about the folding pathway of the proteins studied by their biased sampling method. In these cases, two-state folders show a more diagonal shape representing simultaneous collapse of the chain and the formation of native structure. Proteins with multistate kinetics would exhibit an L-shaped free-energy profile, indicating that the collapse of the chain and the formation of the final native state are two distinct steps.

In our simulations of ubiquitin, trajectories in phase space (Fig. 3) give us insight into the folding kinetics of ubiquitin. When one looks at rmsd and SASA, two distinct elements of structure are probed. Chain collapse can be monitored by looking at the evolution of SASA over time, while formation of native structure can be probed by monitoring rmsd. These quantities, when plotted together, provide a powerful tool for gaining insights into the folding pathway. Plots with a diagonal profile in SASA and rmsd space show that folding takes place with two-state kinetics, while plots which show an L-shaped profile indicate a multistate folding process. This can be likened to the σ factor of Thirumalai and Klimov [7], which measures how close chain collapse and folding are to each other.

Another interesting feature of these plots are the clusters that seem to emerge in these profiles. We define clusters in the SASA–rmsd space as stable states along the folding pathway. Clusters arise due to smaller than the mean fluctuations in rmsd and SASA during the course of the simulations. These small fluctuations must be due to stabilizing interactions which cause these small changes in the structural parameters. From these clusters in structural space, it is evident that stable states are populated on the folding pathway. We can gain insight into the kinetics of the system, and therefore, these plots can be likened to kinetic profiles for the folding of ubiquitin.

Now that we have generated these profiles, what information can we gain from these plots? In six out of the seven simulations, we see an L-shaped profile in SASA–rmsd space. As stated earlier, this is indicative of a noncooperative folding process. One can even say that this shows multistate kinetics, with an initial collapse of the polypeptide chain representing the first phase and the formation of the native state as the second phase. This tells us that ubiquitin folds via the population of an intermediate state that is almost as collapsed as the native state. Further evidence for the population of stable, substructured phases is present in the clusters that appear in the profiles. The structure of the intermediate is shown in Fig. 4. While the secondary structure seems to fluctuate throughout the ensemble, the one unifying feature is the degree of collapse of the structures. After the population of the intermediate, there is no rearrangement of the hydrophobic core. There is the formation of more secondary structural elements followed by relaxation of SASA to the native state value.

How can we reconcile these plots with experimental data? The initial collapse of the hydrophobic core that occurs is observed in six out of the seven simulations; most likely, they occur too quickly to be detected experimentally by conventional stopped-flow techniques. Recent experiments show that a burst phase is present in the folding kinetics of ubiquitin [56]. It has also been shown that it is possible to stabilize an intermediate state of ubiquitin under certain experimental conditions [56]. This indicates the presence of a relatively low energy intermediate state on the ubiquitin energy surface. The data from both experiment and theory provide strong evidence that the folding of ubiquitin is not robustly two state.

Simulations of ubiquitin using a pathway-generator algorithm seem to show apparent two-state kinetics [27]. However, these simulations are carried out on a coarse-grained model with implicit solvent. The simulations mainly probe structure by looking at backbone dynamics and hydrogen bond formation. Very little information about tertiary interactions can be gained from this work. Further simulations on an off-lattice model of ubiquitin seem to show a similar model of folding that we present here [28]. In this case, they observe that ubiquitin folds via the population of a collapsed intermediate state with little secondary structure [28]. All-atom Monte Carlo simulations of protein G which has an ubiquitin-like fold also show multistate kinetics with the population of a collapsed intermediate state [29]. Alonso and Daggett [25] generated nonnative conformations of ubiquitin using high-temperature, all-atom molecular dynamics simulations. In these simulations, they also observe a hydrophobic collapse phase when they quench the simulations to lower temperatures; however, they do not identify putative transition state structures [25].

4.1. Analysis of the rate-determining transition state

The transition state ensemble identified from essential dynamics is in good agreement with the experimentally mapped transition state [57]. The highest experimental ϕ-values are found in the first 28 residues of the protein, indicating that both β-1 and β-2 and part of the α-helix is structured in the transition state. This is, in general, true for the transition state ensemble generated from each simulation as can be seen in Fig. 6.

