Minimum energy compact structures in force-quench polyubiquitin folding are domain swapped

Fei Xia[a,b], D. Thirumalai[c,1], and Frauke Gräter[a,b,1]

[a]Chinese Academy of Sciences Max Planck Gesellschaft Partner Institute and State Key Laboratory for Computational Biology, Shanghai 200031, People's Republic of China; [b]Heidelberg Institute for Theoretical Studies, 69118 Heidelberg, Germany; and [c]Biophysics Program, Institute for Physical Science and Technology, University of Maryland, College Park, MD 20742

Edited* by José N. Onuchic, University of California, San Diego, La Jolla, CA, and approved March 8, 2011 (received for review December 6, 2010)

Single molecule experiments that initiate folding using mechanical force are uniquely suited to reveal the nature of populated states in the folding process. Using a strategy proposed on theoretical grounds, which calls for repeated cycling of force from high to low values using force pulses, it was demonstrated in atomic force spectroscopy (AFM) experiments that an ensemble of minimum energy compact structures (MECS) are sampled during the folding of polyubiquitin. The structures in the ensemble are mechanically resistant to a lesser extent than the native state. Remarkably, forced unfolding of the populated intermediates reveals a broad distribution of extensions including steps up to 30 nm and beyond. We show using molecular simulations that favorable interdomain interactions leading to domain swapping between adjacent ubiquitin modules results in the formation of the ensemble of MECS, whose unfolding leads to an unusually broad distribution of steps. We obtained the domain-swapped structures using coarse-grained ubiquitin dimer models by exchanging native interactions between two monomeric ubiquitin molecules. Brownian dynamics force unfolding of the proposed domain-swapped structures, with mechanical stability that is approximately 100-fold lower than the native state, gives rise to a distribution of extensions from 2 to 30 nm. Our results, which are in quantitative agreement with AFM experiments, suggest that domain swapping may be a general mechanism in the assembly of multi-sub-unit proteins.

folding intermediates | atomic force microscopy | self-organized polymer model | unfolding kinetics | three stage refolding

The remarkable processes governing protein and RNA folding described using polymer concepts, statistical mechanics, and molecular simulations have formed the conceptual basis for understanding a large number of experiments (1–6). Several global predictions including the presence of pathway diversity, the dependence of refolding rates of proteins and RNA on length, collapse transition, and the role of roughness due to topological and energetic frustration (7, 8) on folding rates have been experimentally validated. Control of folding through application of a constant mechanical force (f) at the single molecule level has generated trajectories that have given an unprecedented picture of the folding landscape projected onto molecular extension as a conjugate to f (9). The results of these experiments have confirmed theoretical predictions that refolding pathways are heterogeneous even for proteins that ostensibly fold in an apparent two-state manner.

In typical force-clamp experiments, a constant f is applied to the termini of the protein resulting in an increase in end-to-end length. In polyubiquitin [$(Ub)_n$], a system that has been extensively characterized, forced unfolding occurs in steps of about 20.3 nm (10). Refolding of Ub can be initiated by unfolding the protein at a high stretching force (f_S) and then quenching f a low value, $f_Q \sim 10$ pN, to allow relaxation to the native state (11–13). It is found Ub refolds in multiple steps (11, 14, 15), namely a rapid collapse due to the elasticity of the stretched chain, and a relatively slower refolding process on the time scale of seconds through a broad conformational ensemble of intermediate states (16). Subsequent force-clamp experiments of single ubiquitin (12) also showed a similar mechanism suggesting domain–domain interactions play only a minor role in the collapse dynamics from the stretched to the fully collapsed state. The molecular details of ubiquitin folding have been revealed in a variety of theoretical simulations using coarse-grained and all-atom models (17–21), thereby complementing the experimental investigations.

Single molecule experiments are also well suited to address the vexing question of the role of compact intermediates, whose existence and structures are hard to discern using ensemble experiments. Using theoretical analysis, Barsegov and Thirumalai (BT) (22) proposed that the nature of collapsed intermediates could be probed by using force pulses in which f is quenched from f_S to f_Q for a duration Δt after which f is increased to f_S to induce global unfolding. If this process is repeated multiple times by varying Δt, then the plausible intermediates that are accessed in the folding trajectories can be experimentally characterized by extension, ΔR, which is the natural reaction coordinate in single molecule experiments. The two global time scales that characterize the $f_S \rightarrow f_Q$ quench are the collapse time $\tau_c(f_Q)$ and the folding time $\tau_F(f_Q)$, both of which depend only weakly on f_S (23). If $\tau_c(f_Q) < \Delta t < \tau_F(f_Q)$, then the response of the collapsed [referred to as minimum energy compact structures (MECS)], but not folded structures, to f can be probed using single molecule force spectroscopy experiments. The BT strategy was implemented in insightful experiments on Ub_n (24). By exploiting the differences in Δt-dependent unfolding properties between the MECS and the native state upon $f_Q \rightarrow f_S$ ramp, Garcia-Manyes et al. argued that an ensemble of heterogeneous compact structures must guide refolding, which confirmed predictions that a countable number of MECS are plausible intermediates in the folding reaction (25).

