The Mechanical Unfolding of Ubiquitin through All-Atom Monte Carlo Simulation with a Gō-Type Potential

Ariel Kleiner and Eugene Shakhnovich
Department of Chemistry and Chemical Biology, Harvard University, Cambridge, Massachusetts

ABSTRACT The mechanical unfolding of proteins under a stretching force has an important role in living systems and is a logical extension of the more general protein folding problem. Recent advances in experimental methodology have allowed the stretching of single molecules, thus rendering this process ripe for computational study. We use all-atom Monte Carlo simulation with a Gō-type potential to study the mechanical unfolding pathway of ubiquitin. A detailed, robust, well-defined pathway is found, confirming existing results in this vein though using a different model. Additionally, we identify the protein's fundamental stabilizing secondary structure interactions in the presence of a stretching force and show that this fundamental stabilizing role does not persist in the absence of mechanical stress. The apparent success of simulation methods in studying ubiquitin's mechanical unfolding pathway indicates their potential usefulness for future study of the stretching of other proteins and the relationship between protein structure and the response to mechanical deformation.

INTRODUCTION

The mechanical deformation of proteins is an important process both in living systems and in the study of protein structure. As a primary molecular component of cellular systems, proteins have a fundamental role in such phenomena as muscle elasticity and cell adhesion. In each of these cases, polypeptides unfold under a stretching force applied at specific amino acids and subsequently refold. Thus, by studying the kinetics of this process and its relationship to protein structure, we can gain valuable insight into the mechanisms by which these biological systems function. Furthermore, theoretical and computational work in this vein holds the promise of providing contributions, such as elucidation of polypeptide free-energy landscapes (1) and unfolding rates (2), to the larger body of work on the protein folding problem.

Ubiquitin was selected for study in this context as a result of its small size and the substantial body of experimental work that has focused on it. This 76-residue protein (see Fig. 1 and Table 1) is primarily involved in marking other proteins that have been targeted for degradation within a cell. Its thermal, chemical, and pressure-induced unfolding have been investigated in detail; in particular, it has been shown that ubiquitin unfolds thermally at a temperature of 83°C (3,4). Ubiquitin's folding transition state ensemble has been described and has been shown to contain conformations with a common native-like core (5). Additionally, a ''native-like intermediate'' has been observed during the molecule's low-temperature refolding, though the structure of the intermediate was not ascertained (6).

Despite the fact that ubiquitin is not naturally subjected to mechanical stress, there has recently been an increasing amount of experimental work focusing on its unfolding under stretching force. In particular, atomic force microscopy has allowed the probing of this phenomenon at the single-molecule scale. Earlier work in this vein—which in general considered proteins such as titin and tenascin that, unlike ubiquitin, are biologically subjected to stretching forces—yielded only unfolding trajectories created at constant velocity (i.e., the molecule's end-to-end distance changed at a constant rate), with the stretching force changing accordingly throughout the process (7–14). Although the results thereby produced are valuable in providing insight into the mechanical properties of polypeptides, they present difficulties at the interface with simulation. In the case of Monte Carlo simulation, the notion of an unfolding velocity as a rate of change with respect to time is not well-defined. Thus, because the aforementioned experiments were performed at constant unfolding velocity with variable stretching force, the specification of a potential that will generate the appropriate behavior is not straightforward.

Fortunately, advances in experimental methodology have in recent years allowed the execution of single-molecule mechanical unfolding under constant force. These techniques were of course first applied to proteins such as titin that had already been the focus of study (15). However, experimental work focused on the constant-force mechanical unfolding of ubiquitin has also emerged (16,17). These experiments measure the end-to-end distance of single chains of ubiquitin domains subjected to constant stretching forces. Although the minimum force at which ubiquitin unfolds has not been precisely determined, it has been shown to lie between ~50 pN and 200 pN (17). When the polypeptide is subjected to forces within this range, the unfolding event for a single ubiquitin domain is generally marked by a sharp two-state step in end-to-end distance of 20.3 ± 0.9 nm. However, for ~5% of 821 observed events, three-state unfolding occurred, with steps of size 8.1 ± 0.7 nm and 12.4 ± 1.0 nm (17).

Submitted January 17, 2006, and accepted for publication November 3, 2006.

Address reprint requests to Eugene Shakhnovich, E-mail: eugene@belok.harvard.edu.

0006-3495/07/03/2054/08 $2.00

doi: 10.1529/biophysj.106.081257

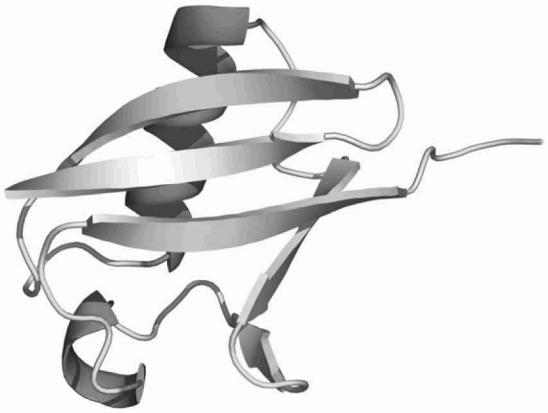

FIGURE 1 Structure of ubiquitin.