In the experimental study of ubiquitin, of the 20 ϕ-values calculated, 8 have values greater than 0.5, indicating that

these regions are structured in the transition state. Of these eight, six (Val5, Thr7, Leu15, Val17, Thr22 and Ile 23) probe the formation of the hydrophobic core and one probes helix formation (mutation of the solvent exposed Ala 28) [57]. The ϕ-values for the hydrophobic core residues, in the N-terminal region of the protein, range between 0.5 and 0.7 and the ϕ-value probing helix formation is 1. The behavior of these residues in the TSE, generated from the simulations, is consistent with the experimental results. As stated in the previous section, residues 5 and 7 are nearly as buried in the TSE as they are in the native state, and they have ϕ-values of 0.5 and 0.7, respectively. Residues 15, 17, 22 and 23 are not consistently buried throughout the ensemble but rather fluctuate from being buried to being exposed. Using both experimental and computational evidence, we can say that residues 5 and 7 are important residues for the folding of ubiquitin and that partial formation of the α-helix is also important in the rate-determining step. In the TSE generated by simulation, the partially formed helix does not interact with β-1 and β-2 to a great extent. However, in three out of the five simulated transition states, Ile 23, which resides at the N-terminal end of the helix, is interacting with the hydrophobic core. This is in general agreement with the experimental data. In Fig. 10, the folding pathway is shown.

In order to further analyze the structure of the simulated transition state, ϕ-values were calculated for the putative transition state using the method of Li and Daggett [8]. The simulated transition states were used to calculate ϕ-values and to compare experimental data. This is summarized in Fig. 9A. In general, the simulated ϕ-values are in good agreement with the experimental values; high ϕ-values are expected in the C-terminus and low values are expected in the N-terminus. However, simulations show some high ϕ-values in the N-terminus, residues 67 and 69, which show ϕ-values of near zero experimentally. The lack of correlation between the experimental and computational ϕ-values for residues 67 and 69 is most likely because the N-terminal region of ubiquitin is disordered. Alonso and Daggett [25] have removed N-terminal residues in their simulations of ubiquitin. Residue 21 also shows a ϕ-value of 1 experimentally, but in silico, it has a low ϕ-value. This is due to the fact that the experimental mutation (Asp to Ala) probes hydrogen bond formation and the technique of Li and Daggett does not account for this. The correlation between experimental ϕ-values and simulated ϕ-values is shown in Fig. 9B and C. In Fig. 9B, one can see there is very little correlation ($R = 0.09$); however, when we remove the ϕ-values for residues 21, 67 and 69, the correlation increases dramatically ($R = 0.81$). The simulated transition states are in good agreement with the experimental transition state and therefore provide a validation of the computational methods used in this study.

Ubiquitin folds via the initial population of a collapsed intermediate state that has very little but some fluctuating

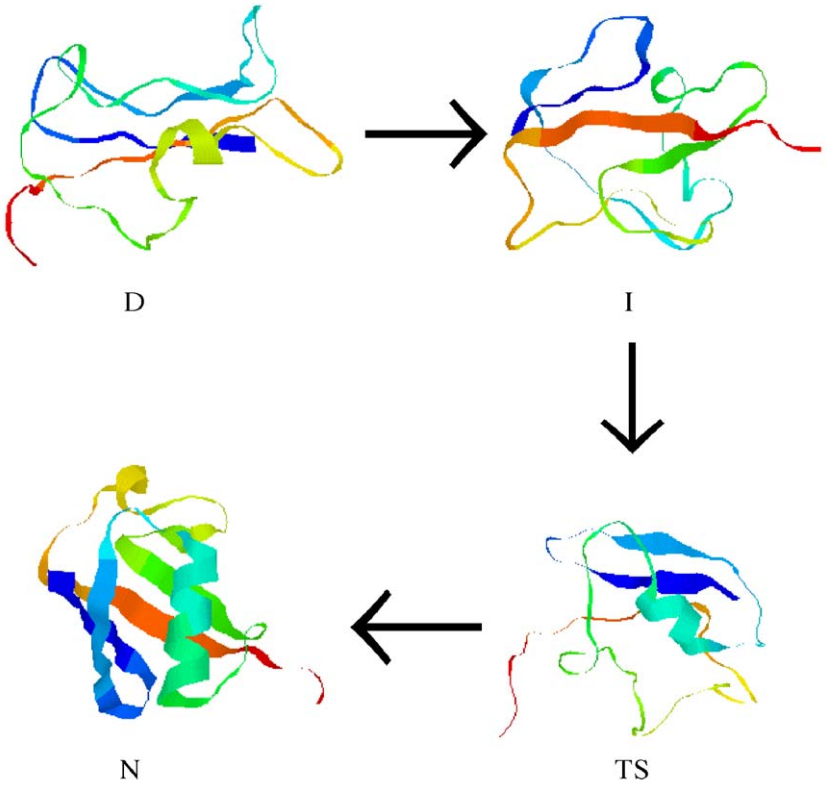

Fig. 10. The folding pathway of ubiquitin starting from the denatured state (D) going to the intermediate (I), to the transition state (TS) and finally to the native state (N).

secondary structure. Secondary structure then becomes consolidated in the transition state. The secondary structure continues to grow to the native structure. After the secondary structure formation, there is a final, small rearrangement of the hydrophobic core to give the native state.

