Available online at www.sciencedirect.com
ScienceDirect
Journal of Structural Biology 157 (2007) 557–569
Journal of Structural Biology
www.elsevier.com/locate/yjsbi

Fluctuations of primary ubiquitin folding intermediates in a force clamp

Frauke Gräter, Helmut Grubmüller *

Theoretical and Computational Biophysics Department, Max-Planck-Institute for Biophysical Chemistry, Am Fassberg 11, 37077 Göttingen, Germany

Received 19 September 2006; received in revised form 19 October 2006; accepted 1 November 2006
Available online 26 November 2006

Abstract

Folding experiments of single ubiquitin molecules under force clamp using an atomic force microscope revealed a dynamic long-lived intermediate with nanometer scale end-to-end distance fluctuations along an unexpectedly complex folding pathway. To examine the nature of this intermediate at the atomic level as well as the driving forces that give rise to the observed fluctuations, we performed molecular dynamics refolding simulations of unfolded ubiquitin under constant force. After an initial fast collapse, we find a highly dynamic, broad ensemble of conformations with partial and continuously changing secondary structure and side chain interactions. This ensemble resembles a molten-globule-like state, similar in nature to the previously described non-native state of ubiquitin in solution, but stretched by the external force. The scale of the end-to-end distance fluctuations derived from the simulations compares well with experiment. Transient formation of unspecific and metastable hydrophobic clusters along the chain are found to give rise to the observed end-to-end distance fluctuations. These distinct collapses, interpreted as folding attempts, imply an upper limit for the folding attempt frequency of ~10 ns. Our results suggest possible relations between force-induced unfolding and temperature or chemically induced denaturation.

Keywords: Ubiquitin; Force probe simulation; Atomic force microscopy; Molten globule; Refolding; Unfolding; Force clamp

1. Introduction

Proteins are complex machines, the biological function of which is determined by their three-dimensional structure, the protein fold. Understanding the basic mechanism and associated driving forces of protein folding remains a major task in biology and has been addressed by a broad variety of experimental techniques and theoretical approaches, most successfully at the molecular level of single proteins (Schuler, 2005; Zhuang and Rief, 2003; Samori et al., 2005; Michalet et al., 2006; Shakhnovich, 2006; Snow et al., 2005). A primary goal is to assess the role and nature of the unfolded and intermediate states along the complex folding pathways. Local interactions formed in the early stages of folding in transient intermediates have been suggested to narrow the conformational space of the unfolded protein towards the native state (Onuchic and Wolynes, 2004). This scenario raises a number of questions, e.g., are primarily hydrophobic or secondary structure interactions involved, and to which extent are early interactions of non-native nature?

Mechanical unfolding of titin immunoglobulin domains and other proteins using an atomic force microscope (AFM) in conjunction with molecular dynamics simulations revealed unfolding intermediates and mechanically stable protein motifs (Rief et al., 1997; Marszalek et al., 1999). Along these lines, new insights into protein folding have been recently obtained from pioneering experiments of Fernandez and Li (2004). Using an atomic force microscope (AFM), they were able to manipulate single protein molecules by means of force to probe the energy landscape underlying protein folding. The natively folded ubiquitin (1 in Fig. 1), was mechanically unfolded to an extended chain (2) under a constant force. Relaxation to a low constant force allowed to monitor the protein folding dynamics in terms of the protein end-to-end distance. The unfolded protein (2) showed a fast collapse to a long-lived intermediate (3), eventually reforming the native state (4).

* Corresponding author. Fax: +49 551 201 3202.
E-mail address: hgrubmu@gwdg.de (H. Grubmüller).

1047-8477/$ - see front matter
doi:10.1016/j.jsb.2006.11.005

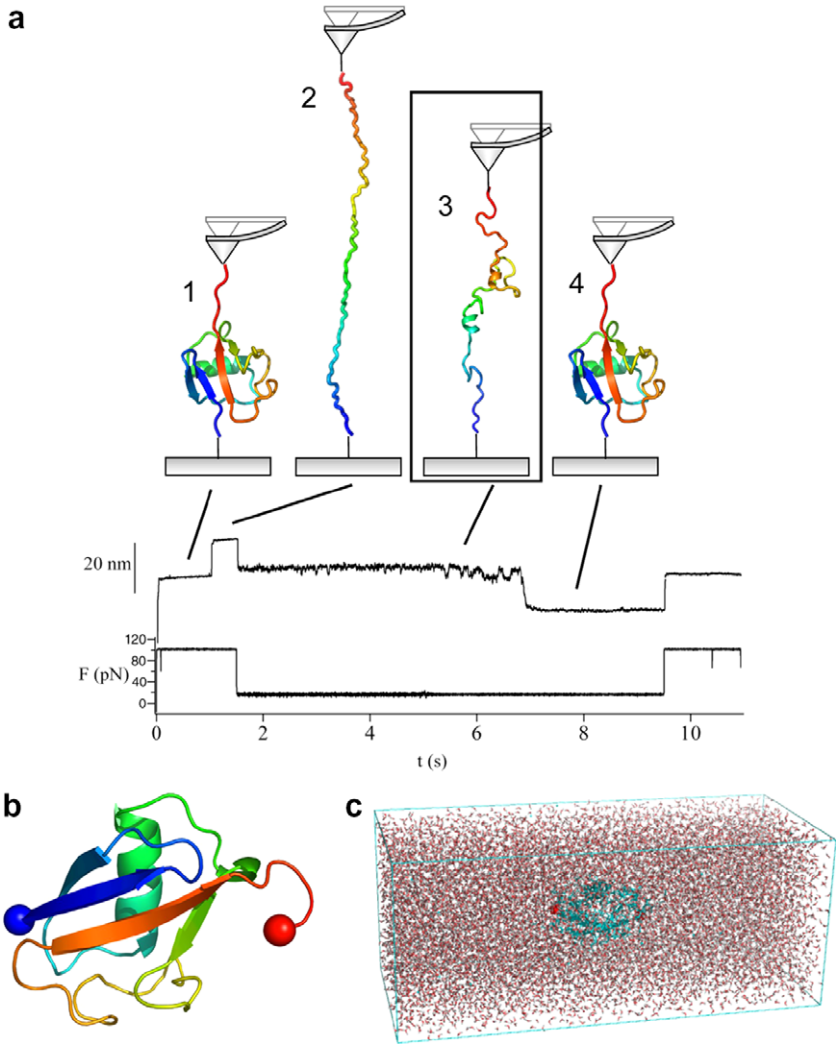

Fig. 1. Ubiquitin unfolding and refolding under constant force in an atomic-force microscope. (a) A large force triggering ubiquitin unfolding is after unfolding reduced in a single step to a lower constant force to monitor refolding (lower panel). The peptide end-to-end distance is monitored in the AFM to follow the protein (un)folding dynamics (upper panel) (1) stretched folded ubiquitin, (2) stretched unfolded ubiquitin, (3) relaxed unfolded ubiquitin, (4) refolded ubiquitin. Adopted from (Fernandez and Li, 2004). While AFM experiments were carried out for polymeric ubiquitin construct, they include trajectories showing unfolding and folding events of single domains as the one shown here. (b) X-ray structure of ubiquitin (pdb-entry 1UBQ). Color coding from blue (N-terminus) to red (C-terminus). C_α-atoms of the terminal residues (spheres) have been subjected to a constant force and are shown as spheres. (c) Simulation system for force-clamp molecular dynamics simulations of ubiquitin.

Interestingly, the intermediate featured discrete nanometer-scale fluctuations in the end-to-end distance, which intensified during the progress of folding close to the final folding event. The microscopic nature of these large fluctuations is unknown. Strictly speaking, a poly-ubiquitin construct was used; however, the features described above were also observed in events involving unfolding and folding of only a single ubiquitin domain (Fernandez and Li, 2004). We therefore refer the above description to the folding of single domains, and show such a measurement and a corresponding scheme in Fig. 1a.

Accordingly, and in the light of the simple two-state behavior found for ubiquitin in bulk measurements (Krantz et al., 2000; Krantz and Sosnick, 2000; Sosnick et al., 2004), the complex mechanical folding pathway involving several states is unexpected, and highly controversial. It has been argued that the difference in the reaction coordinate between the end-to-end distance and hydrogen exchange or fluorescence quenching gives rise to different incomparable mechanisms. Another possible explanation is that the intermediate state corresponds to an early collapsed unfolded state the formation of which has remained undetected so far (Sosnick, 2004). Reconciling the apparent discrepancy between bulk- and single-molecule experiments requires a molecular characterization of the chain of events operating in protein folding under force. The underlying molecular process, however, is inaccessible to experiments. In particular, the nature of the long-lived collapsed state and its large-scale fluctuations in chain length remains elusive. First, does the intermediate represent a purely entropically coiled state, a molten globule, or a partially structured folding nucleus? Second, might this

intermediate similarly occur in solution folding studies, or is it substantially different from any states involved in folding in solution due to the different reaction coordinate?