Experiments using force pulses (24) yielded two key results that are relevant to the present study. (*i*) Forces required to disrupt the MECS are less than those needed to unfold the native structure. The reduced mechanical stability implies that not all of the interactions that stabilize the folded monomer are fully realized in the MECS. (*ii*) Remarkably, the distribution of unfolding lengths upon increasing $f_Q \rightarrow f_S$ when $\frac{\Delta t}{\tau_F(f_q)} < 1$ is extremely broad, with ΔR ranging from 1 to 60 nm, although in the majority of the molecules (~88%) the unfolding lengths extend from 1 to 30 nm. Step sizes up to 30 nm are remarkable because any imaginable nonnative collapsed topology of ubiquitin comprising of only one domain can give rise to extensions of at most

Author contributions: D.T. and F.G. designed research; F.X. performed research; F.X., D.T., and F.G. analyzed data; and F.X., D.T., and F.G. wrote the paper.

The authors declare no conflict of interest.

*This Direct Submission article had a prearranged editor.

Freely available online through the PNAS open access option.

[1]To whom correspondence may be addressed. E-mail: thirum@umd.edu or frauke@picb.ac.cn.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1018177108/-/DCSupplemental.

approximately 25 nm, the protein length of a stretched ubiquitin protein at approximately 100 pN, minus its length in the collapsed state, which is approximately 5 nm in the native ubiquitin topology (26).

Here, we hypothesize that domain swapping (27–29) involving adjacent Ub monomers is responsible for the broad distribution of extensions (24). It occurs if the topology of a homodimer, consisting of either covalently linked or individual monomers, allows a certain sequence motif of one monomer to fold onto the second domain and vice versa. Because identical sequences are exchanged, they can in principle pack onto the other monomer as optimally as in the nonswapped structures. Although some proteins have been reported to swap whole tertiary globular domains (30), protein monomers generally tend to exchange a single element such as a strand of β-sheet or an α-helix with their identical counterpart. We tested our hypothesis using force-clamp Brownian dynamics simulations of a coarse-grained Ub_2 dimer construct. Because the structures of swapped Ub_2 are not known, we constructed a number of putative domain-swapped (DS) structures (Fig. 1) by noting that protein topology is the major determinant of forced-unfolding pathways as inferred from unfolding step sizes ΔR (31). Using the putative DS structures (Fig. 1) as plausible models for MECS, we simulated their unfolding behavior and distribution of unfolding lengths upon application of constant f. The various domain-swapped dimers, which form readily after force quench, are mechanically weak compared to the native fold. As f is increased, they unfold with a broad $P(\Delta R)$ of stepwise extensions. Our results establish that DS polyubiquitin conformations are likely candidates for the ensemble of compact folding intermediates probed in force spectroscopy experiments using the force pulse protocol.

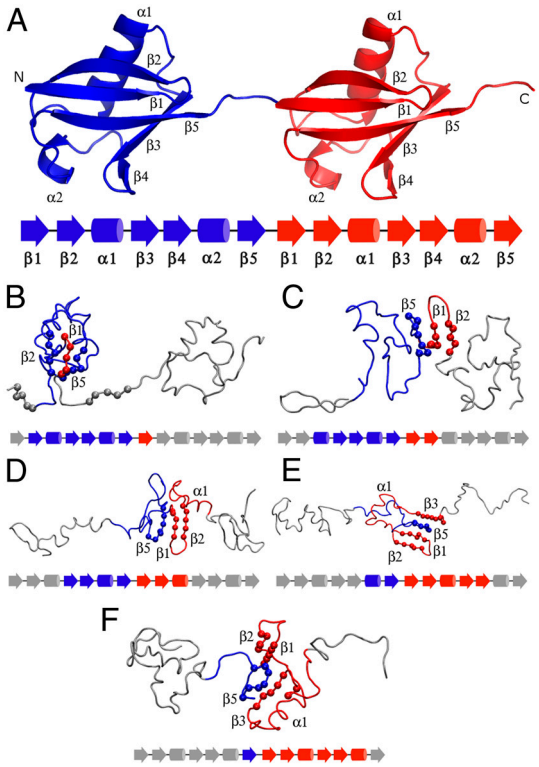

Fig. 1. Structures of Ub_2. (*A*) Schematic illustration of the ubiquitin dimer constructed from the ubiquitin crystal structure (PDB ID code 1UBQ). The five β-sheets are labeled from β1 to β5 and the two α-helices are α1 and α2. The first domain in blue consists of β1(2–6), β2(12–16) α1(23–34), β3(41–45) β4(48–49), α2(56–59), and β5(65–71); the second red domain contains β1(78–82), β2 (88–92), α1(99–110), β3(117–121), β4(124–125), α2(132–135), and β5(141–147). (*B–F*) The five structures are used as models for domain swapping. The sequence shows the secondary structure elements of the ubiquitin dimer from the first blue to the second red, where the blue and red parts indicate the proposed DS structures and the gray represents the unswapped region.

Results and Discussion

Unfolding and Refolding of Natively Folded Ubiquitin Dimer. In order to validate the two-domain ubiquitin model, we first compared the unfolding and folding kinetics of the native nonswapped topology under force clamp to experiments and previous simulations of polyubiquitin and monomers. Ub_2 was initially pulled over a range of forces from 160 to 200 pN by subjecting the C terminus of ubiquitin to a stretching force. The simulations were started with random velocities, and we generated 50 unfolding trajectories at each force. The average end-to-end distance $\langle R(t)\rangle$ as a function of time (t) is well fit with a single exponential function (Fig. 2*A*). Based on the Bell model, the distance to the transition state is found to be $\Delta x_U \approx 0.20$ nm, which is in the range of (0.16–0.25) nm, found in several previous studies (10, 32–35). Fits using linear cubic free energy profile give $\Delta x_U \approx 0.15$ nm (Fig. 2*A*) (36, 37). Thus, there is consistency between the theoretical models and experiments. In addition, we find a narrow distribution of unfolding step lengths with a peak at 20 nm, which is the fingerprint for ubiquitin unfolding (Fig. S1) (12).