Such experimental work has provided a basis for the development of models and accompanying simulations of the mechanical unfolding process. Molecular dynamics has been applied in this context using models and yielding results having varying degrees of detail (8,18–21). To our knowledge, molecular dynamics has not yet been used to study the mechanical unfolding pathway of ubiquitin. Monte Carlo simulation has also been employed in the investigation of mechanical unfolding, though in general to test relatively coarse kinetic models (2,17,22). In particular, Szymczak and Cieplak have recently used a coarse Gō-like model to examine the statistics of the mechanical unfolding times of ubiquitin and integrin (23). Furthermore, Paci et al. have made a number of contributions to the investigation of the mechanical unfolding of a number of different proteins using molecular dynamics, Gō-type models, and other methods, in general to the end of examining molecules' energetic and mechanical properties such as resistance to force (24–28). More in the vein of this article, recent work by Irbäck et al. has examined the unfolding of ubiquitin under a stretching force using all-atom Monte Carlo simulation with a sequence-based potential (29). That work predicts an unfolding pathway for the protein, and so our results may be considered complementary.

We have used all-atom Monte Carlo simulation with a Gō-type potential to elucidate in detail the mechanical unfolding pathway of ubiquitin. The computationally tractable nature of this simulation method has allowed us to obtain robust results that hold at different forces across a large number of simulations. Our model gives good qualitative agreement with experimental results for a number of measurable aspects of the process and provides an explanation of the underlying manner in which it may occur. Additionally, our proposed unfolding pathway agrees with that proposed by Irbäck et al. (29) despite our use of a fundamentally different potential, thereby further confirming the robustness of these results. In contrast to such previous studies of the ubiquitin unfolding pathway, we have used a Gō-type model that guarantees that the ground state of the model is in fact the protein's native state, thus rendering our simulations a truly equilibrium study of the adiabatic unfolding pathway.

TABLE 1 Secondary structure elements of ubiquitin

Label	Residue range
β_1	1–7
β_2	10–17
α_1	23–34
β_3	40–45
β_4	48–50
α_2	56–59
β_5	64–72

Labels of the form α_i denote α-helices, and labels of the form β_i denote strands of the β-sheet.

METHODS

We followed the work of Shimada et al. (30,31) in constructing our simulations. All nonhydrogen atoms in the residue chain were represented as hard spheres having radii proportional to their van der Waals radii; a proportionality factor of 0.75 was used. No hydrogen atoms or solvent (e.g., water) molecules were explicitly included in the model. At each Monte Carlo step, one backbone and ten side-chain moves were performed. Only local backbone moves and torsional side-chain rotations were permitted, and moves were accepted or rejected based on the Metropolis rule (through which the simulation "temperature" T entered). Each backbone move consisted of a small random rotation of the ϕ-ψ angles of up to three nearby residues. Thus, although these moves were not strictly enforced to be wholly local, the majority of them resulted in largely local changes. Please see Shimada et al. (30) for more details regarding the structure of the Monte Carlo simulations. Because we only studied unfolding, all simulations were started with ubiquitin in its native state as specified by the structure given by Protein Data Bank code 1UBQ (32).

The potential used has the form

$$E = E_{G\bar{o}} + E_H + E_F,$$

where $E_{G\bar{o}}$ is a Gō-type (33) square well contribution, E_H is a backbone hydrogen bonding contribution, and E_F is the force term. For any atom a, let $\#(a)$ denote the index of the residue of which a is a member. Thus, with $k \geq 0$ a constant,

$$E_{G\bar{o}} = \sum_{\substack{\text{allatoms } a,b \\ a \neq b \\ |\#(a) - \#(b)| \geq k}} \epsilon(a,b),$$

where, for any pair of atoms a, b with hard core distance σ_{ab} and interatom distance r,

$$\epsilon(a,b) = \begin{cases} \infty & r < \sigma_{ab} \\ \delta(i,j) & \sigma_{ab} \leq r < \lambda\sigma_{ab} \\ 0 & r \geq \lambda\sigma_{ab} \end{cases},$$

with $\delta(i,j) = -1$ if i,j are in contact (i.e., $\sigma_{ab} \leq r < \lambda\sigma_{ab}$) in the native state and $\delta(i,j) = 1$ otherwise. The hard core distance σ_{ab} is simply the sum of the aforementioned hard sphere radii for atoms a and b, and the value λ is a constant. We use $\lambda = 1.8$ in accordance with the findings of Shimada et al. (30). The backbone hydrogen-bonding contribution to the potential was included primarily to allow tuning of secondary structure stability and has the form

$$E_H = hN_{NO},$$

where h is a constant and N_{NO} is the number of distinct pairs of nitrogen and oxygen atoms in the backbone that are in contact in a given conformation. Finally, the force term is given by

$$E_F = -\vec{F} \cdot \vec{r},$$

where $\vec{r}$ is a vector extending from the N-terminal nitrogen atom to the C-terminal carbon atom of the backbone. The vector $\vec{F}$ is the Monte Carlo stretching force; in our simulations, $\vec{F}$ remains constant. Note that these definitions imply that the stretching force is applied at the terminal amino acids of the protein, in accordance with experimental methodology (16,17).