Attempts were made using simplified topology-based Gō-like models (Gō and Abe, 1981) to clarify the folding mechanism under force and the source of the end-to-end distance fluctuations during the long lived intermediate states (Li et al., 2006; Best and Hummer, 2005; Cieplak and Szymczak, 2006). However, previous experimental and theoretical studies emphasized the contributions of non-native contacts (Nishimura et al., 2006; Gsponer et al., 2006; Sosnick et al., 1994) and solvation effects (Rank and Baker, 1997; Cheung et al., 2002; Liu and Chan, 2005; Rodriguez-Larrea et al., 2006) to the roughness of the folding energy landscape, which is not captured by these attempts. Rather, the delicate balance between different enthalpic and entropic contributions along the protein folding pathway calls for an atomistic description of the system.

This study therefore aimed at elucidating the molecular structure and dynamical fluctuations of the intermediate state at the atomic level. To this end, we described the experiments of ubiquitin refolding in the AFM by means of microsecond force-clamp molecular dynamics (FCMD) simulations in explicit water. We restricted ourselves to the intermediate state since monitoring the final folding event itself requires time-scales of seconds and thus is computationally unfeasible. We focused on a single ubiquitin domain, and thereby exclude contributions from inter-domain interaction to the fluctuations, in accordance to the experimental observation of the highly fluctuating intermediate state also for single folding events of ubiquitin (Fernandez and Li, 2004).

Beyond the virtue of direct comparability to force-clamp experiments, studying the collapsed non-native state under external force has the advantage of a well-defined reaction coordinate restricting the non-native state to extended conformations. The reduced compactness of the non-native state in the force clamp should facilitate transitions between conformations, making them accessible during our MD simulation in explicit solvent at room temperature at a microsecond time-scale. In the absence of external force, a millisecond time-scale has been suggested by hydrogen exchange and NMR measurements for the exploration of dihedral space and other conformational transitions, that are hindered in a molten globule due to its relative compactness (Baum et al., 1989; Canet et al., 2002). As a consequence, to date non-native states could be only addressed by simulations using simplified models, additional perturbation potentials, elevated temperatures, or small peptides (Smith, 2004). To our knowledge, our study represents the first large-scale atomic-detail MD simulation of a non-native state of a globular protein without the usage of artificial perturbations or model simplifications.

After relaxation of the partially unfolded ubiquitin molecule to a low constant force, we find collapse to a stretched molten-globule-like state, a large heterogeneous ensemble with a broad distribution of secondary structure and side chain packing interactions. Formation and rupture of—mostly non-native—hydrophobic contacts are seen to give rise to nanometer fluctuations of the end-to-end distance. No specific interactions suggesting a distinct folding nucleus are seen. The heterogeneous nature of contact formation events are interpreted as folding attempts in a rough energy well, and an attempt frequency for folding under force is derived. The high similarity of the stretched molten-globule ensemble with the unfolded collapsed state of ubiquitin in solution—as previously described (Hoerner et al., 2005; Shi et al., 2006; Chung et al., 2005; Crespo et al., 2006)—suggests that the mechanical folding pathway and the folding pathway in solution might share common unfolded or intermediate states.

2. Methods

2.1. MD simulations

All simulations were carried out using the MD software package GROMACS 3.1.4 (Lindahl et al., 2001) and were run in the NpT ensemble. The temperature was kept constant at $T = 300$ K by coupling to a Berendsen thermostat with a coupling time of $\tau_T = 0.1$ ps (Berendsen et al., 1984). The pressure was kept constant at 1 bar by coupling to a Berendsen barostat with $\tau_p = 1.0$ ps and a compressibility of 4.5×10^{-5} bar^{-1} (Berendsen et al., 1984). All bonds were constrained using the LINCS algorithm (Hess et al., 1997). An integration step of 2 fs was used. Non-bonded interactions were calculated using a cutoff of 10 Å. Long-range electrostatic interactions were calculated by Particle-Mesh Ewald summation (Darden et al., 1993) with a grid spacing of 0.12 nm and cubic interpolation.

The protonation state of the only titratable group in ubiquitin, His68, was determined by calculating its pK_a value for 1UBQ (Vijaykumar et al., 1987) with Whatif (Vriend, 1990) and its interface to DelPhi (Nicholls et al., 1990). His68 was found to be neutral at pH 7. The GROMOS96 force field (van Gunsteren et al., 1996) for the protein and the SPC water model (Berendsen et al., 1981) were used. The protein (pdb-entry 1UBQ) (Vijaykumar et al., 1987) was solvated in 5120 SPC water (Berendsen et al., 1981) molecules in a cubic box of size $5.2 \times 5.6 \times 5.9$ nm^3. The resulting simulation system comprised ~32,000 atoms. After 1000 steps of steepest descent energy minimization, the solvent and ions were equilibrated during a 0.5 ns MD simulation with the protein heavy atoms subjected to harmonical constraints with a force constant of $k = 1000$ kJ mol^{-1} nm^{-2}. Subsequently, the system was equilibrated for 65 ns.

2.2. Force-clamp MD simulations

For the force-clamp MD (FPMC) simulations, a simulation system large enough to accommodate the unfolded protein was required. After rotating the equilibrated

system such that the termini were aligned along the z-axis of the simulation box, water was added to obtain a larger cuboid box of $5.8 \times 5.5 \times 20.0\,\mathrm{nm}^3$ in size, containing ~63,000 atoms (see Fig. 1). With 20 nm of length in pulling direction, this box can accommodate a substantially unfolded, largely stretched ubiquitin molecule (fully stretched ubiquitin has a contour length of 30 nm). The chosen box size allowed us to reach an accumulated simulation time of a microsecond for the collapsed folding intermediate under force. It restricts the simulations, however, to a not yet fully stretched ubiquitin chain (see below and Section 3).

A subsequent 500 ps MD simulation with harmonic constraints on the protein heavy atoms was carried out to equilibrate the newly added water and ions. The final structure of this run served as starting structure for the FCMD simulations.

Constant force was applied to each of the C_α-atoms of the two terminal amino acids (N, C). This was achieved by means of harmonic spring potentials with finite spring constants ($500\,\mathrm{kJ\,mol^{-1}\,nm^{-2}}$) acting on the two termini. The positions of the springs were continuously adjusted at each 2 fs time step such that the extension of the spring, i.e., the pulling force, remained constant throughout the simulation. We note that the instantaneous adjustment of the springs to yield constant force differs from the experimental AFM setup, which allows readjustment of the force not faster than within a time-scale of several 100 μs (Fernandez and Li, 2004). To minimize artificial friction forces which are absent in the AFM, both springs are moved depending on the protein dynamics in contrast to the situation in the AFM, in which one side, the cantilever position, is held fixed. Our approach has proved useful in previous force-probe MD simulations (Gräter et al., 2005).

To induce unfolding the equilibrated ubiquitin structure was subjected to a high stretching force of 500 pN. An FPMC simulation using 400 pN was also carried out, but did not yield unfolding events within a 40 ns simulation. Given the limitation of an end-to-end length of ubiquitin to less than 20 nm, unfolding was stopped at extensions of 10.4–17.3 nm of partially stretched ubiquitin. Four starting structures were extracted from the unfolded state at 16.4, 16.6, 16.8, and 17.2 ns. The resulting unfolded structures were subjected to extensive subsequent FPMC simulations at low force allowing for relaxation to the intermediate state as given in Table 1.

Table 1
Details of the FCMD simulations described in the text

No.	Starting extension (nm)	Force (pN)	Time (ns)
L.1	10.4	23	220
L.2	13.6	23	148
L.3	14.9	23	134
L.4	17.3	23	137
H.1	10.4	40	156
H.2	13.6	40	102
H.3	14.9	40	95
H.4	17.3	40	99

Eight FPMC simulations with two different constant relaxation forces F starting from the four partially unfolded ubiquitin conformations were performed, totaling ~1.1 μs of FCMD simulation time. The fluctuations in end-to-end length during the FPMC simulations was monitored and compared to the force-clamp measurements using the AFM.

All protein structures were plotted with Pymol (DeLano, 2001).

3. Results

3.1. Characterization of end-to-end distance fluctuations

In accordance with the force-clamp experiments using the AFM, we subjected the termini of the ubiquitin molecule to a constant force to induce unfolding (see Section 2). To yield unfolding on the nanosecond timescale, a force of 500 pN was employed. Unfolding was reflected in a one-step increase in the end-to-end distance starting from 5 nm for the stretched native fold (Fig. 2a, black curve). Such a two-state unfolding has been found by previous experiments and simulations to be the main mechanical unfolding route of ubiquitin (Schlierf et al., 2004; Li and Makarov, 2004; Irback et al., 2005).

We here focus on the ubiquitin refolding dynamics after relaxation to low force. Instead of starting the relaxation from a totally stretched protein, here, a series of partially unfolded conformations (crosses in Fig. 2a) were chosen as starting structures, the reason for which being twofold. First, this procedure allowed us to use a box of only 20 nm in length, which, though still sufficiently large, made several 100 ns of accumulated simulation time accessible. Second, the first step of the relaxation, the collapse towards the intermediate, is shortened. Again, this allows a more extensive sampling of the intermediate of interest, and of early events in the folding process itself, as well as, in particular, of its fluctuations. The four starting structures are shown in Fig. 2b.