The TSE for ubiquitin mapped by both experiment and simulation shows that ubiquitin does not fold by a strict nucleation–condensation mechanism [31]. CI2, which folds by a nucleation–condensation mechanism, has a gradient of ϕ-values which indicates that its folding nucleus is stabilized by long-range interactions [32,33]. The TSE for CI2 is, therefore, an expanded form of the native state with some fraying of the β structure. The ϕ-values for ubiquitin measured outside the first 30 residues are either zero or close to zero, indicating this region of the protein is largely unstructured in the TSE. The TSE for ubiquitin has some native-like structure but it cannot be considered an expanded form of the native state (see Fig. 6).

The TSE for ubiquitin is more closely related to that of barnase. This protein exhibits high ϕ-values in certain regions, with ϕ-values of zero outside of these regions [34]. This is similar to the pattern of ϕ-values observed for ubiquitin. This pattern of ϕ-values is reasoned by Fersht and Daggett [35], to allow for a noncooperative folding mechanism. Barnase folds noncooperatively by parts, in a hierarchical fashion [30] rather than in one concerted step, which is indicative of the two-state mechanism observed for CI2. This pattern of ϕ-values would indicate multistate kinetics for ubiquitin.

The denatured state ensemble for ubiquitin shows flickering native and nonnative regions of secondary structure. The calculation of z-values per residue indicates that a large fraction of residues is buried in the denatured state. This contributes to the low R_g of the denatured state. The expected R_g for a 76-residue random coil is about 30 Å [36]. It has been shown that the stability and compactness of the denatured state ensemble for ribonuclease Sa can be influenced by electrostatic interactions [37]. However, for ubiquitin, there are no charged residues buried in the denatured state. This seems to indicate that the compactness of the denatured state is due to the clustering of hydrophobic residues. This is currently under investigation experimentally using NMR spectroscopy.

The denatured state of CI2 has been studied both by NMR and simulation [38]. In contrast to the simulated denatured state of ubiquitin, CI2 has an expanded denatured state, with very little residual structure. The simulations show a 40% increase in the R_g for CI2 upon unfolding, compared with only a very slight increase for ubiquitin. The denatured state of CI2 is close to a random coil, whereas that of ubiquitin is more aptly described as a relatively compact globule stabilized by hydrophobic interactions.

The denatured state of barnase has also been studied using both experiment and theory [39]. These studies suggest that it has a compact denatured state with flickering regions of native structure and a region of nonnative helical structure. This is quite similar to the simulated denatured state of ubiquitin characterized here. The similarity between the denatured states of both proteins may give us some insight into the folding kinetics of ubiquitin because barnase also folds with multistate kinetics. Residual structure in the denatured state maybe a prerequisite for the observation of three-state folding kinetics. To investigate this correlation further, and see whether this is a general phenomenon, the properties of the denatured state of proteins and their kinetic models were studied (see Table 1). This strongly suggests that proteins, which have significant residual structure in their denatured states, will populate intermediate states on their folding pathways.

Table 1
List of proteins whose folding kinetics and residual structure have been determined

Protein	Residual structure in denatured state	Kinetics
CI2	no [38]	two-state [40]
Barnase	yes [39]	multistate [34]
Staph nuclease	yes [41]	multistate [42]
WW domains	no [43]	two-state [43]
Lysozyme	yes [44]	multistate [45]
Barstar	yes [46]	multistate [47]
Fibronectin type III	no [48]	two-state [49]
Protein G	yes [50]	multistate [51]

5. Conclusions

In this work, we have shown that by monitoring the simultaneous evolution of distinct structural probes over time, much information can be gained about the folding process. By plotting these distinct parameters against each other, it is possible to gain insight into folding kinetics. This study has shown that the folding of ubiquitin is a multiphase process, consisting of an initial collapse of the polypeptide chain and of a fine structural rearrangement to the native state. The initial phase most likely occurs within the dead time of conventional stopped-flow instruments and therefore cannot be measured experimentally.