To complete the energy model's specification, it is necessary to calibrate it to the observed thermodynamics of ubiquitin (in the absence of a stretching force) by selecting values for the constants k and h such that the appropriate relative stabilities of the entire protein and its constituent secondary structures across the temperature range of interest are achieved. Let T_c be the Monte Carlo temperature at which ubiquitin transitions thermally from its folded to its unfolded state. Thus, T_c corresponds to a physical temperature of 83°C = 356 K (3). To allow comparison with experimental results, we performed our simulations of mechanical unfolding at the Monte Carlo temperature T_o corresponding to a physical temperature of 21°C = 294 K. Note that T_o is therefore determined by T_c as $T_o = (294/356)T_c$, and we must ensure that our model exhibits the appropriate thermodynamic behavior for $T \sim T_o$.

As predicted by AGADIR (34), α_1 (see Table 1 for definitions of secondary structure labels) in isolation (i.e., a free helix with the residue sequence of α_1) has helical propensity <1.2% at all temperatures between 273 and 373 K (and pH = 7). Although α_2 is predicted to have no helical propensity at any of these temperatures, we will not explicitly consider this fact because the extremely small size of α_2 causes it to always have stability significantly less than that of α_1. To ensure that an isolated α_1 is not stable at T_o, simulations of length 50×10^6 steps were run for a variety of values of k, h, and T to screen across values of T and determine T_c and T_{α_1}, the location of the thermal unfolding transition of the isolated α_1 helix. Selected results of these screens are shown in Table 2.

It is clear that α_1 is vastly overstabilized when $k = 2$; this relatively low value of k allows the plethora of local interactions in the helix to dominate its energetics. As expected, this effect is significantly mitigated, though not eliminated, when k is increased to 3. Additionally, increasing k from 2 to 3 does not significantly destabilize the remainder of the protein, because the existence and contribution of local (i.e., $i - (i + 2)$ in this case) interactions outside of the helices in the native state are minimal. However, even with $k = 3$ (and $h = 0$), $T_{\alpha_1} > T_o$. As a result, we introduce $h > 0$ to further destabilize α_1 relative to the remainder of the protein by lessening the stabilizing energetic contribution of the $i - (i + 4)$ hydrogen bonds of the α-helix. Although $h = 1.5$ yields a significant decrease in helix stability, inspection of the resulting thermal unfolding simulation data indicates that this relatively high value of h causes the remainder of the molecule to become excessively unstable. It is worth noting that a similar problem was observed for $k \geq 4$, even with $h = 0$. Because $h = 0.5$ yields $T_{\alpha_1} < T_o$ and increasing h to 0.75 or 1.0 does not give decreased helix stability, we use $h = 0.5$. As a result, the parameter values selected are

$$k = 3 \quad h = 0.5,$$

and we have thus now fully specified the potential function. Fig. 2 shows the results of simulations performed with these parameter values throughout a range of temperatures in the absence of a stretching force. As seen from these data as well as Table 2, $T_c = 2.6$ and so the temperature at which all mechanical unfolding simulations were performed is

$$T_o = 2.1.$$

Simple physical considerations indicate that, when subjected to a stretching force applied at its terminal residues, a protein first rotates to align the axis defined by its terminal residues with the direction of the applied force. However, such an initial global rotation cannot easily occur for an arbitrarily directed stretching force in our simulations because only local backbone moves are permitted. Thus, all mechanical unfolding simulations were performed with the Monte Carlo stretching force $\vec{F}$ applied in the direction of the vector $\vec{r}_o$ extending from the N-terminal nitrogen atom to the C-terminal carbon atom of the polypeptide's backbone in the native state. Given this directionality, the stretching force in any simulation can now be specified merely by giving its magnitude $F = |\vec{F}|$.

To ascertain the force magnitude F_c at which ubiquitin transitions from its folded to its unfolded state at temperature T_o, simulations were run on the molecule for 50×10^6 steps with stretching force magnitudes throughout a range of values. The results of this screen across force magnitudes are shown in Fig. 3 and indicate that $13 < F_c < 14$; although F_c has not yet been precisely determined experimentally in physical units, it has been shown approximately that 50 pN $< F_c <$ 200 pN (17). Therefore, to examine the mechanical unfolding process, large collections of independent simulations were run at force magnitudes of $F = 16$ and $F = 26$. Each simulation was allowed to run until the polypeptide unfolded and subsequently for an additional 5×10^6 steps.