Refolding simulations were initiated from a stretched conformation taken from a snapshot at $f_S = 200$ pN. We varied from $f_Q = 2.5$ pN to 15.0 pN. Fig. 2*B*, which shows the end-to-end distance $\langle R(t)\rangle$ averaged over refolding trajectories, decreases by a multistep process (11, 14, 15) and is well fit using a biexponential function. Refolding time follows $\tau_F \approx \tau(0)\exp(\Delta x_F f_Q/k_B T)$, where Δx_F is the distance between the stretched and the transi-

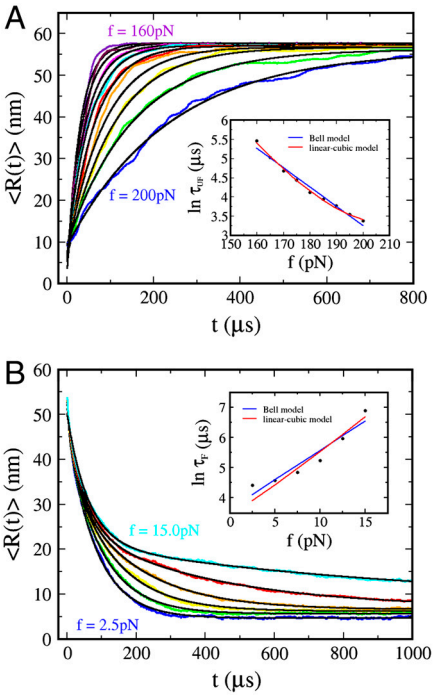

Fig. 2. Time dependence of the end-to-end distance, $R(t)$. (*A*) Increase in $\langle R(t)\rangle$ averaged over all the trajectories upon forced unfolding of Ub_2. Forces vary from 160 to 200 pN. The averaged end-to-end distances $\langle R(t)\rangle$ are fit using $\langle R(t)\rangle = A_0 - A_1\exp(-t/\tau_{UF})$, where τ_{UF} denotes the unfolding time. (*Inset*) A fit to the logarithm of unfolding times τ_{UF} using the Bell model (blue) and a cubic potential (red). (*B*) Refolding of the Ub_2 under quench forces (f_Qs) ranging from 2.5 to 15.0 pN. Decreases in $\langle R(t)\rangle$ are fit using biexponential function $\langle R(t)\rangle = A_0 - A_1\exp(-t/\tau_1) - A_2\exp(-t/\tau_2)$, where τ_2 refers to f_Q-dependent folding time τ_F. (*Inset*) A fit to τ_F as a function of f_Q using the Bell model (blue) and a cubic potential (red).

tion states. From the linear fit in Fig. 2*B*, we obtain $\Delta x_F \approx$ 0.77 nm, which is close to previous estimations from monomeric and trimeric ubiquitin simulations (34) and is also in agreement with the experimental value of 0.80 nm (11). Nonlinear dependencies between the logarithm of the folding times and f as observed here have been previously found for other proteins including polyimmunoglobulin (12). The f_Q dependence of refolding rates upon force quench can also be derived by assuming a one-dimensional cubic potential (see *SI Text* in ref. 23). Fit of the data in Fig. 2*B* using the theoretical result to describe refolding yields $\Delta x_F \approx 0.85$ nm, which is close to the value obtained using the standard Bell model. Overall, our results establish that the dimer model exhibits unfolding and folding kinetics both in agreement with previous experiments and theoretical simulations.

Unfolding Mechanism of Domain-Swapped Structures. In order to test whether the ensemble of domain-swapped ubiquitin structures can account for the broad distribution of unfolding lengths observed in experiments (24), we examined the forced unfolding of five DS ubiquitin dimer topologies. We generated representative structures of the putative MECS by starting from the stretched state and quenching the force to $f_Q = 10$ pN to allow folding into the respective swapped structures (Fig. 1 *B–F*). Because we predefined the extent of swapping by native contacts in the self-organized polymer (SOP) model our simulations cannot predict the relative likelihood of forming the proposed DS structures. However, the contact orders (38) of the swapped topologies range between 0.24 and 0.30, which is close to the value of the native nonswapped fold (0.29), indicating a reasonable degree of foldability of the proposed native-like DS structures.

The mechanical stability and unfolding mechanism of the swapped ubiquitin topologies were probed by subjecting the structures in Fig. 1 *B–F* to a series of low forces to generate an ensemble of unfolding trajectories. The domain-swapped conformations are mechanically stable at low forces (10 pN) and show instantaneous unfolding and absence of a plateau in the folding trajectories at high forces typically used for disrupting the native ubiquitin fold ($f > 100$ pN). Interestingly, at medium forces (50–70 pN depending on the swapped topology), we found a subset of trajectories with distinct plateaus in the end-to-end length, which implies that the DS structures unfold in a stepwise manner. Representative end-to-end distance curves for the unfolding at 60 pN of the topology for which a single β-strand, namely β_1, was swapped (Fig. 1*B*) are shown in Fig. 3*A*. The unfolding mechanism varies between trajectories, involving different steps in the end-to-end length and with distinct unfolding intermediates. In the case of the black curve shown in Fig. 3*A*, $R(t)$ shows a stable plateau and two unfolding steps with length $R_1 = 9.7$ nm and $R_2 = 28.4$ nm. Native contacts involving domain-swapped segments β_2 and β_1 (compare snapshot in Fig. 3*A*) partly rupture in the first unfolding step, which leads to a small extension R_1 and a drop in Q_{12} within the first 50 μs (Fig. 3*B*). The remainder of the folded structure withstands forces up to 60 pN for approximately 300 μs, when the second rip leads to a cooperative unraveling of all native contacts including the interdomain interaction between β_1 and β_5 as indicated by the final drop in Q_{12}. By contrast, the red curve exhibits an unfolding scenario with two comparable step sizes $R_1 = 20.8$ nm and $R_2 = 17.0$ nm.