The folding pathway of this protein was probed further by identifying the rate-determining TSE. The TSE was identified using essential dynamics which looks at the important or essential motions of proteins in phase space. The simulated TSE is in good agreement with the experimental studies and comparisons with the TSE for barnase; and CI2 have given us insights into the folding mechanism for ubiquitin. The simulated denatured state for ubiquitin was also analyzed and compared with other proteins. This analysis seems to indicate that ubiquitin folds with multistate kinetics in a noncooperative manner. There is also a strong correlation which indicates that proteins with residual structure in the denatured state fold with multistate kinetics. The data presented in this paper gives ample evidence that

the description of the ubiquitin folding pathway using a two-state model is inadequate.

Acknowledgements

The authors would like to thank the Welton Foundation for funding during the course of this work. We would also like to acknowledge The Unilever Centre for Molecular Informatics, Cambridge, for computational time. We would like to thank Kresten Lindorff-Larsen for help in calculating ϕ_{sim}-values. We are grateful to Dr. Charlotte Bolton and Dr. David Wales for help in setting up GROMACS. NJM would like to thank Andrew Brown and Dr. Robert Best for useful discussions.

References

[1] J.E. Shea, C.L. Brooks III, From folding theories to folding proteins: a review and assessment of simulation studies of protein folding and unfolding, Annu. Rev. Phys. Chem. 52 (2001) 499–535.
[2] L. Mirny, E. Shakhnovich, Protein folding theory: from lattice to all-atom models, Annu. Rev. Biophys. Biomol. Struct. 30 (2001) 361–396.
[3] K.F. Lau, K.A. Dill, A lattice statistical mechanical model of the conformational and sequence spaces of proteins, Macromolecules 22 (1989) 3986–3997.
[4] B. Derrida, Random-energy model: an exactly solvable model of disordered systems, Phys. Rev. B24 (1981) 2613–2624.
[5] J.D. Bryngelson, P.G. Wolynes, Spin glasses and the statistical mechanics of protein folding, Proc. Natl. Acad. Sci. U. S. A. 84 (21).
[6] E.I. Shakhnovich, A.M. Gutin, Formation of unique structure in polypeptide chains. Theoretical investigation with the aid of a replica approach, Biophys. Chemist. 34 (3) (1989) 187–199.
[7] D. Thirumalai, D.K. Klimov, Deciphering the timescales and mechanisms of protein folding using minimal off-lattice models, Curr. Opin. Struck. Biol. 9 (2) (1999) 197–207.
[8] A. Li, V. Daggett, Characterization of the transition state of protein unfolding by use of molecular dynamics: chymotrypsin inhibitor 2, Proc. Natl. Acad. Sci. U. S. A. 91 (22) (1994) 10430–10434.
[9] A. Li, V. Daggett, Molecular dynamics simulation of the unfolding of barnase: characterization of the major intermediate, J. Mol. Biol. 275 (4) (1998) 677–694.
[10] K.F. Fulton, E.R. Main, V. Daggett, S.E. Jackson, Mapping the interactions present in the transition state for unfolding/folding of FKBP12, J. Mol. Biol. 291 (2) (1999) 445–461.
[11] A. Li, V. Daggett, Identification and characterization of the unfolding transition state of chymotrypsin inhibitor 2 by molecular dynamics simulations, J. Mol. Biol. 257 (2) (1996) 412–429.
[12] Y. Pan, V. Daggett, Direct comparison of experimental and calculated folding free energies for hydrophobic deletion mutants of chymotrypsin inhibitor 2: free energy perturbation calculations using transition and denatured states from molecular dynamics simulations of unfolding, Biochemistry 40 (9) (2001) 2723–2731.
[13] A.G. Ladurner, L.S. Itzhaki, V. Daggett, A.R. Fersht, Synergy between simulation and experiment in describing the energy landscape of protein folding, Proc. Natl. Acad. Sci. U. S. A. 95 (15) (1998) 8473–8478.
[14] S.L. Kazmirski, A. Li, V. Dagget, Analysis methods for comparison of multiple molecular dynamics trajectories: applications to protein unfolding pathways and denatured ensembles, J. Mol. Biol. 290 (1) (1999) 283–304.