Each of the four stretched ubiquitin molecules was relaxed under two different constant forces, 27 and 40 pN, respectively, covering the range of experimentally employed forces (Fernandez and Li, 2004). Thus, here, no artificial acceleration of the folding process was applied. As can be seen, the end-to-end distances during the eight independent trajectories are plotted in color in Fig. 2a. The end-to-end distance decreased abruptly in all cases. The decrease leveled off to a plateau at a medium end-to-end distance for a subset of trajectories (L.2, L.3, H.2, H.3). Upon relaxation, large end-to-end distance fluctuations in the nanometer scale (up to 3 nm) were observed. Notably, the fluctuations were most pronounced and discrete for the relaxation to an intermediate state reaching a plateau value (L.2, L.3, H.2, H.3). An average standard deviation of 0.6 nm in the chain length was obtained for this subset of four trajectories. In contrast, the folded ubiquitin molecule shows a significantly smaller standard

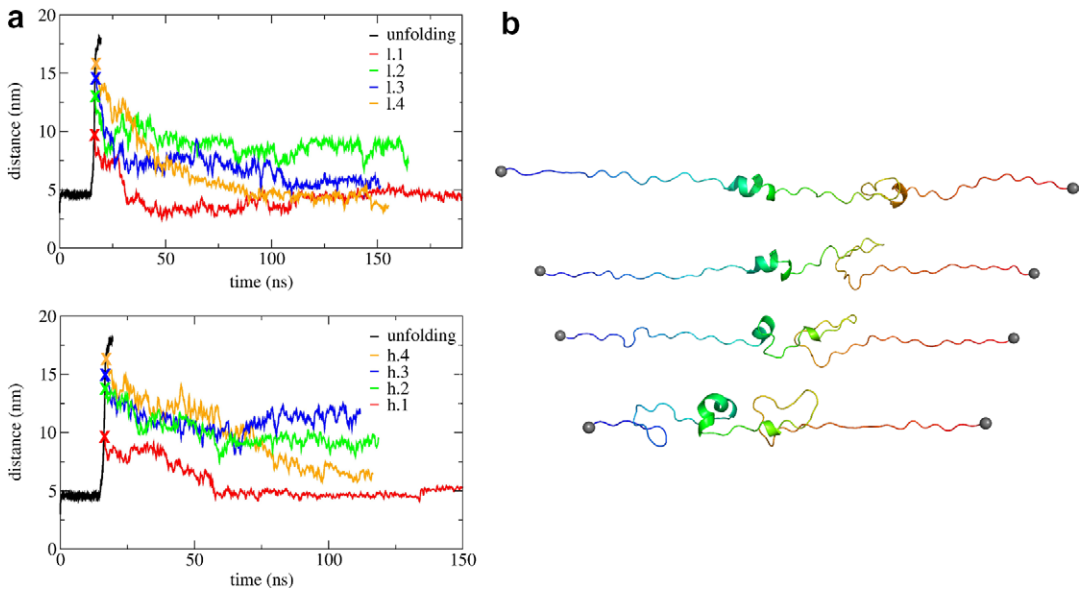

Fig. 2. End-to-end distances during unfolding and subsequent relaxation of single ubiquitin domains in force-clamp molecular dynamics simulations. (a) Applying a force 500 pN to folded ubiquitin results in an unfolding event associated with a sudden increase in length (black curve). Relaxation to a force of 40 pN (upper panel) and 66 pN (lower panel), respectively, was simulated starting from four different partially unfolded structures, end-to-end distances of the structures prior to relaxation are shown as crosses, those during relaxation as colored curves. (b) Unfolded starting structures prior to relaxation, with end-to-end distances between 10.4 and 17.3 nm (from bottom to top).

deviation of 0.06 nm under force (black curve, before unfolding). This difference in end-to-end fluctuations of the intermediate with respect to the folded states hints at a much broader energy potential well along the pulling coordinate. A significant dependency of the fluctuations on the applied force was not observed.

The extremely reduced fluctuations in the folded state seen in the simulations might be partly due to the very high force applied to induce unfolding on a short time-scale. For the sake of better comparison to experiment, we performed additional FPMC simulations subjecting the folded ubiquitin to the smaller unfolding forces (100 pN) used in experiments. We obtained a standard deviations of 0.15 nm at 100 pN (80 ns simulation time; 0.09 nm at 200 pN during 20 ns of simulation time). As expected, slightly higher standard deviations of the end-to-end distance of the folded state were found at lower stretching forces. In summary, we find a fourfold increase in the end-to-end distance fluctuations when going from the stretched folded to the relaxed unfolded state. This increase is in good agreement with the fluctuations of single unfolded ubiquitin molecules observed experimentally. Under force clamp in the AFM (Fig. 1), a single ubiquitin domain in the folded state stretched at 110 pN shows a standard deviation of 0.3 nm, and the intermediate state at 10 pN, 0.8 nm of end-to-end distance fluctuations (Jasna Brujic and Sergi Garcia-Manyes, private communication). We note that both the time-scale and the time resolution differ between our FCMD simulations and the AFM experiments, rendering a direct comparison of above numbers impractical. Nevertheless, the good agreement of the end-to-end distance trajectory, in terms of the large size and discrete nature of the fluctuations, also relative to the folded state, strongly suggests that the observed collapsed state and his conformational dynamics are sufficiently close to the mechanical folding intermediate and his fluctuations to justify further analysis.

Other starting conformations, namely the least and most unfolded ubiquitin structures (L.1, L.4, H.1, H.4), do not level off at a medium distance corresponding to an intermediate state upon relaxation. Instead, they collapse to a non-native state with an end-to-end distance close to the native state. Within the achieved simulation time of roughly 100 ns each, no subsequent distance increase or large fluctuations are seen. Apparently, the strongly collapsed non-native state is kinetically trapped in a deeper energy well comparing to the characteristic intermediate state featuring the nanometer scale fluctuations. Also, the force-clamp experiments suggest the collapse to non-native states to be a stochastic process, as a small fraction of trajectories showing a drop to the end-to-end distance of the folded state do not exhibit the fingerprint of folded ubiquitin. Having trajectories of these incidents at hand, we are now able to characterize and rationalize also these collapse events (see below).

The agreement between the fluctuations observed during the FPMD simulations with those obtained experimentally suggest, indeed, that the simulations correctly describe this aspect of the refolding dynamics of ubiquitin under force. We now address the question of the driving forces that cause the nanometer-scale length fluctuations. Fig. 3 shows the superposition of the end-to-end length and (a) the solvent-accessible surface area, or (b) the number of secondary structure contacts for one representative relaxation

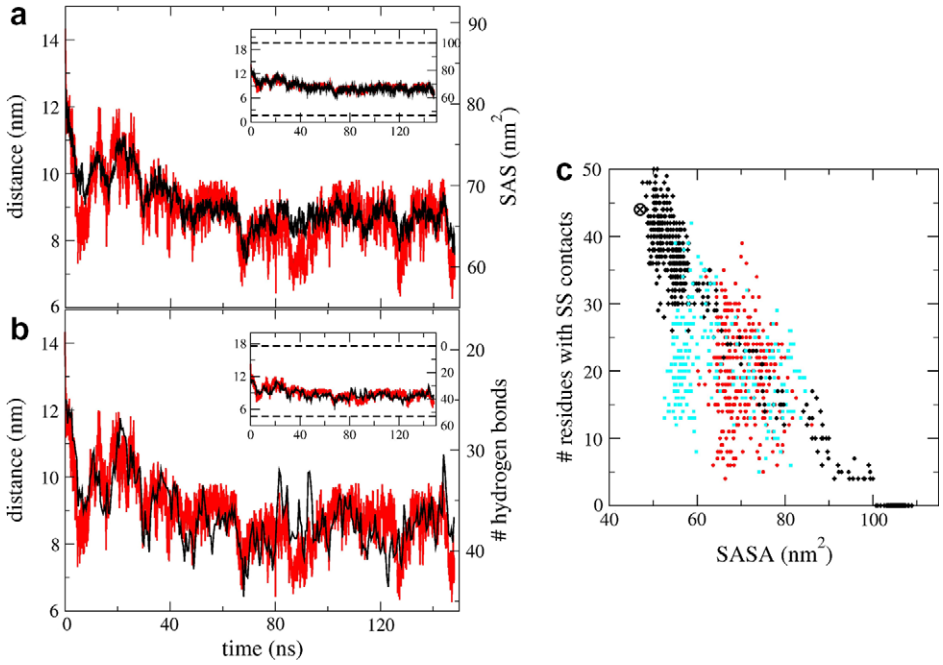

Fig. 3. Characterization of the end-to-end fluctuations. Fluctuations in (a) the solvent-accessible surface (black) and (b) the number of secondary structure contacts (black) are compared to the end-to-end distance (red) for one representative relaxation. The inset shows the same data together with the values for the native and completely stretched protein as upper and lower limits as dashed lines. (c) Number or secondary structure contacts are plotted against the solvent accessible surface for one unfolding trajectories (black), relaxation trajectories with fluctuating intermediates (L.2, L.3, H.2, H.3, red), and relaxation trajectories leading to compact non-native states (L.4, H.4, cyan).

trajectory. Both the buried surface area as well as the number of backbone hydrogen bond interactions change with the distance. This suggests both as driving forces for the end-to-end distance fluctuations of the intermediate state, with a higher correlation seen for the hydrophobic interactions (*correlation coefficients of 0.8 compared to −0.2 for hydrogen bond number*). Thus, driving forces are of both hydrophobic and electrostatic nature. As can be seen in the insets of Fig. 3a and b, in terms of both measures, the buried surface and the total backbone interactions, the intermediate is closer to the folded than to the fully extended state.