We measured the step sizes observed in a total of 50 unfolding trajectories at the pulling forces 60 and 65 pN. The histograms of step length distributions, presented in Fig. 3 *C* and *D*, are broad with ΔR ranging from 2 to 30 nm, thus covering the experimental range found for folding intermediates (24). The separate distributions of R_1 and R_2 at 60 and 65 pN are provided in Fig. S2. We also performed unfolding simulations for the other domain-swapped candidates in Fig. 1 *C–F*. The resulting histograms

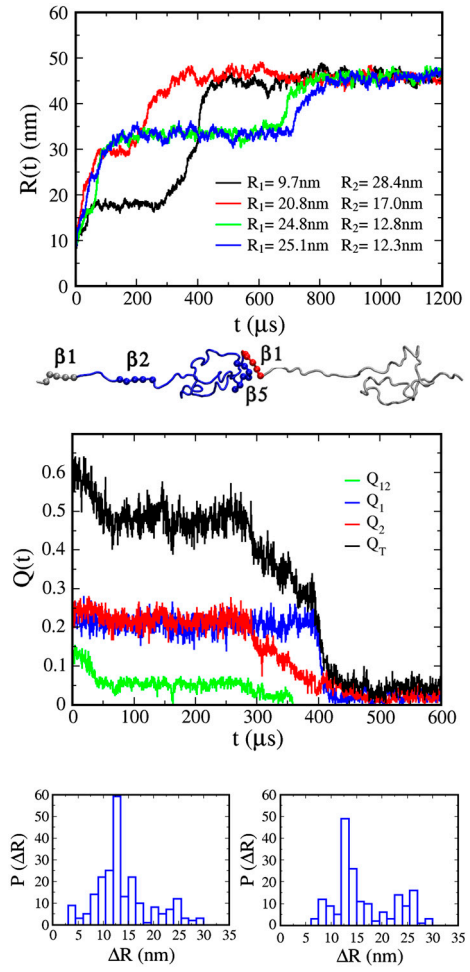

Fig. 3. Heterogeneity in forced unfolding and refolding pathways. (*A*) Four $R(t)$ traces for unfolding the domain-swapped structure in the Fig. 1*B* at $f = 60$ pN, where the plateaus indicate stable intermediates. The snapshot of intermediates at plateaus for the black curve is shown below. (*B*) Time dependence of fraction of native contacts Q_{12}, Q_1, Q_2, and Q_T for the black curve in *A*. (*C*) Distribution of extensions obtained from unfolding trajectories at 60 pN. (*D*) $P(\Delta R)$ for unfolding trajectories at 65 pN.

are shown in Fig. 4 *A–D*. The swapping topologies in Fig. 1 *C* and *D*, both with strands β_1 and β_2 swapped, exhibit broad distributions covering $\Delta R = 8$–26 nm upon unfolding, with two major peaks and a predominant unfolding intermediate with the swapped interactions. With an increasing number of strands swapped between the two domains (topologies in Fig. 1 *E* and *F*), we obtained a narrower distribution of step sizes with $\Delta R =$ 10–25 nm. Unfolding in this case occurs via an intermediate such that the two unfolding steps involve similar extensions (Fig. S3).

Previous studies of single Ub forced unfolding showed protein rips through a uniquely defined sequence of steps. The dominant step sizes were peaked at 8.1 ± 0.7 nm and 12.4 ± 1.0 nm (10, 39). The appearance of a plateau was observed with low probability. In contrast, we find numerous intermediates with various step sizes when domain-swapped ubiquitin dimer conformations are forced to unfold. Apparently, domain swapping lowers the mechanical resistance of the individual modules relative to the nonswapped topology of native ubiquitin. At $f \sim 60$ pN, contacts formed within and between adjacent domains rip simultaneously or in discrete steps of varying lengths thus yielding a broad $P(\Delta R)$.

Unfolding Kinetics of Domain-Swapped Structures. To determine the dynamics of the weakly mechanically resistant ensemble of MECS probed using the force pulse sequences, Garcia-Manyes

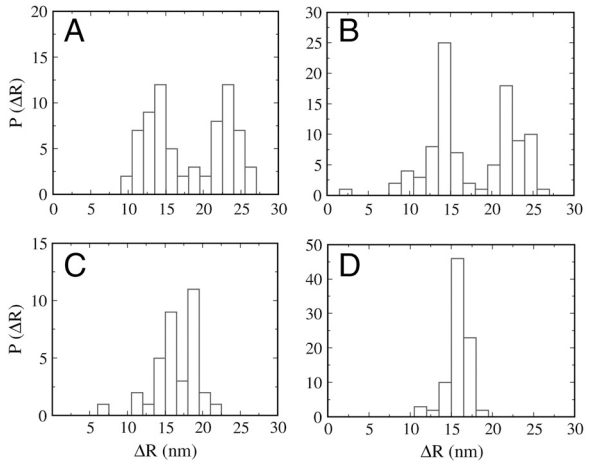

Fig. 4. Broad distribution of extension (Δ*R*) for DS structures. (*A–D*) Distributions of Δ*R* estimated from trajectories of unfolding domain-swapped structures in Fig. 1 *C–F*, respectively. *A–C* are obtained for Fig. 1 *C–E* at 60 pN, whereas *D* is for Fig. 1*F* at $f = 50$ pN.