[15] V.S. Pande, D.S. Rokhsar, Molecular dynamics simulations of unfolding and refolding of a beta-hairpin fragment of protein G, Proc. Natl. Acad. Sci. U. S. A. 96 (16) (1999) 9062–9067.
[16] R. Du, V.S. Pande, A.Y. Grosber, T. Tanaka, E.S. Shakhnovich, On the reaction coordinate for protein folding, J. Chem. Phys. 108 (1998) 334–350.
[17] J. Williams, A. Rohrbacher, J. Seong, N. Marianayagam, K.C. Janda, R. Burcl, M.M. Szczesniak, G. Chalasinski, S.M. Cybulski, N. Halberstadt, A three dimensional potential energy surface for He+Cl2: ab initio calculations and a multiproperty fin, J. Chem. Phys. 111 (1999) 997–1007.
[18] A. Rohrbacher, N. Halberstadt, K.C. Janda, The dynamics of noble gas di-halogen molecules and clusters, Annu. Rev. Phys. Chem. 51 (2001) 405–433.
[19] J.N. Onuchic, Z. Luthey-Schulten, P.G. Wolynes, Theory of protein folding: the energy landscape perspective, Annu. Rev. Phys. Chem. 48 (1997) 545–600.
[20] J.N. Onuchic, P.G. Wolynes, Z. Luthey-Schulten, Toward an outline of the topography of a realistic protein-folding funnel, Proc. Natl. Acad. Sci. U. S. A. 92 (8) (1995) 3626–3630.
[21] S. Vijay-Kumar, C.E. Bugg, W.J. Cook, Structure of ubiquitin refined at 1.8 Å, J. Mol. Biol. 194 (1987) 531–544.
[22] S. Khorasanizadeh, I.D. Peters, H. Roder, Evidence for a three-state model of protein folding from kinetic analysis of ubiquitin variants with altered core residues, Nat. Struct. Biol. 3 (2) (1996) 193–205.
[23] B.A. Krantz, T.R. Sosnick, Distinguishing between two-state and three-state models for ubiquitin folding, Biochemistry 39 (38) (2000) 11696–11701.
[24] A. Amadei, A.B. Linssen, H.J. Berendsen, Essential dynamics of proteins, Proteins 17 (1993) 412–425.
[25] V.O. Alonso, V. Daggett, Molecular dynamics simulations of the hydrophobic collapse of ubiquitin, Protein Sci. 7 (1998) 860–874.
[26] T. Lazaridis, M. Karplus, "New view" of protein folding reconciled with the old through multiple unfolding simulations, Science 278 (1997) 1928–1931.
[27] A. Colubri, A. Fernandez, Pathway diversity and concertedness in protein folding: an ab-initio approach, J. Biomol. Struct. Dyn. 19 (2002) 739–764.
[28] J.M. Sorenson, T. Head-Gordon, Toward minimalist models of larger proteins: a ubiquitin-like protein, Proteins 46 (2002) 368–379.
[29] J. Shimada, E.I. Shakhnovich, The ensemble folding kinetics of protein G from an all-atom Monte Carlo simulation, Proc. Natl. Acad. Sci. U. S. A. 99 (17) (2002) 11175–11180.
[30] P.S. Kim, R.L. Baldwin, Intermediates in the folding reactions of small proteins, Annu. Rev. Biochem. 59 (1990) 631–660.
[31] A.R. Fersht, Optimization of rates of protein folding: the nucleation–condensation mechanism and its implications, Proc. Natl. Acad. Sci. U. S. A. 92 (1995) 10869–10873.
[32] S.E. Jackson, N. elMasry, A.R. Fersht, Structure of the hydrophobic core in the transition state for folding of chymotrypsin inhibitor 2: a critical test of the protein engineering method of analysis, Biochemistry 32 (1993) 11270–11278.
[33] L.S. Itzhaki, D.E. Otzen, A.R. Fersht, The structure of the transition state for folding of chymotrypsin inhibitor 2 analysed by protein engineering methods: evidence for a nucleation-condensation mechanism for protein folding, J. Mol. Biol. 254 (1995) 260–288.
[34] L. Serrano, A. Matouschek, A.R. Fersht, The folding of an enzyme: VI. The folding pathway of barnase: comparison with theoretical models, J. Mol. Biol. 224 (1992) 847–859.
[35] A.R. Fersht, V. Daggett, Protein folding and unfolding at atomic resolution, Cell 108 (2002) 573–582.
[36] T.E. Creighton, Proteins: Structure and Molecular Properties, Freeman, New York, 1984.
[37] C.N. Pace, R.W. Alston, K.L. Shaw, Charge-charge interactions influence the denatured state ensemble and contribute to protein stability, Protein Sci. 9 (2000) 1395–1398.
[38] S.L. Kazmirski, K.B. Wong, S.M. Freund, Y.J. Tan, A.R. Fersht, V.