TABLE 2 Monte Carlo thermal transition temperatures of isolated α_1 and ubiquitin for selected values of k and h

k	h	T_{α_1}	T_c	T_o
2	0	3.5	3.1	2.6
3	0	2.7	2.7	2.2
3	0.5	1.9	2.6	2.1
3	0.75	1.9	2.5	2.1
3	1.0	1.9	2.5	2.1
3	1.5	<1.6	2.3	1.9

As seen in the text, $T_o = (294/356)\ T_c$ is the Monte Carlo temperature corresponding to 294 K. All temperature values have uncertainties of ±0.1.

RESULTS

Fig. 4 shows a representative mechanical unfolding trajectory for ubiquitin. Qualitatively, we obtain good agreement with experimentally derived trajectories (17), with the transition from the folded to the unfolded state appearing

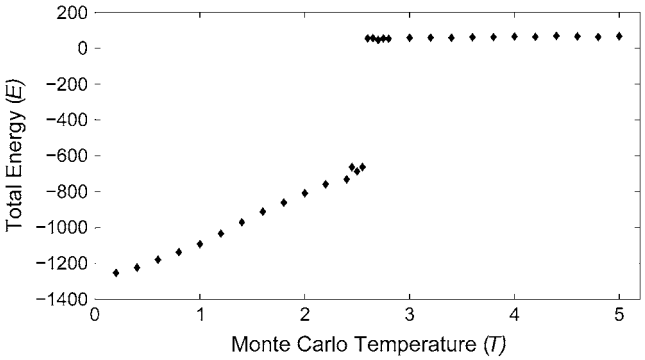

FIGURE 2 The results of simulations of ubiquitin performed at various Monte Carlo temperatures with $k = 2$ and $h = 0.5$ in the absence of a stretching force. Each point gives the average energy of the molecule during the final 10×10^6 steps of a single 50×10^6 step simulation. The energy of the native state of ubiquitin in these simulations is −1253.5.

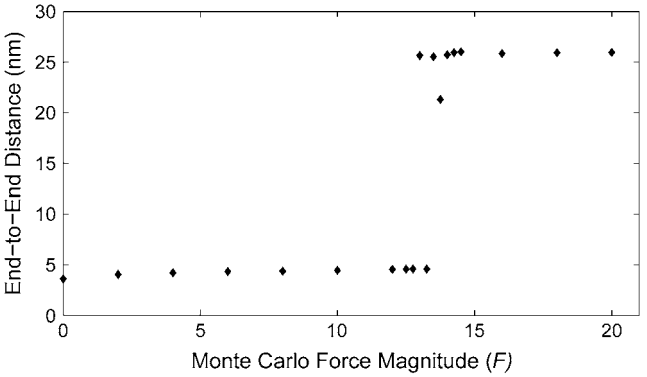

FIGURE 3 The results of simulations of ubiquitin performed at various stretching force magnitudes with $T = T_o = 2.1$. Each point gives the average end-to-end distance ($|\vec{r}|$) of the molecule during the final 10×10^6 steps of a single 50×10^6 step simulation. The end-to-end distance of the native state of ubiquitin is $|\vec{r}_o| = 3.847$ nm.

as a sharp ~20 nm step. However, closer inspection (see Fig. 4 *b*) reveals a subtle intermediate plateau in the value of $|\vec{r}|$ that occurs during the primary unfolding event. Its existence suggests the possibility of an unstable intermediate on the mechanical unfolding pathway. Although the majority

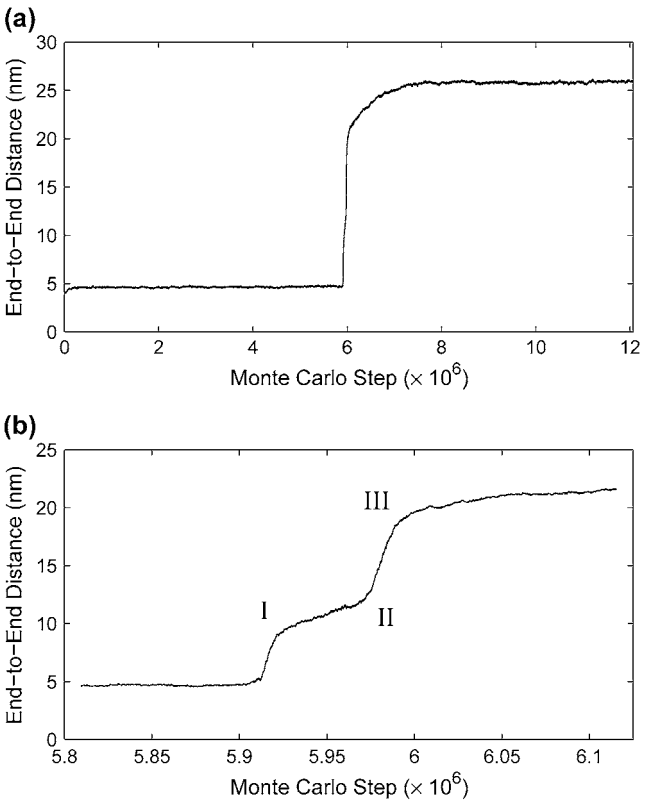

FIGURE 4 (*a*) A typical simulated mechanical unfolding trajectory for ubiquitin with $F = 16$. The protein is fully unfolded when its end-to-end distance ($|\vec{r}|$) becomes >25 nm. (*b*) A closer look at the major unfolding step. The structures corresponding to points I, II, and III are shown in Fig. 5.