et al. (24) have estimated the unfolding rates at low forces, which are insufficient to unfold native Ub. The measured f-dependent unfolding rates $[k_{u;DS}(f)]$ of the collapsed intermediate states were larger than the unfolding rates $[k_{u;N}(f)]$ of the native state by almost a factor of 100. We made the same comparison of the unfolding rates between the native and swapped structures. First, the domain-swapped structures unfold at forces that are considerably less than those required to rupture the native state (Fig. 5). More importantly, the calculated values of $k_{u;DS}(f)$ (squares in Fig. 5) are between 90- and 100-fold greater than the rates $k_{u;N}(f)$ (dots in Fig. 5) when extrapolating to the same unraveling forces. The predicted trends (both the unfolding forces and the ratios of $k_{u;DS}(f)$ and $k_{u;N}(f)$ are in quantitative agreement with experiments (24), which further lend credence to our proposal that the MECS sampled using the force pulse refolding are domain swapped. From a quantitative analysis of the data in Fig. 5 using the Bell model, we obtain an unfolding rate at zero force for the domain-swapped collapsed states, $k_{DS}(0)$, which is approximately 96 times larger than the corresponding rate of the native state, $k_N(0)$. Assuming similar prefactors A in the Arrhenius equations for $k_{N,DS} = A\exp(-\Delta G_{N,DS}/k_BT)$, the difference in the unfolding free energy barriers between the native or domain-swapped MECS is approximately $\Delta\Delta G = \Delta G_N - \Delta G_{DS} \approx$ 2.4 kcal/mol, in good agreement with the experimentally derived value of 2.6 kcal/mol (24). In addition, the estimated distances to transition states of Δx_{DS} and Δx_N for the collapsed structures and the native state from our model are 0.23 nm and 0.20 nm,

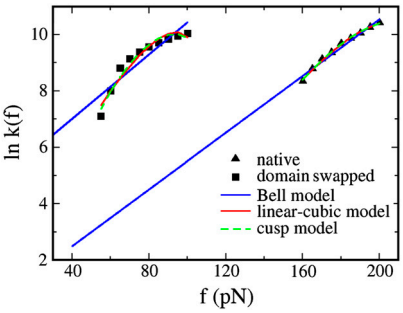

Fig. 5. Comparison of the force-dependent rates between the DS and native structures. The rate constants $k(f)$ were calculated as $1/\tau_{UF}$, where the τ_{UF} were obtained by fitting the averaged unfolding trajectories using $\langle R(t)\rangle = A_0 - A_1\exp(-t/\tau_{UF})$. The logarithms of $k(f)$ were fitted to the forces f with the Bell model, the linear-cubic, and cusp potentials.

respectively. A shift of mere $\Delta\Delta x = \Delta x_{DS} - \Delta x_N \approx 0.03$ nm, which is also in good agreement with the experimentally measured 0.04 nm (24), suggests a specific collapse of the MECS to a native-like topology. By combining the broad length distribution of unfolding lengths of the mechanically compromised compact structures that unfold rapidly at low forces we envision the MECS most straightforwardly as adjacent domains with swapped β-strands. We note that use of linear-cubic (36, 37) and cusp models (36) result in fits that follow more closely the calculated force dependence of the rates (Fig. 5). However, as expected, they result in differences between the swapped and native states in terms of free energy, $\Delta\Delta G$ (6.3–8.7 kcal/mol), and distance to the transition state, $\Delta\Delta x$ (0.011–0.025 nm), which deviate more from the values extracted using experiments, which were obtained using the Bell model.

Further evidence that swapping gives rise to the folding intermediates, which explains the broad distribution of unfolding lengths, we find that the unfolding rates are independent of the collapse time (24). Fig. S4 shows the averaged $R(t)$ obtained by pulling at a force of $f_S = 60$ pN. The waiting time between quenching the force to $f_Q = 10$ pN for collapse and stepping the force back to 60 pN was varied from 0.4 to 1.6 ms. We obtained the two rate constants by fitting $R(t)$ with a biexponential function, as was done in the atomic force spectroscopy (AFM) experiments (24). Fig. S4, *Inset* shows that the fast unfolding rate k_1 (black dots) does not depend on the waiting time, which is also in agreement with the data in figure 2*B* of ref. 24.