Does the first part of the folding pathway, the instantaneous collapse upon relaxation and the subsequent fluctuations in the folding intermediate, largely represent a hydrophobic collapse, or, alternatively, feature early formation of secondary structure elements? To address this question, Fig. 3c plots two measures for the progress of folding towards a compact (non-native or native) state, the solvent-accessible surface and the number of secondary structure contacts, against each other for trajectories obtained from unfolding (black), relaxation to a fluctuating intermediate state (red), and relaxation to a highly compact non-native state (cyan). Apparently, a broad range between the fully extended and the native state is covered by the different simulation runs. Hydrophobic collapse is associated with, but not tightly coupled to backbone hydrogen bond formation. Hence, the folding mechanism under force clamp is a mixture between the two extreme cases without a clear order in the sequence of contact formations. Comparing to the fluctuating intermediates (L.2, L.3, H.2, H.3), collapse to a more compact non-native state with low end-to-end distances close to the native state (L.4, H.4, cyan) leads to a further decrease in the solvent-accessible surface area. However, the strong collapse is not associated with a further increase in hydrogen bonding as in the native structure. The compact structures resulting from such far-reaching collapses show extensive clustering of hydrophobic residues in the protein core. The hydrophobic packing apparently allows less optimal secondary structure conformations regarding hydrogen bonding comparing to the native state. It must be emphasized that we here restrict ourselves to the first part of the protein folding mechanism, such that no conclusions about later folding steps from the intermediate to the natively folded state are possible.

3.2. Heterogeneous contraction–extension events

The fluctuations in end-to-end length are discrete, with separate jumps between a more extended chain and a more contracted chain. The discreteness of the fluctuations might reflect repeated transitions between two discrete well-defined states, or instead just represent a random diffusion in a broad conformational subspace. To decide between these two alternatives, we examined both the structural properties of the intermediate state, and the extent to which the fluctuations, if at all, add to the diversity.

Fig. 4 shows the results for three representative relaxation trajectories, L.2, H.2, and H.3. We dissected the ubiquitin protein chain into smaller fragments of several amino

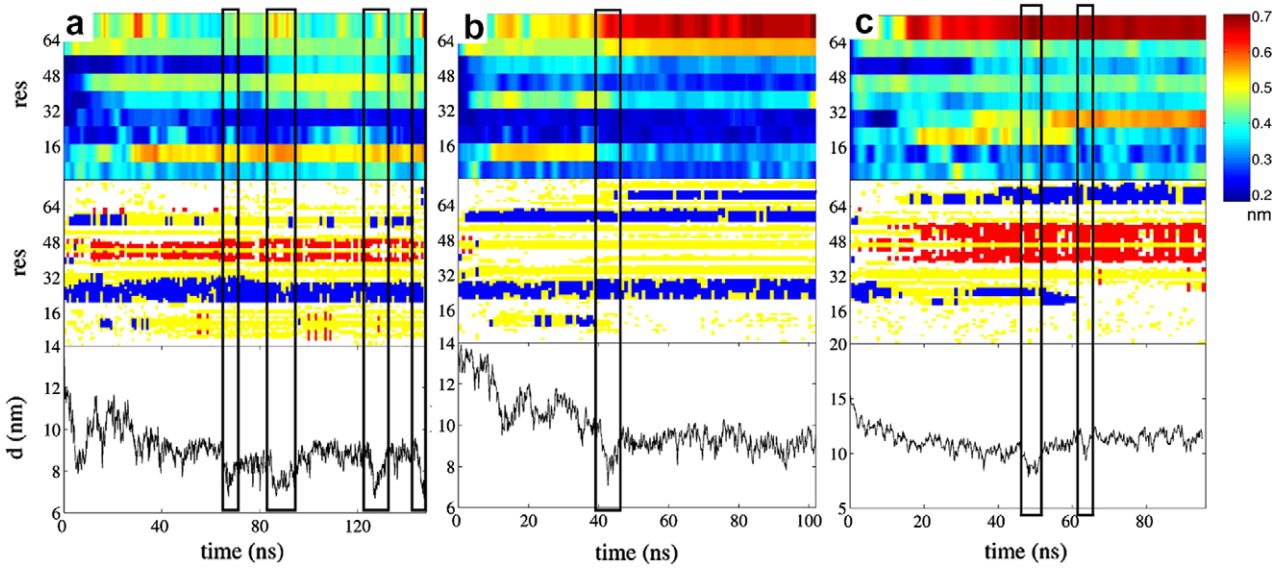

Fig. 4. Conformational changes along the peptide chain and secondary structure during relaxation of a subset of three relaxation trajectories (a, L.2; b, H.2; c, H.3). Discrete collapses of the peptide chain are boxed. Lower panel: end-to-end distances. Middle panel: secondary structure. Blue indicates β-sheets, red helical structure, white coil, bend or no structure. Upper panel: root-mean-square deviation of the backbone atoms of peptide fragments along the ubiquitin chain as indicated on the *y*-axis. Deviations were calculated after least-square fit of the segment to the respective conformation at the start of the relaxation.

acids, and monitored the conformational change of each fragment in terms of root-mean-square deviation prior to relaxation (upper panel). Strikingly, contraction and extension events are not restricted to certain fragments along the peptide chain. Instead, distances vary broadly for a given fragment, among the chain, and in between the different independent simulations. Fluctuations in the total protein length were traced back to transient contractions at very different sites along the chain, with subsequent re-opening of this site or yet another site of the chain. The heterogeneous patterns of fluctuations is associated with varying secondary structure patterns, as shown in the middle panel of Fig. 4. For instance, H.2 shows transient formation and rupture of non-native helices both near the N and C-terminus, whereas L.2 shows remainders of native helices, and additional formation of primary contacts between β-strands β1 and β2.

To further locate the peptide fragments of transient contact formations, we calculated the inter-residue distances averaged over alternating periods of extension and contraction. Fig. 5 shows the resulting contact maps for L.2 as an example. The periods of relative extension (*above diagonal*,

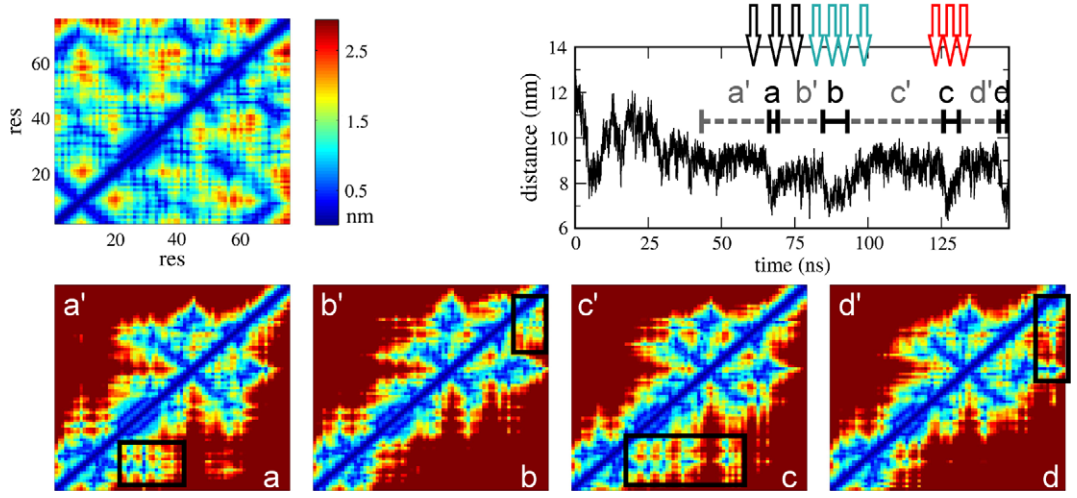

Fig. 5. Contact formation during collapse events giving rise to discrete distance fluctuations. Contact maps show the residue-residue distances for the folded conformation (top) and for the unfolded collapsed state (bottom) during a representative relaxation trajectory (L.2). They have been obtained from averaging over eight different intervals of the trajectory as indicated in the end-to-end distance plot (top, dashed line: extended state, solid line: collapsed state). Contact maps of extended and collapsed conformations alternate, and residues involved in contact formation (black boxes) vary among the collapse events.

a′–d′) and contraction (*below diagonal*, a–d) are indicated in the distance plot (inset). The upper left panel shows the contact map of the folded state for comparison. Corresponding structures before, during, and after the contraction events, indicated by arrows in the distance plot, are shown in Fig. 6. Residues involved in contact formation during the respective contraction event are depicted as spheres. First and most importantly, contact formation upon a contraction event occurs at alternating fragments along the protein chain (black boxes in contact maps). Hence, contacts are inherently transient, a characteristic of molten globules. Second, contact orders, defined as the distance of interacting residues along the peptide sequence, of up to 20 residues are occasionally reached. Large contact orders (the native state reaches 60 residues in β1–β5) result in larger chain contractions and are thus inhibited by the applied force-clamp as long as not counterbalanced by sufficiently beneficial interactions, such as for the final folding to the native state. Yet, contacts in the intermediate do, though only transiently, go beyond local interactions between a few residues very close in sequence. The nanometer scale end-to-end distance fluctuations are the immediate consequence.