of experimentally derived mechanical unfolding trajectories are similar to that seen in Fig. 4 *a*, three-state unfolding as suggested by Fig. 4 *b* is observed for ~5% of trajectories (17).

In describing the mechanical unfolding pathway of ubiquitin, we focus on the interactions among secondary structure elements. Fig. 6 shows the secondary structure unfolding trajectory corresponding to the trajectory shown in Fig. 4. Unsurprisingly, comparison of the two figures indicates that the initiation of the molecule's unfolding coincides with the first secondary structure unfolding event (in this case, β_1 separates from α_1). More significantly, the unfolding of secondary structure involving the β-strands is consistently characterized by sharp transitions from the presence of a finite number of native-state contacts to the absence of such contacts, with few fluctuations thereafter. As a result, we define the step in a simulation during which a given secondary structure interaction (e.g., α_1–β_1) unfolds to be the step after which the fraction of native-state contacts intact within that interaction remains beneath 0.1 for the remainder of the simulation. Unfolding trajectories for α_1 and α_2 are not included in Fig. 6 because these structures are consistently the final secondary structure elements to unfold. Their unraveling occurs more gradually in comparison to the aforementioned sharper transitions characterizing interactions involving the β-strands. In general, when $|\vec{r}| \approx 20$ nm for ubiquitin (as in Fig. 4), the entire molecule with the exception of the α-helices has unfolded; a subsequent more gradual increase of $|\vec{r}|$ from 20 to 25 nm consists of the subsequent unfolding of α_1 and α_2.

To systematically determine the events on the mechanical unfolding pathway of ubiquitin, an ensemble of 113 simulations were run with $F = 16$. Table 3 shows the four most common sequences of unfolding events that were observed, accounting for 84 of the 113 runs (74%). The remaining runs exhibit very similar pathways, generally differing from those shown only in the relative positions of two to three events. However, a number of results hold for the event sequences of all or nearly all runs. Most importantly, the pathway can

TABLE 3 The four most common sequences in which secondary structure interactions were observed to unfold in an ensemble of mechanical unfolding simulations with $F = 16$

53/113 Runs	18/113 Runs	7/113 Runs	6/113 Runs
α_1–β_1	β_1–β_5	α_1–β_1	α_1–β_1
β_1–β_5	α_1–β_1	β_1–β_5	β_1–β_5
α_1–β_2	α_1–β_2	α_1–β_2	β_1–β_2
β_1–β_2	β_1–β_2	α_1–β_5	α_1–β_2
α_1–β_5	α_1–β_5	β_1–β_2	α_1–β_5
β_3–β_5	β_3–β_5	β_3–β_5	β_3–β_5
α_1–β_4	α_1–β_4	α_1–β_4	α_1–β_4
α_1–β_3	α_1–β_3	α_1–β_3	α_1–β_3
β_3–β_4	β_3–β_4	β_3–β_4	β_3–β_4
α_2–α_2	α_2–α_2	α_2–α_2	α_2–α_2
α_1–α_1	α_1–α_1	α_1–α_1	α_1–α_1

be divided into three blocks of unfolding events that do not comingle in 112 of the 113 simulations (in the single anomalous simulation, the positions of the β_3–β_4 and α_2–α_2 unfolding events are simply exchanged):

$$\left.\begin{array}{l}\alpha_1-\beta_1\\ \beta_1-\beta_5\\ \alpha_1-\beta_2\\ \beta_1-\beta_2\\ \alpha_1-\beta_5\\ \beta_3-\beta_5\end{array}\right\}\text{block 1}$$
$$\left.\begin{array}{l}\alpha_1-\beta_4\\ \alpha_1-\beta_3\\ \beta_3-\beta_4\end{array}\right\}\text{block 2}$$
$$\left.\begin{array}{l}\alpha_2-\alpha_2\\ \alpha_1-\alpha_1\end{array}\right\}\text{block 3}$$

The events in block 2 are always observed to occur in the order shown. In block 3, α_2 unfolds before α_1 in all but 4/113 simulations. Furthermore, the final unfolding event of block 1 is β_3–β_5 in 112/113 simulations. Thus, the vast majority of the variability in the sequence of unfolding events arises in the remainder of block 1, as seen in Table 3.