Conclusions

The ensemble of minimum energy compact structures that is populated during the folding process using force pulse sequence in AFM experiments is remarkably resistant toward force, which implies that the collapsed intermediates are ordered and perhaps have native-like features. These findings and the surprising observation that the unfolding of the populated intermediates results in a broad distribution of step sizes, ranging from (1–30) nm (24), are explained in this work based on the assumption that the structures of the compact states involve domain swapping. Interestingly, swapping of only the β_1 strand (Fig. 1) leads to a topology whose unfolding results in heterogeneous pathways observed in experiments. It is possible that, besides the proposed DS structures, forced unfolding of other compact mechanically weak conformations could also explain the observed broad $P(\Delta R)$. Our work makes it clear that such structures must also involve interdomain interactions that are strong enough to stabilize the intermediates for detection in AFM experiments. However, given that ubiquitin is optimally packed in the native fold, the domain-swapped topologies with native-like contacts appear to be the simplest candidates for the MECS detected in AFM experiments. The unusually large steps in the unfolding lengths involves rupture of contacts beyond a single ubiquitin domain resulting in an intermediate mechanical stability by maintaining β-sheet structures. We speculate that domain swapping could also be observed in the folding of polyproteins (or heterodimers with similar topology) using force-quench methods (22, 24). Our predictions can be tested using force spectroscopy of polyubiquitin by inserting additional amino acids into the domain connecting linker, which should lead to a further broadening of the length distribution beyond the 30 nm observed in AFM experiments (24). Another consequence of our work, which is amenable to experimental test, is that if AFM or laser optical tweezer experiments are performed by attaching handles (DNA or polyimmunoglobulin domains) to a monomeric Ub using force pulse sequence, then the distribution of unfolding step lengths of Ub would be narrow.

It is known that Ub can aggregate at modest concentrations, which has complicated the interpretation of ensemble experiments. In general, high protein concentrations can lead to transient domain swapping during folding (40). It has been proposed

that the enhanced local concentration of Ub in Ub_n constructs used in single molecule experiments can result in interdomain interactions (41). Although this scenario is unlikely when $\frac{\Delta t}{\tau_F(f_Q)} \gg 1$, as was the case in ref. 11, it is indeed highly probable if Δt is short. If the quench time is long compared to the refolding time, then the metastable MECS will anneal to form native structures. Our work predicts that by lengthening Δt the probability of domain-swapped structure formation can be greatly reduced. A corollary of our work is that even weak interdomain interactions can impede refolding. Interestingly, the folding rate of I27 monomers is higher than in the polyprotein counterparts (2.9 s^{-1} versus 1.9 s^{-1} at $f_Q = 0$, respectively) (12). Our study readily explains these observations by attributing slower folding in polyprotein constructs to transient formation of domain-swapped intermediates, which in a folding trajectory cannot be easily distinguished from other MECS.

Although we focused on the origin of broad $P(\Delta R)$ in force-quench folding of Ub_n, our work also highlights the mechanism of folding upon force quench. We find that typically folding occurs in three stages (Fig. 3*B*) with the final stage involving a transition from the compact intermediates to the native state being the most abrupt. The three-stage approach to the native state involving heterogeneous folding pathways accords well with theoretical predictions made some time ago using simplified representation of polypeptide chains (42, 43).

Methods

Models. We performed simulations using a variant of the SOP model (44), which accurately predicts the unfolding and refolding of RNA and proteins triggered by *f* (45). In the SOP model, the 76 residues of the Ub monomer are represented using the C_α atoms as the interaction centers. We constructed the 152 residue Ub_2 by covalently linking the last residue GLY of the first domain (46) to the first residue MET of the second domain. The total effective energy in the SOP representation as a function of coordinates $\{r_i\}(i = 1,2,\cdots,N = 152)$ of the interaction centers is mainly composed of three terms: The first term V_{FENE} accounts for the finite extensible nonlinear elastic (FENE) potential that describes the chain connectivity, where $r_{i,i+1}$ is the distance between two neighboring C_α atoms and $r^o_{i,i+1}$ is the corresponding value in the crystal structure. The parameters k and R_0 in V_{FENE} (Eq. **1**) are 20.0 kcal/mol · Å and 0.2 nm. The second term in Eq. **1** is the Lennard–Jones potential, which describes the attractive interaction of native contacts that exist if the distance, r_{ij}, between two nonbonded C_α atoms is less than a cutoff distance $R_c = 0.8$ nm. If r_{ij} between the *i*th and *j*th atoms separated by more than two beads is less than R_c ($r_{ij} < R_c$), then Δ_{ij} is equal to 1 and their interaction is defined as being native. If $r_{ij} > R_c$, then $\Delta_{ij} = 0$. Consequently, their interaction is nonnative and assumed to be represented as a repulsive potential with the sixth power (see the third term of Eq. **1**). In order to prevent unphysical interchain crossing, a specific constraint potential is imposed on the bond angle between the three beads of i, $i+1$, and $i+2$ with the parameter σ whose value is 0.38 nm. In the ubiquitin dimer, a value of 0.2 nm was used for σ for the angles involving beads from 71 to 75 in the hinge loop to account for the increased flexibility of the Gly-rich loop. The values of ϵ_h and ϵ_l are 1.4 kcal/mol and 1.0 kcal/mol, respectively.

$$\begin{aligned} V_T &= V_{\text{FENE}} + V_{\text{NB}}^{\text{ATT}} + V_{\text{NB}}^{\text{REP}} \\ &= -\sum_{i=1}^{N-1} \frac{k}{2} R_0^2 \log\left[1 - \frac{(r_{i,i+1} - r^0_{i,i+1})^2}{R_0^2}\right] \\ &\quad + \sum_{i=1}^{N-3} \sum_{j=i+3}^{N} \epsilon_h \left[\left(\frac{r^0_{ij}}{r_{ij}}\right)^{12} - 2\left(\frac{r^0_{ij}}{r_{ij}}\right)^{6}\right] \\ &\quad + \sum_{i=1}^{N-3} \sum_{j=i+3}^{N} \epsilon_l \left(\frac{\sigma}{r_{ij}}\right)^6 (1 - \Delta_{ij}) + \sum_{i=1}^{N-2} \epsilon_l \left(\frac{\sigma_{i,i+2}}{r_{i,i+2}}\right)^6. \end{aligned} \quad [1]$$