Analysis of the residues located in collapsed peptide fragments (compare Fig. 6) reveals that hydrophobic amino acids predominate the clusters formed upon contraction events. Apparently, formation of small hydrophobic cores are the main driving forces for the fluctuations in the intermediate state, in line with the high correlation between the hydrophobic surface area the intermediate and the end-to-end distance (Fig. 3). Collapses also incidentally involve the formation of salt bridges, favorable, and long-range interactions, which might be a main driving force for contact formations of higher contact order. Both native (Lys11-Glu34), and non-native salt bridges (Lys29-Glu18, Lys29-Asp39) are found. Within the time-scale of the simulations, however, a statistically significant increase in native contacts was not observed. Considering the milliseconds to seconds time scale of ubiquitin folding in the force clamp, completion of the folding process towards the native structure can of course not be expected to be observed in the simulations. Overall, interactions are found to be distinct, yet transient and unspecific. The observed intermediate state does not feature a specific folding nucleus, around which conformational fluctuations occur; rather, the complete protein undergoes continuous reformation of interactions.

Beyond hydrophobic and polar contacts between side chains, what is the role of backbone hydrogen bonding and secondary structure formation in the observed intermediate state? The temporal changes of α-helical and β-sheet structures (Fig. 4) show that, again, no characteristic pattern dominates the intermediate state;instead, continuous reformation is seen. Further analysis of backbone hydrogen bonding revealed dynamics not necessarily coupled to the end-to-end distance fluctuations. One example is shown in Fig. 7 for the N-terminal β1–β2 sheet during L.2. Within the 150 ns long dynamics of the intermediate

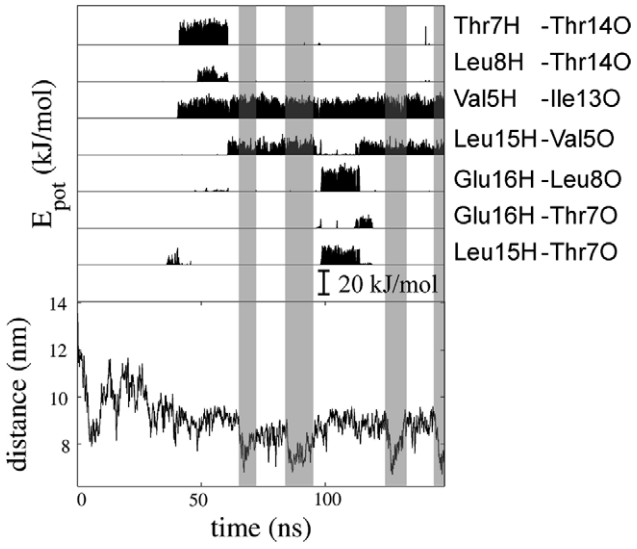

Fig. 7. Dynamics of transient backbone hydrogen bond formation during fluctuations in the end-to-end distance. Hydrogen bond potentials are shown for an N-terminal fragment of ubiquitin as observed for trajectory L.2 (top). Changes in the hydrogen bond patch are not correlated to fluctuations in the end-to-end distance (bottom, gray shades).

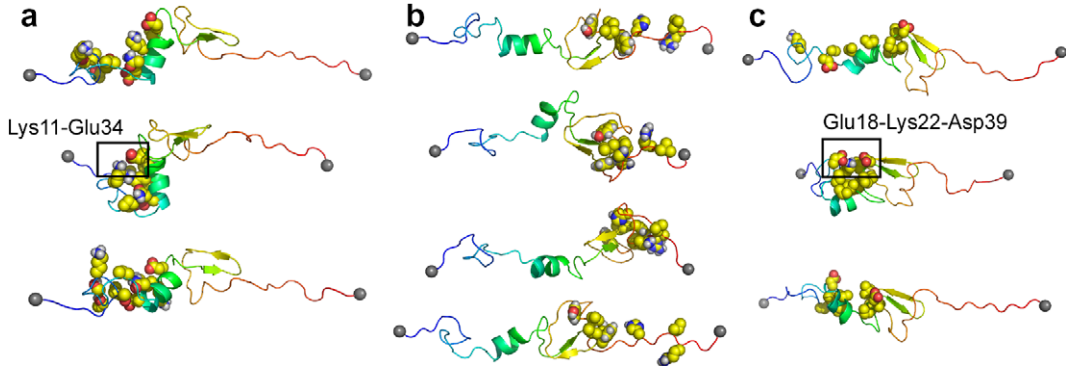

Fig. 6. Structural changes upon collapse giving rise to end-to-end distance fluctuations. Representative structures are shown before, during, and after three collapses corresponding to the contact maps a, b, and c in Fig. 5. Residues involved in contact formation (boxed in Fig. 5) are shown as spheres. Newly formed salt bridges are indicated.

state, inter-strand hydrogen bonds are formed and repeatedly reorganized. Changes in the hydrogen bonding pattern, however, do not always involve end-to-end distance changes, but instead only slight relative adjustments of the chains sliding along each other. Thus, the large-scale contraction and extension dynamics are superimposed by additional subtle backbone dynamics exploring secondary structure space, which are invisible in the AFM.

3.3. Folding attempt frequency and folding energy landscape

Our findings suggest the fluctuations in the intermediate state to reflect a search for favorable backbone conformations and side chain packing, driven by both electrostatic and hydrophobic forces. Repeated contractions of the chain represent protein attempt to form additional contacts. Insufficient stabilization upon additional contact formation leads to re-opening of contacts at this or another site under external force. These characteristics of a dynamical, stretched molten-globule suggest the following folding scenario. The stretched unfolded ubiquitin molecule recoils to a molten-globule state upon force quench. Repeated trials of the stretched molten-globule to collapse might represent motions towards the folded state without finally overcoming the energy barrier separating the molten globule from the native state. The conformational heterogeneity of the molten globule in terms of native and non-native contacts is due to motion along the well of the molten globule in dimensions orthogonal to the folding coordinate of end-to-end distances. During this conformational search along the well of the stretched molten-globule, the barrier is eventually overcome during one of the folding attempts resulting in folding.

Within this folding scenario, we calculated an attempt frequency for folding from the end-to-end distance fluctuations. Fig. 8a shows the frequency spectrum of the end-to-end distance fluctuations of the subset of four trajectories, L.2, L.3, H.2, and H.3. From the main contributions to the spectrum, we predict a folding attempt frequency k_a in the range of 0.05–0.2 ns^{-1}. We here assume any collapse event to represent a folding attempt. The frequency spectrum does not suggest any faster large-scale motion of the intermediate state along the folding coordinate to be significant in terms of distance fluctuations. On the other hand, only a part of the fluctuations taken into account might constitute folding attempts. Thus, the given frequency can only serve as an upper limit for folding attempts.

From the attempt frequency, we deduced an activation free energy barrier via Kramers' equation (Kramers, 1940; Hänggi et al., 1990). Despite the complex nature of protein free energy landscapes and assuming a single collective reaction coordinate with a single barrier and a harmonic well, remarkably good results for the kinetics of protein folding have been reported (Klimov and Thirumalai, 1997; Socci et al., 1996). A similar approach has been previously applied to estimate folding free energies for a cold-shock protein from FRET measurements (Schuler et al., 2002; Schuler, 2005) and for a helix bundle modeled with the Go-model (Best and Hummer, 2006). Here, we define the end-to-end distance as the reaction coordinate, and assume a single barrier between the molten globule (MG) and the native state (N). In the strong friction regime, the Kramers' equation reads (Hänggi et al., 1990)