Therefore, the mechanical unfolding pathway of ubiquitin is in general as follows, with only minor deviations:

1. β_1 and β_2 separate from α_1, β_5, and each other. β_5 may separate from α_1 either concurrently or immediately thereafter, while remaining in contact with β_3. The end of this step in the pathway marks the beginning of the plateau seen in Fig. 4 (point I); the resulting structure is given by structure I in Fig. 5.
2. β_5 separates from β_3. The completion of this separation yields the end of the plateau seen in Fig. 4 (point II); the resulting structure is given by structure II in Fig. 5.
3. β_3 and β_4 separate from α_1 and only subsequently separate from each other. The completion of this step occurs at point III in Fig. 4, with the resulting structure shown in Fig. 5.
4. α_1 and α_2 unfold. The completion of this final step yields the complete unfolding of the protein and is indicated by $|\vec{r}|$ becoming greater than ~25 nm.

These results and the form of ubiquitin's mechanical unfolding trajectory suggest that structure I is an unstable intermediate in the stretching process; we also expect that structure II is similarly unstable. To validate these expectations, we created two structures corresponding to structures I and II by eliminating secondary structure elements from the original molecule. The first, which we name UBQiA, corresponds to structure I, consists of ubiquitin's residues 21–76, and was formed by removing β_1 and β_2 from the original structure of ubiquitin. The second, UBQiB, corresponds to structure II and was formed by also eliminating β_5, leaving only residues 21–63 of the original ubiquitin structure. Simulations with parameters identical to those of the simulations discussed above (i.e., $T = T_o = 2.1$) were performed throughout a range of forces for each structure; the results are shown in Fig. 7. As expected, both structures are unstable at low forces ($F < 4$ for UBQiA and $F < 2$ for UBQiB). Furthermore, UBQiB is less stable under a stretching force than UBQiA, as expected from the roles of structures I and II in ubiquitin's mechanical unfolding pathway: the formation of structure I marks the beginning of the plateau in the value of $|\vec{r}|$, and the formation of structure II marks its end.

The simulations and analysis discussed above were performed with $F = 16$. We also performed an identical investigation with $F = 26$, obtaining the same results for the form of ubiquitin's mechanical unfolding pathway. These results obtained at a higher stretching force further illustrate the robustness of our findings.

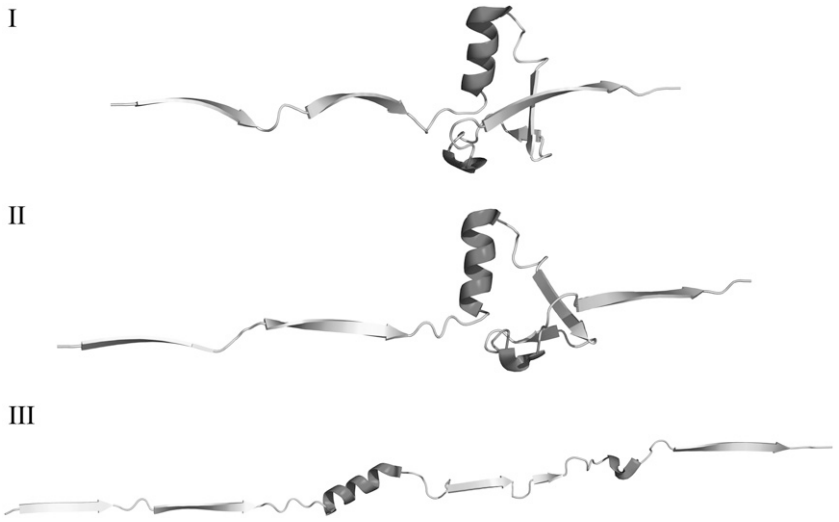

FIGURE 5 The structures corresponding to points I, II, and III in the mechanical unfolding trajectory shown in Fig. 4.