Braun, Protein folding from a highly disordered denatured state: the folding pathway of chymotrypsin inhibitor 2 at atomic resolution, Proc. Natl. Acad. Sci. U. S. A. 98 (8) (2001) 4349–4354.
[39] K.B. Wong, J. Clarke, C.J. Bond, J.L. Neira, S.M. Freund, A.R. Fersht, V. Daggett, Towards a complete description of the structural and dynamic properties of the denatured state of barnase and the role of residual structure in folding, J. Mol. Biol. 296 (2000) 1257–1282.
[40] S.E. Jackson, A.R. Fersht, Folding of chymotrypsin inhibitor 2: 1. Evidence for a two-state transition, Biochemistry 30 (1991) 10428–10435.
[41] W.F. Walkenhorst, S.M. Green, H. Roder, Kinetic evidence for folding and unfolding intermediates in staphylococcal nuclease, Biochemistry 36 (19) (1997) 5795–5805.
[42] D. Shortle, M.S. Ackerman, Persistence of native-like topology in a denatured protein in 8 M urea, Science 293 (2001) 487–489.
[43] N. Ferguson, C.M. Johnson, M. Macias, H. Oschkinat, A. Fersht, Ultrafast folding of WW domains without structured aromatic clusters in the denatured state, Proc. Natl. Acad. Sci. U. S. A. 98 (2001) 13002–13007.
[44] J. Klein-Seetharaman, M. Oikawa, S.B. Grimshaw, J. Wirmer, E. Duchardt, T. Ueda, T. Imoto, L.J. Smith, C.M. Dobson, H. Schwalbe, Long-range interactions within a non-native protein, Science 295 (5560) (2002) 1719–1722.
[45] S.E. Radford, C.M. Dobson, P.A. Evans, The folding of hen lysozyme involves partially structured intermediates and multiple pathways, Nature 358 (1992) 302–307.
[46] B. Nolting, R. Golbik, A.S. Soler-Gonzalez, A.R. Fersht, Circular dichroism of denatured barstar suggests residual structure, Biochemistry 36 (1997) 9899–9905.
[47] A.K. Bhuyan, J.B. Udgaonkar, Observation of multistate kinetics during the slow folding and unfolding of barstar, Biochemistry 38 (1999) 9158–9168.
[48] A.E. Meekhof, S.M. Freund, Probing residual structure and backbone dynamics on the milli- to picosecond timescale in a urea-denatured fibronectin type III domain, J. Mol. Biol. 286 (1999) 579–592.
[49] J. Clarke, E. Cota, S.B. Fowler, S.J. Hamill, Folding studies of immunoglobulin-like beta-sandwich proteins suggest that they share a common folding pathway, Struct. Fold. Des. 7 (1999) 1145–1153.
[50] N. Sari, P. Alexander, P.N. Bryan, J. Orban, Structure and dynamics of an acid-denatured protein G mutant, Biochemistry 39 (2000) 965–977.
[51] S.H. Park, M.C. Shastry, H. Roder, Folding dynamics of the B1 domain of protein G explored by ultrarapid mixing, Nat. Struct. Biol. 6 (1999) 943–947.
[52] W.F. van Gunstren, S.R. Billeter, A.A. Eising, P.H. Humberger, P. Kruger, A.E. Mark, W.R.P. Scott, I.G. Tironi, Biomolecular Simulation: the GROMOS96 Manual and User Guide, Hochschulverlag AG an der ETH, Zurich, 1996.
[53] H.J.C. Berendsen, D. van der Spoel, R. van Drunen, GROMACS: a message-passing parallel molecular dynamics implementation, Comp. Phys. Comm. 91 (1995) 43–56.
[54] H.J.C. Berendsen, J.D.M. Postman, W.F. van Gunstren, J. Hermans, Intermolecular Forces. Interaction models for water in relation to protein hydration, Reidel, Dordrecht, 1981, pp. 331–342.
[55] R. Fraczkiewicz, W. Braun, Exact and efficient analytical calculation of the accessible surface areas and their gradients for macromolecules, J. Comp. Chem. 19 (1998) 319–333.
[56] H.M. Went, C.G. Carodza-Benitez, S.E. Jackson, FEBS Lett. (in press).
[57] H.M. Went, S.E. Jackson, unpublished results.