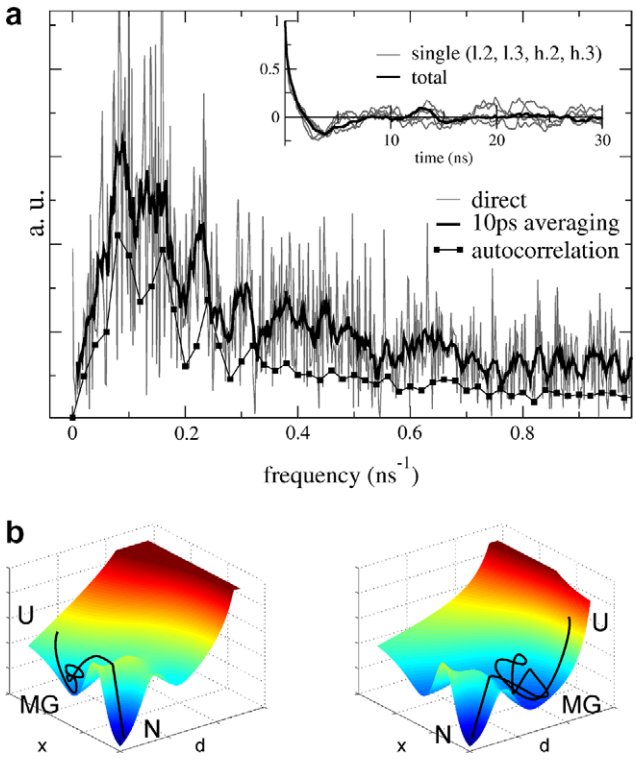

Fig. 8. Folding attempt frequency and folding energy landscape. (a) A frequency spectrum for the conformational fluctuations is obtained from a Fourier-transformation of (i) the end-to-end distance directly (gray curve), from (ii) a 10 ps average (black curve), and (iii) from the auto-correlation of the end-to-end distance fluctuations (black squares). A dominant contribution is seen for frequencies in the range of 0.05–0.2 ns^{-1}. The auto-correlation is shown in the inset. (b) Schematic folding free energy landscape in two dimensions, the reaction coordinate for folding in the force clamp, the end-to-end distance d, and a different conformational coordinate for folding in solution, x. The main features of the free energy profile along both pathways were assumed to be an unfolded molten-globule state (MG), one major barrier towards folding and a narrow native state (N). The free energy landscape in the absence of force (left) is tilted by external force along d (right). For details see text.

$$k_f = \frac{\omega_{MG}\omega_{TS}}{2\pi\gamma} e^{-\Delta G^{\neq}/k_B T}, \quad (1)$$

where ω_{MG} and ω_{TS} are the angular frequencies of the system at the well of the stretched molten globule and at the transition state barrier, respectively, both assumed to be of harmonic shape. γ is the damping rate, $\Delta G^{\neq}$ the free energy barrier, k_B the Boltzmann constant, and T the temperature. The Kramers' pre-factor $\omega_{MG}\omega_{TS}/2\pi\gamma$ can be regarded as the folding attempt frequency in the one-dimensional harmonic well of the molten globule during conformational diffusion as given by the shape of the well and the damping. Assuming ω_{MG} and ω_{TS} to be equal, the pre-factor simplifies to $\omega_{MG}^2/2\pi\gamma = k_a/2\pi$, where k_a is

the attempt frequency or rate of conformational reorganization of the molten globule (Socci et al., 1996; Schuler et al., 2002; Best and Hummer, 2006). Here, k_a as given by the distance fluctuations in the FPMC simulations serves as an upper bound for the pre-factor. With a folding rate of $k_f = 100\,s^{-1}$ from force-clamp measurements extrapolated to zero force (Fernandez and Li, 2004), we obtained an upper limit for the activation free energy of $\Delta G^{\neq}$ of 12 k_BT.

Whereas folding attempt frequencies and activation free energies are the parameters of choice to characterize folding free energy landscapes, they are not straightforward available from kinetic measurements. To date, few attempts have been made to quantify attempt frequencies, resulting in estimates that greatly differ by up to several orders of magnitude. For the heat-shock protein CspTm, a lower bound of $0.04\,\mu s^{-1}$ was recently obtained experimentally, and an upper bound of $0.02\,ns^{-1}$ from a Gaussian chain model (Schuler et al., 2002). *In agreement, calorimetry experiments gave 0.04* μs^{-1} *for the related protein CspB* (Naganathan et al., 2005). A reorganizational frequency of $2\,\mu s^{-1}$ was found from diffusive model for a small helix bundle (Best and Hummer, 2006), *close to the previously determined pre-factor of 2.5* μs^{-1} *on average based on a similar lattice model for a set of globular proteins* (Li et al., 2004). Nanosecond temperature spikes leading to protein denaturation recently suggested an upper bound of $0.1\,ns^{-1}$ (Steel et al., 2006). The discrepancy in times scales for folding attempts might be partly attributed to their quality as mere lower and upper bounds, to the different reaction coordinates used, as well as to the fact that different proteins were studied. Further studies in this direction will surely help to reconcile these discrepancies and to narrow down the estimated range of potential reconfigurational frequencies. Eventual dependencies of these parameters on size and topology of the protein might also be expected. Further efforts along these lines will surely help to refine our picture of folding free energy landscapes.

4. Discussion and conclusion

This study aimed at characterizing the structure and dynamics of the strikingly long-lived and fluctuating intermediate state of folding in the force clamp as observed in AFM measurements. Indeed, we find large-scale discrete end-to-end distance fluctuations in the experimental length scale of several nanometers, strongly suggesting that the conformational ensemble sampled in the force-clamp molecular dynamics simulations corresponds to the folding intermediate seen in the experiments.

One major result is the significant contribution of local non-native contacts to the intermediate ensemble. Specific interactions between oppositely charged side chains and distinctive hydrophobic cores are formed, however, with a distribution too broad to be fully captured by our sub-microsecond simulations. Certainly, the conformational fluctuations substantially go beyond purely entropic chain fluctuations. The eventual rupture of non-native interactions to form the folding nucleus leading to the native state is likely to represent the main folding barrier. The degree of non-native interactions yet is restricted by holding the chain in a force clamp, which limits the main barrier. By starting the relaxation to low force from not fully extended ubiquitin with residual native secondary structure, our simulations might overestimate the extent of native backbone hydrogen interactions in the intermediate. However, we observe alterations in the secondary structure including the reduction of present and the formation of new helix and sheet motifs. This strongly hints at a role of both native and non-native secondary structure contacts and their interconversion as integral parts of the intermediate ensemble under force.

The role of non-native interactions in early folding events is the subject of ongoing discussion. Intermediate states with a broad distribution of non-native contacts have been observed using NMR and hydrogen exchange (Feng et al., 2003; Feng et al., 2005 Bai, 2006; Tang et al., 2006). Since non-native contacts have to be broken during the folding progress and thus inherently lead to energy barriers, they have been *dismissed* for fast two-state proteins folding on a millisecond scale (Sosnick et al., 1994; Jackson, 1998). Ubiquitin folds within milliseconds in the absence of force (Krantz and Sosnick, 2000), and yet here, under mechanical force, non-native contacts in the intermediate state are found to be important. Large-scale conformational fluctuations involving the continuous reformation of interactions including non-native contacts are shown here to occur on a microsecond time-scale and faster. They thus can be imagined for fast folders as well, supposedly under the guidance of funneling towards an increased native structure content. The significance of non-native structural motifs on the folding pathway also hint towards the weakness of empirical Go-models to describe ubiquitin folding in a force clamp. Being constructed from information from native contacts only, these models seem not to be able to adequately describe the highly fluctuating intermediate states (Best and Hummer, 2005; Cieplak and Szymczak, 2006).

The conformational fluctuations reflected by the end-to-end distance variation involve changes in the protein's hydrophobic and electrostatic interaction patterns. The observed intermediate state features a broad and heterogeneous conformational ensemble regarding both backbone dihedral space and side chain packing. Energetic barriers within this ensemble are low enough to be overcome within the time-scale of the force-clamp MD simulations. The characteristics of the intermediate state, distinct interactions which continuously evolve, including side chain and secondary structure conformational changes of native and non-native nature, and lacking any distinct folding nucleus, are the very features of a molten-globule state (Ptitsyn, 1995a,b). Here, due to the external force, the molten-globule state is stretched out from a globular to a rod-like shape.

The major properties of the conformationally broad and highly dynamic stretched molten-globule state found here

are comparable to non-native states of ubiquitin previously found in solution studies under very different conditions of denaturation. Mass spectrometry and hydrogen exchange studies revealed a partially structured—yet dynamic—molten-globule-like state with a non-uniform distribution of fluctuations along the chain at low pH in methanol (Hoerner et al., 2005). Ubiquitin fragments, that similar to the globular rod here lack substantial tertiary interaction, were found by NMR and circular dichroism studies to locally form native and non-native secondary structure in water as observed here, predominated by β-like and poly-proline II motifs (Shi et al., 2006). Notably, N-terminal fragments comprising β1, β2, or both, show a predisposition for β-structures (Shi et al., 2006; Zerella et al., 1999; Maynard et al., 1998). Infrared-spectroscopy techniques revealed detailed information on the non-native conformation of ubiquitin after temperature and pressure denaturation (Chung et al., 2005; Herberhold and Winter, 2002). Again, the ensemble of the molten globule obtained here under force includes their observed feature of residual structure in the unfolded state, namely a β1–β2 strand as part of a hydrophobic core (compare Fig. 6a and c). In close agreement to these findings, we find partial early onset of β-turn formation and a very low α-helix propensity for the N-terminus (Figs. 4 and 7). Clear indications of hydrophobic clusters in non-native states of ubiquitin detected as folding intermediates, as also found here in the force-clamp simulations, were obtained for mutants with stabilized hydrophobic cores, such as Tyr45Trp, in temperature jump and chemical denaturation experiments, monitoring fluorescence quenching and other variables (Crespo et al., 2006; Khorasanizadeh et al., 1993; Bofill and Searle, 2005; Larios et al., 2004).