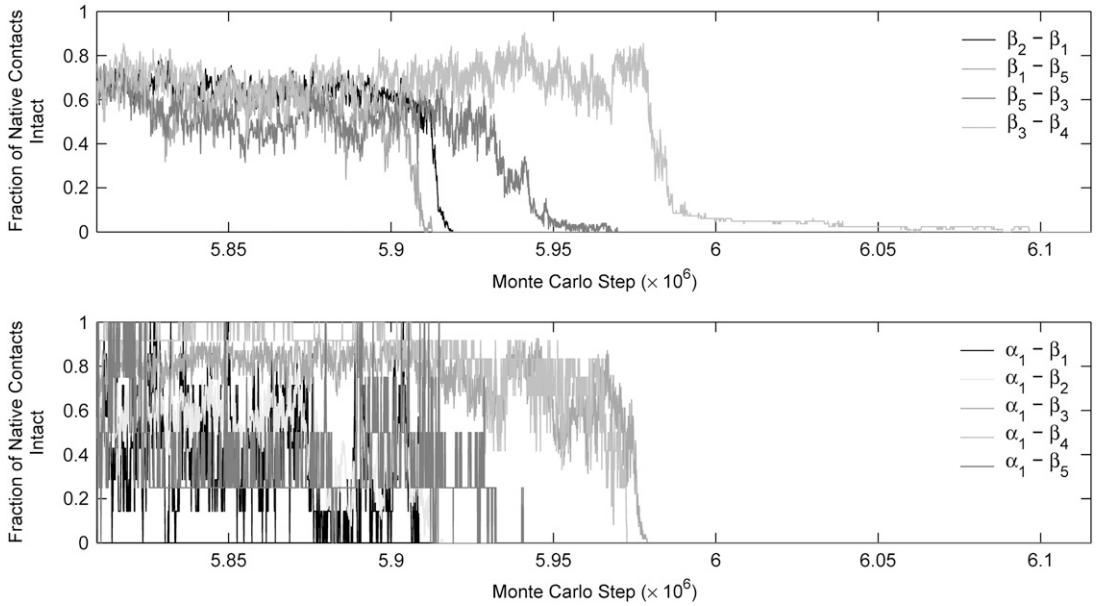

FIGURE 6 The secondary structure unfolding trajectory corresponding to the mechanical unfolding trajectory shown in Fig. 4. Labels of the form $\beta_i - \beta_j$ indicate interactions between atoms in strands of the β-sheet, and labels of the form $\alpha_1 - \beta_i$ indicate interactions between α_1 and strands of the β-sheet. This figure shows only the range of interesting Monte Carlo steps surrounding the unfolding transition. Intrahelix interaction trajectories are not shown because the helices remain intact while the remainder of the secondary structure unfolds and only subsequently unfold more gradually themselves.

DISCUSSION

Our results indicate that the mechanical unfolding of ubiquitin is triggered by the separation of β_1, β_2, or β_5 from either α_1 or the remainder of the β-sheet. It follows that the interaction between these β-strands and the other secondary structure of the protein provides a fundamental stabilizing contribution in the presence of a stretching force. This conclusion is further supported by our examination of the structures UBQiA and UBQiB, which are drastically destabilized under a stretching force as a result of their lack of β_1, β_2, and β_5 (in the case of UBQiB). Interestingly, the seemingly fundamental stabilizing energetic contribution of these β-strands to the protein is only essential to its stability in the presence of a stretching force. Simulations of UBQiA and UBQiB at various Monte Carlo temperatures in the absence of a stretching force indicate that both structures remain stable at temperatures up to ~2.2 (see Fig. 8). Although this value is less than $T_c = 2.6$, the difference between these transition temperatures is sufficiently small to allow us to conclude that the removed secondary structure elements and their interactions with the remainder of the protein are not fundamental to the stability of ubiquitin in the absence of a stretching force. Rather, the small amount of

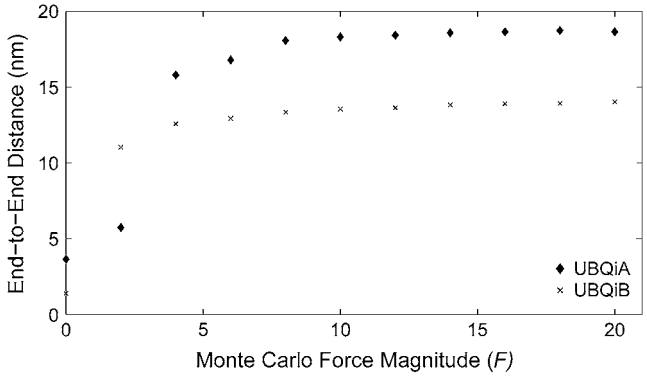

FIGURE 7 The results of simulations of UBQiA and UBQiB performed at various stretching force magnitudes with $T = T_o = 2.1$. Each point gives the average end-to-end distance ($|\vec{r}|$) of a structure during the final 10×10^6 steps of a single 50×10^6 step simulation. The points at $F = 0$ give the end-to-end distances of the structures' native states.

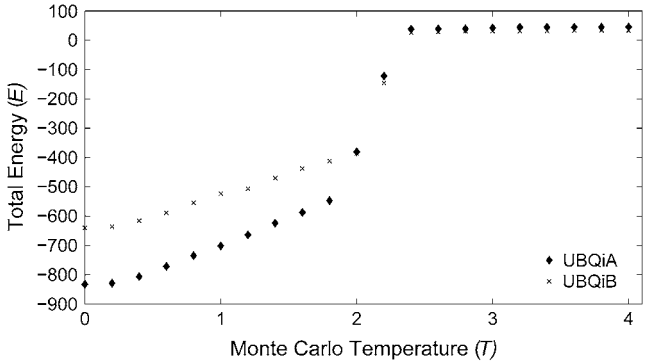

FIGURE 8 The results of simulations of UBQiA and UBQiB performed at various Monte Carlo temperatures in the absence of a stretching force. Each point gives the average energy of a structure during the final 10×10^6 steps of a single 50×10^6 step simulation. The points at $T = 0$ give the energies of the structures' native states.

observed destabilization is presumably largely the result of a decrease in the number of bonds (or contacts, in our case) that must be broken to unfold the structure thermally.