Topologies of Domain-Swapped Structures. The crystal structure of the Ub monomer [Protein Data Bank (PDB) ID code 1UBQ] (26) has five beta sheets from β1 to β5 and two helixes α1 and α2. With the definition of the cutoff for native contacts in the SOP model, there are 177 native contacts in a monomer. The DS structures for Ub_2 are shown in Fig. 1*A*, where the blue and red denote the first and the second domain, respectively. In the topology of the nonswapped Ub_2, 354 native contacts exist only within each domain. We considered five DS structures (displayed in Fig. 1 *B–F*), for which one or several β-strands are swapped between the two domains. In the DS structures, native contacts of the substituted motif (e.g., the blue β1 in case 1) of the first domain are replaced by native contacts of a motif in the second domain (e.g., β1 in gray) of the second domain.

Simulations. Simulations were performed at a temperature of $T = 300$ K with the force-clamp Brownian dynamics algorithm (47, 48). The characteristic time of overdamped motion is $\tau_H = ((\zeta\epsilon_h h)/(k_B T))\tau_L$, where $\tau_L = 2.77$ ps and the friction coefficient $\zeta = 83.0$ for residues. A time step in Brownian simulation corresponds to 5.4 ps in real time, and therefore the common integration time step of 100 equals 540 μs. Just as in AFM experiments a constant pulling force is applied to the C terminal of Ub_2, whereas the N terminal is kept fixed. We also repeated the simulations by applying *f* to the N terminus instead and fixing the C terminus and show data of the merged results if not otherwise noted. The (un)folding process is monitored in terms of native contacts $Q_i(t) = \sum_{j(|j-i|>2)}^{N} \Theta(R_c - r_{ij}(t))\Delta_{ij}$, where $r_{ij}(t)$ denotes the distance between the *i*th and *j*th beads and $\Delta_{ij} = 1$ for native contacts. In the simulation of domain-swapped structures, the native contacts in the first and second domains are monitored by $Q_1(t)$ and $Q_2(t)$, respectively, $Q_{12}(t)$ counts interdomain native contacts, and Q_T is the total number of native contacts in the dimer. The time dependence of the end-to-end distance, $R(t)$, is a natural variable in experiments and is used to monitor the dynamics of the unfolding and refolding processes.

ACKNOWLEDGMENTS. We are grateful to the National Institutes of Health (Grant GM089685), the National Science Foundation (Grant CHE 09-14033), the Klaus Tschira Foundation, and the German Academic Exchange Service for partial support of this work.

1. Onuchic JN, Wolynes PG (2004) Theory of protein folding. *Curr Opin Struct Biol* 14:70–75.
2. Thirumalai D, Hyeon C (2005) RNA and protein folding: Common themes and variations. *Biochemistry* 44:4957–4970.
3. Shakhnovich E (2006) Protein folding thermodynamics and dynamics: Where physics, chemistry, and biology meet. *Chem Rev* 106:1559–1588.
4. Dill KA, Ozkan B, Shell MS, Weikl TR (2008) The protein folding problem. *Annu Rev Biophys* 37:289–316.
5. Schuler B, Eaton WA (2008) Protein folding studied by single-molecule FRET. *Curr Opin Struct Biol* 18:16–26.
6. Thirumalai D, O'Brien EP, Morrison G, Hyeon C (2010) Theoretical perspectives on protein folding. *Annu Rev Biophys* 39:159–183.
7. Clementi C, Jennings PA, Onuchic JN (2000) How native-state topology affects the folding of dihydrofolate reductase and interleukin-1 beta. *Proc Natl Acad Sci USA* 97:5871–5876.
8. Hyeon CB, Thirumalai D (2003) Can energy landscape roughness of proteins and RNA be measured by using mechanical unfolding experiments. *Proc Natl Acad Sci USA* 100:10249–10253.
9. Borgia A, Williams PM, Clarke J (2008) Single-molecule studies of protein folding. *Ann Rev Biophys* 77:101–125.
10. Schlierf M, Li H, Fernández JM (2004) The unfolding kinetics of ubiquitin captured with single-molecule force-clamp techniques. *Proc Natl Acad Sci USA* 101:7299–7304.
11. Fernández JM, Li H (2004) Force-clamp spectroscopy monitors the folding trajectory of a single protein. *Science* 303:1674–1678.
12. Garcia-Manyes S, Brujić J, Badilla CL, Fernández JM (2007) Force-clamp spectroscopy of single-protein monomers reveals the individual unfolding and folding pathways of I27 and ubiquitin. *Biophys J* 93:2436–2446.
13. Walther KA, et al. (2007) Signatures of hydrophobic collapse in extended proteins captured with force spectroscopy. *Proc Natl Acad Sci USA* 104:7916–7921.
14. Zhang J, Qin M, Wang W (2005) Multiple folding mechanisms of protein ubiquitin. *Proteins* 59:565–579.
15. Li MS, Hu CK, Klimov DK, Thirumalai D (2006) Multiple stepwise refolding of immunoglobulin domain I27 upon force quench depends on initial conditions. *Proc Natl Acad Sci USA* 103:93–98.
16. Gräter F, Grubmüller H (2007) Fluctuations of primary ubiquitin folding intermediates in a force clamp. *J Struct Biol* 157:557–569.
17. Alonso DOV, Daggett V (1998) Molecular dynamics simulations of hydrophobic collapse of ubiquitin. *Protein Sci* 7:860–874.
18. Klimov DK, Thirumalai D (1999) Stretching single-domain proteins: Phase diagram and kinetics of force-induced unfolding. *Proc Natl Acad Sci USA* 96:6166–6170.
19. Larios E, Li JS, 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.