An obvious assumption is that an early hydrophobic collapse to a molten globule is a general phenomenon in ubiquitin folding, only hidden for the wild-type under certain experimental conditions leading to apparent two-state folding, *in agreement to recent findings for I27 using a coarse-grained Gō-model* (Li et al., 2004; Li et al., 2006). Here, mechanical force clearly discloses the significant role of hydrophobic interactions in the unfolded collapsed state. A purely random coil, which is highly flexible throughout the chain and does not show any structure or contacts, was only observed under very high urea concentrations in a recent NMR study (Wirmer et al., 2006). Instead, under the less extreme conditions conventionally used in folding experiments, growing evidence for broad dynamic and partially structured ensembles has corroborated the general notion of a non-random unfolded state (Shi et al., 2006). Thus, the stretched molten-globule characterized in this study might be closely related to the molten-globule state characterized in previous studies of non-native ubiquitin. The manifold structural parallels could be interpreted as a proximity or an overlap of the non-native ensembles on the folding free energy landscape probed under the different denaturing conditions. We here examine the collapse from an only partially stretched ubiquitin molecule. It remains to be examined how the initial unfolding conditions, the protein length prior to the force quench as given by the unfolding force, influence the structure and dynamics of the intermediate state. With regard to the conformational diversity in the intermediate observed here, we believe that the finding of an extended molten-globule state as intermediate also holds for initially fully stretched proteins, though assumed on different time scales.

A putative picture of the energy landscape underlying mechanical folding of ubiquitin emerges, a scheme of which is shown in Fig. 8b, reconciling the apparent disagreement between the folding kinetics observed when inducing denaturation by means of force and chemical denaturant (Sosnick, 2004; Fernandez et al., 2004). We map the high-dimensional free energy landscape into two conformational degrees of freedom, which describe the progression variable of folding for (i) force-clamp measurements, d, in terms of the end-to-end distance d, and for (ii) bulk experiments using denaturant, x, a conformational variable such as fluorescence quenching. The free energy profile along d, as outlined in the Results Section, shows a steep increase in free energy for large distances, as described, e.g., by the worm-like chain model, due to the entropic penalty of highly stretched chains. A partially collapsed molten-globule folds over a barrier to the folded state. Similarly, along the different conformational coordinate x, a broad energy well describing the dynamics non-native state ensemble is separated from the folded state via another barrier. In the absence of force, largely extended non-native states are unfavored since less unspecific interactions can be formed in the molten globule. Application of mechanical force now tilts the free energy landscape downwards with $G(x,d) - Fd$, where $G(x,d)$ is the free energy at a given point, and F is the applied force. Consequently, the molten globule is shifted towards larger end-to-end distances by means of force.

Now, two variants need to be considered. First, the two folding pathways probed by mechanical unfolding and unfolding in solution, along d and x, respectively, could be separated from each other by barriers, in which case they would not share states except for the native state. Alternatively, and the similarities of the molten-globule states hint towards this notion, folding might proceed via certain pathways with intermediate values x and d, depending on the details of the free energy landscape, in particular the barrier heights for different combinations of progress variables x and d. The common property of the folding mechanism then would be the molten-globule state, a dynamic and partially structured, more or less collapsed state, and the barrier-limited transition to the folded state. Details of roughness have been omitted in this simplified picture, and might well lead to additional folding intermediates. We predict energy roughness along x to decrease with the distance d, since a less compact state apparently facilitates conformational reorganizations, as discussed above.

We were able to gain molecular insight into the intermediate state during ubiquitin folding in the force clamp,

a hitherto undescribed extended molten-globule. The simulations provided first extensive insight into the complex dynamics of the non-native state of a globular protein under force. A next obvious step is to examine the folding process itself under force clamp to understand what the properties of the molten globule during the last events in the folding process towards the native state are. An ultimate connection between folding from differently (mechanically, chemically, thermally etc.) denatured states will require the integration of the different folding progress variables into the experimental measurements and theoretical treatments of protein folding.

Acknowledgments

We thank Julio Fernandez and all members of the Fernandez lab, as well as Udo Schmitt for stimulating discussions. This work was supported by Volkswagen Foundation Grants I/78 420 and I/80 585 (to H.G. and F.G.) and a Ph.D. scholarship of the Boehringer Ingelheim Fonds (to F.G.).

References

Bai, Y.W., 2006. Energy barriers, cooperativity, and hidden intermediates in the folding of small proteins. Biochem. Biophys. Res. Commun. 340, 976–983.

Baum, J., Dobson, C.M., Evans, P.A., Hanley, C., 1989. Characterization of a partly folded protein by NMR methods—studies on the molten globule state of guinea-pig alpha-lactalbumin. Biochemistry 28, 7–13.

Berendsen, H.J.C., Postma, J.P.M., van Gunsteren, W.F., Hermans, J., 1981. Interaction Model for Water in Relation to Protein Hydration. D. Reidel Publishing Company, Dordrecht, The Netherlands.

Berendsen, H.J.C., Postma, J.P.M., van Gunsteren, W.F., Di Nola, A., Haak, J.R., 1984. Molecular dynamics with coupling to an external bath. J. Chem. Phys. 81 (8), 3684–3690.

Best, R.B., Hummer, G., 2005. Comment on "Force-clamp spectroscopy monitors the folding trajectory of a single protein. Science 308.

Best, R.B., Hummer, G., 2006. Diffusive model of protein folding dynamics with Kramers turnover in rate. Phys. Rev. Lett. 96.

Bofill, R., Searle, M.S., 2005. Engineering stabilising beta-sheet interactions into a conformationally flexible region of the folding transition state of ubiquitin. J. Mol. Biol. 353, 373–384.

Canet, D., Last, A.M., Tito, P., Sunde, M., Spencer, A., Archer, D.B., Redfield, C., Robinson, C.V., Dobson, C.M., 2002. Local cooperativity in the unfolding of an amyloidogenic variant of human lysozyme. Nat. Struct. Biol. 9, 308–315.

Cheung, M.S., Garcia, A.E., Onuchic, J.N., 2002. Protein folding mediated by solvation: water expulsion and formation of the hydrophobic core occur after the structural collapse. Proc. Natl. Acad. Sci. USA 99, 685–690.

Chung, H.S., Khalil, M., Smith, A.W., Ganim, Z., Tokmakoff, A., 2005. Conformational changes during the nanosecond-to-millisecond unfolding of ubiquitin. Proc. Natl. Acad. Sci. USA 102, 612–617.

Cieplak, M., Szymczak, P., 2006. Protein folding in a force clamp. J. Chem. Phys. 124.

Crespo, M.D., Simpson, E.R., Searle, M.S., 2006. Population of on-pathway intermediates in the folding of ubiquitin. J. Mol. Biol. 360, 1053–1066.

Darden, T., York, D., Pedersen, L., 1993. Particle Mesh Ewald—An Nlog(N) method for Ewald sums in large systems. J. Chem. Phys. 98, 10089–10092.

DeLano, W.L., 2001. PyMOL Manual. DeLano Scientific, San Carlos, CA, USA. <http://www.delanoscientific.com>.

Feng, H.Q., Takei, J., Lipsitz, R., Tjandra, N., Bai, Y.W., 2003. Specific non-native hydrophobic interactions in a hidden folding intermediate: Implications for protein folding. Biochemistry 42, 12461–12465.

Feng, H.Q., Zhou, Z., Bai, Y.W., 2005. A protein folding pathway with multiple folding intermediates at atomic resolution. Proc. Natl. Acad. Sci. USA 102, 5026–5031.

Fernandez, J.M., Li, H.B., 2004. Force-clamp spectroscopy monitors the folding trajectory of a single protein. Science 303, 1674–1678.

Fernandez, J.M., Li, H.B., Brujic, J., 2004. Response to comment on Force-clamp spectroscopy monitors the folding trajectory of a single protein. Science 306, 411.

Gō, N., Abe, H., 1981. Non-interacting local-structure model of folding and unfolding transition in globular-proteins .1. Formulation. Biopolymers 20, 991–1011.

Gräter, F., Shen, J.H., Jiang, H.L., Gautel, M., Grubmüller, H., 2005. Mechanically induced titin kinase activation studied by force-probe molecular dynamics simulations. Biophys. J. 88, 790–804.

Gsponer, J., Hopearuoho, H., Whittaker, S.B.M., Spence, G.R., Moore, G.R., Paci, E., Radford, S.E., Vendruscolo, M., 2006. Determination of an ensemble of structures representing the intermediate state of the bacterial immunity protein Im7. Proc. Natl. Acad. Sci. USA 103, 99–104.

Hänggi, P., Talkner, P., Borkovec, M., 1990. Reaction-rate theory—50 years after Kramers. Rev. Modern Phys. 62, 251–341.

Herberhold, H., Winter, R., 2002. Temperature- and pressure-induced unfolding and refolding of ubiquit. A static and kinetic fourier transform infrared spectroscopy study. Biochemistry 41, 2396–2401.

Hess, B., Bekker, H., Berendsen, H.J.C., Fraaije, J.G.E.M., 1997. LINCS: a linear constraint solver for molecular simulations. J. Comput. Chem. 18, 1463–1472.