20. Marianayagam NJ, Jackson SE (2004) The folding pathway of ubiquitin from all-atom molecular dynamics simulations. *Biophys Chem* 111:159–171.
21. Best RB, Paci E, Hummer G, Dudko OK (2008) Pulling direction as a reaction coordinate for the mechanical unfolding of single molecules. *J Phys Chem B* 112:5968–5976.
22. Barsegov V, Thirumalai D (2005) Probing protein-protein interactions by dynamics force correlation spectroscopy. *Phys Rev Lett* 95:168302.
23. Hyeon C, Morrsion G, Pincus D, Thirumalai D (2009) Refolding dynamics of stretched biopolymers upon force quench. *Proc Natl Acad Sci USA* 106:20288–20293.
24. Garcia-Manyes S, Dougan L, Badilla CL, Brujić J, Fernández JM (2009) Direct observation of an ensemble of stable collapsed states in the mechanical folding of ubiquitin. *Proc Natl Acad Sci USA* 106:10534–10539.
25. Camacho CJ, Thirumalai D (1993) Minimum energy compact structures of random sequences of heteropolymers. *Phys Rev Lett* 71:2505–2508.
26. Vijay-Kumar S, Bugg CE, Cook WJ (1987) Structure of ubiquitin refined at 1.8 Å resolution. *J Mol Biol* 194:531–544.
27. Bennett MJ, Choe S, Eisenberg D (1994) Domain swapping: Entangling alliances between proteins. *Proc Natl Acad Sci USA* 91:3127–3131.
28. Rousseau F, Schymkowitz JWH, Itzhaki LS (2003) The unfolding story of three-dimensional domain swapping. *Structure* 11:243–251.
29. Bennett MJ, Schlunegger MP, Eisenberg D (1995) 3D domain swapping: A mechanism for oligomer assembly. *Protein Sci* 4:2455–2468.
30. Bennett MJ, Sawaya MR, Eisenberg D (2006) Deposition diseases and 3D domain swapping. *Structure* 14:811–824.
31. Klimov DK, Thirumalai D (2000) Native topology determines force-induced unfolding pathways in globular proteins. *Proc Natl Acad Sci USA* 97:7254–7259.
32. Chyan C-L, et al. (2004) Reversible mechanical unfolding of single ubiquitin molecules. *Biophys J* 87:3995–4006.
33. Li MS, Kouza M, Hu CK (2007) Refolding upon force quench and pathways of mechanical and thermal unfolding of ubiquitin. *Biophys J* 92:547–561.
34. Kouza M, Hu C-K, Li MS (2008) New force replica exchange method and protein folding pathways probed by force-clamp technique. *J Chem Phys* 128:045103.
35. Carrion-Vazquez M, et al. (2003) The mechanical stability of ubiquitin is linkage dependent. *Nat Struct Biol* 10:738–743.
36. Dudko OK, Hummer G, Szabo A (2006) Intrinsic rates and activation free energies from single-molecule pulling experiments. *Phys Rev Lett* 96:108101–108104.
37. Garg A (1995) Escape-field distribution for escape from a metastable potential well subject to a steadily increasing bias field. *Phys Rev B* 51:15592–15595.
38. Plaxco KW, Simons KT, Baker D (1998) Contact order, transition state placement and the refolding rates of single domain proteins. *J Mol Biol* 277:985–994.
39. Irbäck A, Mitternacht S, Mohanty S (2005) Dissecting the mechanical unfolding of ubiquitin. *Proc Natl Acad Sci USA* 102:13427–13432.
40. Khorasanizadel S, Peters ID, Roder H (1996) Evidence for a three-state model of protein folding from kinetic analysis of ubiquitin variants with altered core residues. *Nat Struct Biol* 3:193–205.
41. Sosnick TR (2004) Comment on force-clamp spectroscopy monitors the folding trajectory of a single protein. *Science* 306:411.
42. Camacho CJ, Thirumalai D (1993) Kinetics and thermodynamics of folding in model proteins. *Proc Natl Acad Sci USA* 90:6369–6372.
43. Guo Z, Thirumalai D (1996) Kinetics and thermodynamics of folding of a de Novo designed four-helix bundle protein. *J Mol Biol* 263:323–343.
44. Hyeon C, Dima RI, Thirumalai D (2006) Pathways and kinetic barriers in mechanical unfolding and refolding of RNA and proteins. *Structure* 14:1633–1645.
45. Mickler M, et al. (2007) Revealing the bifurcation in the unfolding pathways of GFP by using single-molecule experiments and simulations. *Proc Natl Acad Sci USA* 104:20268–20273.
46. Wiborg O, et al. (1985) The human ubiquitin multigene family: Some genes contain multiple directly repeated ubiquitin coding sequences. *EMBO J* 4:755–759.
47. Ermack DL, McCammon JA (1978) Brownian dynamics with hydrodynamic interactions. *J Chem Phys* 69:1352–1369.
48. Klimov DK, Betancourt MR, Thirumalai D (1998) Virtual atom representation of hydrogen bonds in minimal off-lattice models of alpha helices: Effect on stability, cooperativity and kinetics. *Fold Des* 3:481–496.