Hoerner, J.K., Xiao, H., Kaltashov, I.A., 2005. Structural and dynamic characteristics of a partially folded state of ubiquitin revealed by hydrogen exchange mass spectrometry. Biochemistry 44, 11286–11294.

Irback, A., Mitternacht, S., Mohanty, S., 2005. Dissecting the mechanical unfolding of ubiquitin. Proc. Natl. Acad. Sci. USA 102, 13427–13432.

Jackson, S.E., 1998. How do small single-domain proteins fold? Fold. & Des. 3, 81–91.

Khorasanizadeh, S., Peters, I.D., Butt, T.R., Roder, H., 1993. Folding and stability of a tryptophan-containing mutant of ubiquitin. Biochemistry 32, 7054–7063.

Klimov, D.K., Thirumalai, D., 1997. Viscosity dependence of the folding rates of proteins. Phys. Rev. Lett. 79, 317–320.

Kramers, H.A., 1940. Brownian motion in a field of force and the diffusion model of chemical reactions. Physica (Utrecht), VII, 284–304.

Krantz, B.A., Sosnick, T.R., 2000. Distinguishing between two-state and three-state models for ubiquitin folding. Biochemistry 39, 11696–11701.

Krantz, B.A., Moran, L.B., Kentsis, A., Sosnick, T.R., 2000. D/H amide kinetic isotope effects reveal when hydrogen bonds form during protein folding. Nat. Struct. Biol. 7, 62–71.

Larios, E., Li, J.S., Schulten, K., Kihara, H., Gruebele, M., 2004. Multiple probes reveal a native-like intermediate during low-temperature refolding of ubiquitin. J. Mol. Biol. 340, 115–125.

Li, P.C., Makarov, D.E., 2004. Ubiquitin-like protein domains show high resistance to mechanical unfolding similar to that of the I27 domain in titin: Evidence from simulations. J. Phys. Chem. B 108, 745–749.

Li, M.S., Klimov, D.K., Thirumalai, D., 2004. Thermal denaturation and folding rates of single domain proteins: size matters. Polymer 45, 573–579.

Li, M.S., Hu, C.K., Klimov, D.K., Thirumalai, D., 2006. PNAS Multiple stepwise refolding of immunoglobulin domain i27 upon force quench depends on initial conditions. Proc. Nat. Acad. Sci. 103, 93–98.

Lindahl, E., Hess, B., van der Spoel, D., 2001. GROMACS 3.0: a package for molecular simulation and trajectory analysis. J. Mol. Model. 7, 306–317.

Liu, Z.R., Chan, H.S., 2005. Desolvation is a likely origin of robust enthalpic barriers to protein folding. J. Mol. Biol. 349, 872–889.

Marszalek, P.E., Lu, H., Li, H., Carrion-Vazquez, M., Oberhauser, A.F., Schulten, K., Fernandez, J.M., 1999. Mechanical unfolding intermediates in titin modules. Nature 402, 100–103.

Maynard, A.J., Sharman, G.J., Searle, M.S., 1998. Origin of beta-hairpin stability in solution: Structural and thermodynamic analysis of the folding of model peptide supports hydrophobic stabilization in water. J. Am. Chem. Soc. 120, 1996–2007.

Michalet, X., Weiss, S., Jager, M., 2006. Single-molecule fluorescence studies of protein folding and conformational dynamics. Chem. Rev. 106, 1785–1813.

Naganathan, A.N., Sanchez-Ruiz, J.M., Munoz, V., 2005. Direct measurement of barrier heights in protein folding. J. Am. Chem. Soc. 127, 17970–17971.

Nicholls, A., Sharp, K.A., Honig, B., 1990. DelPhi V3.0. Columbia University, New York.

Nishimura, C., Dyson, H.J., Wright, P.E., 2006. Identification of native and non-native structure in kinetic folding intermediates of apomyoglobin. J. Mol. Biol. 355, 139–156.

Onuchic, J.N., Wolynes, P.G., 2004. Theory of protein folding. Curr. Opin. Struct. Biol. 14, 70–75.

Ptitsyn, O.B., 1995a. Molten globule and protein folding. Adv. Protein Chem. 47, 83–229.

Ptitsyn, O.B., 1995b. How the molten globule became. Trends Biochem. Sci. 20, 376–379.

Rank, J.A., Baker, D., 1997. A desolvation barrier to hydrophobic cluster formation may contribute to the rate-limiting step in protein folding. Protein Sci. 6, 347–354.

Rief, M., Gautel, M., Oesterhelt, F., Fernandez, J.M., Gaub, H.E., 1997. Reversible unfolding of individual titin immunoglobulin domains by AFM. Science 276, 1109–1112.

Rodriguez-Larrea, D., Minning, S., Borchert, T.V., Sanchez-Ruiz, J.M., 2006. Role of solvation barriers in protein kinetic stability. J. Mol. Biol. 360, 715–724.

Samori, B., Zuccheri, G., Baschieri, P., 2005. Protein unfolding and refolding under force: Methodologies for nanomechanics. ChemPhysChem 6, 29–34.

Schlierf, M., Li, H.B., Fernandez, J.M., 2004. The unfolding kinetics of ubiquitin captured with single-molecule force-clamp techniques. Proc. Natl. Acad. Sci. USA 101, 7299–7304.

Schuler, B., 2005. Single-molecule fluorescence spectroscopy of protein folding. ChemPhysChem 6, 1206–1220.

Schuler, B., Lipman, E.A., Eaton, W.A., 2002. Probing the free-energy surface for protein folding with single-molecule fluorescence spectroscopy. Nature 419, 743–747.

Shakhnovich, E., 2006. Protein folding thermodynamics and dynamics: where physics, chemistry, and biology meet. Chem. Rev. 106, 1559–1588.

Shi, Z.S., Chen, K., Liu, Z.G., Sosnick, T.R., Kallenbach, N.R., 2006. Pii structure in the model peptides for unfolded proteins: studies on ubiquitin fragments and several alanine-rich peptides containing QQQ, SSS, FFF, and VVV. Proteins 63, 312–321.

Shi, Z.S., Chen, K., Liu, Z.G., Kallenbach, N.R., 2006. Conformation of the backbone in unfolded proteins. Chem. Rev. 106, 1877–1897.

Smith, L.J., 2004. Computational methods for generating models of denatured and partially folded proteins. Methods 34, 144–150.

Snow, C.D., Sorin, E.J., Rhee, Y.M., Pande, V.S., 2005. How well can simulation predict protein folding kinetics and thermodynamics? Ann. Rev. Biophys. Biomol. Struct. 34, 43–69.

Socci, N.D., Onuchic, J.N., Wolynes, P.G., 1996. Diffusive dynamics of the reaction coordinate for protein folding funnels. J. Chem. Phys. 104, 5860–5868.

Sosnick, T.R., 2004. Comment on "Force-clamp spectroscopy monitors the folding trajectory of a single protein. Science 306, 411.

Sosnick, T.R., Mayne, L., Hiller, R., Englander, S.W., 1994. The barriers in protein-folding. Nat. Struct. Biol. 1, 149–156.

Sosnick, T.R., Dothager, R.S., Krantz, B.A., 2004. Differences in the folding transition state of ubiquitin indicated by phi and psi analyses. Proc. Natl. Acad. Sci. USA 101, 17377–17382.

Steel, B.C., McKenzie, D.R., Bilek, M.M., Nosworthy, N.J., dos Remedios, C.G., 2006. Nanosecond responses of proteins to ultra-high temperature pulses. Biophys. J. 91, L66–L68.

Tang, Y.F., Goger, M.J., Raleigh, D.P., 2006. NMR characterization of a peptide model provides evidence for significant structure in the unfolded state of the villin headpiece helical subdomain. Biochemistry 45, 6940–6946.

van Gunsteren, W.F., Billeter, S.R., Eising, A.A., Hünenberger, P.H., Krüger, P., Mark, A.E., Scott, W.R.P., Tironi, I.G., 1996. Biomolecular Simulation: The GROMOS96 Manual and User Guide. Vdf Hochschulverlag AG an der ETH Zürich, Zürich, Switzerland.

Vijaykumar, S., Bugg, C.E., Cook, W.J., 1987. Structure of ubiquitin refined at 1.8 Å, resolution. J. Mol. Biol. 194, 531–544.

Vriend, G., 1990. WHAT IF: a molecular modelling and drug design program. J. Mol. Graph. 8, 52–56.

Wirmer, J., Petiund W., Schwalbe, H., 2006. Motional properties of unfolded ubiquitin: a model for a random coil protein. J. Biomol. NMR. Epub ahead of print.

Zerella, R., Evans, P.A., Ionides, J.M.C., Packman, L.C., Trotter, B.W., Mackay, J.P., Williams, D.H., 1999. Autonomous folding of a peptide corresponding to the N-terminal beta-hairpin from ubiquitin. Protein Sci. 8, 1320–1331.

Zhuang, X.W., Rief, M., 2003. Single-molecule folding. Curr. Opin. Struct. Biol. 13, 88–97.