The fact that, in our simulations, the α-helices unfold gradually in comparison to the remainder of ubiquitin's secondary structure is also of interest. This observation is at first glance somewhat surprising given that, as seen above, the helices have negligible helical propensity in isolation within the temperature range of interest in the absence of a stretching force. However, the timescales and simulation lengths characteristic of thermally induced unfolding are significantly greater than those for force-induced unfolding. Furthermore, the structure of the helices and the interactions within them differ greatly from the structure and interactions of the β-sheet. The vast majority of the interactions within the β-sheet and between β-strands and α_1 are nonlocal. Additionally, the number of contacts per atom is on average less for atoms in the β-sheet than for atoms in the α-helices. The helices are in contrast characterized by dense local interactions. These physical considerations provide some insight into the causes of the more gradual unfolding observed for the helices even after the remainder of the protein has unfolded. It is also possible that the helices are to some extent shielded from the applied force by the presence of a significant portion of the amino acid chain between their ends and the terminal residues of the protein. Finally, it must be noted that the unfolding behavior of the helices could simply be an artifact of the model and potential chosen, with the helices resultantly overstabilized.

As seen above, our computationally derived mechanical unfolding trajectories for ubiquitin yield good qualitative agreement with those that are observed experimentally. The unfolding transition is in both cases generally characterized by a sudden ~20 nm step. However, upon examining the simulated transition at higher resolution—which is not currently possible experimentally—we observe a plateau in the value of $|\vec{r}|$ when $|\vec{r}| - |\vec{r}_o| \approx 7$ nm. Analagously, three-state unfolding is seen in ~5% of experimentally derived trajectories (17). In these cases, a marked plateau in the value of $|\vec{r}|$ occurs when $|\vec{r}| - |\vec{r}_o| = 8.1 \pm 0.7$ nm; this value is an average over events observed in a large collection of experiments. Our results in this vein are thus within the realm of quantitative agreement with experiment, thereby providing a potential explanation of the empirically observed three-state unfolding: this phenomenon is observed in the relatively few instances in which an unstable intermediate consisting of structure I transitioning to structure II persists for a detectably long period of time.

In addition to being consistent with experimental observations, our results are complementary to and provide confirmation of existing computational results regarding the mechanical unfolding of ubiquitin. Also using all-atom Monte Carlo simulation but with a sequence-based potential, Irbäck et al. (29) have studied this process. Our work differs fundamentally in that we use a Gō-type rather than a sequence-based potential, thus ensuring that our model's ground state is the protein's native state; as a result, our simulations do in fact constitute an equilibrium study of the unfolding pathway. We have nevertheless obtained similar results regarding the qualitative properties of unfolding trajectories, the location of the plateau in the value of $|\vec{r}|$, and the form of the mechanical unfolding pathway. Therefore, we have strikingly demonstrated the robustness of these findings.

CONCLUSION

Using all-atom Monte Carlo simulation with a Gō-type potential, we have elucidated a detailed and robust mechanical unfolding pathway for ubiquitin. Our results also identify the protein's essential stabilizing interactions under a stretching force and suggest the structure of an unstable intermediate on the pathway. In addition to obtaining good agreement with and a possible explanation for experimental observations, we have confirmed a number of existing results in this vein, though using a fundamentally different energy model.

As applications of all-atom simulation methods such as those presented here to the detailed study of the mechanical unfolding of proteins are only now emerging, rich opportunities for future work exist. In particular, it will be interesting to examine other proteins such as titin, which, unlike ubiquitin, has a role in natural processes that involves mechanical deformation. Consideration of other polypeptides will allow a more thorough assessment of the structural properties that determine the behavior of a protein under a stretching force. Furthermore, the simulations presented here can be extended to probe other interesting features of the process, including the effects of applying the stretching force at nonterminal amino acids and the results of allowing a protein to refold under low force after it has been unfolded at a higher force. These variants of the mechanical unfolding process have already been examined experimentally (8,16) and have bearing on the mechanical unfolding of proteins as it occurs in biological systems.

We thank the Shakhnovich group at Harvard University and, in particular, Lucas Nivon for valuable discussion during the course of this work.

REFERENCES

1. Kirmizialtin, S., L. Huang, and D. E. Makarov. 2005. Topography of the free-energy landscape probed via mechanical unfolding of proteins. *J. Chem. Phys.* 122:234915-1–234915-12.
2. Carrion-Vazquez, M., A. F. Oberhauser, S. B. Fowler, P. E. Marszalek, S. E. Broedel, J. Clarke, and J. M. Fernandez. 1999. Mechanical and chemical unfolding of a single protein: a comparison. *Proc. Natl. Acad. Sci. USA.* 96:3694–3699.
3. Herberhold, H., and R. Winter. 2002. Temperature- and pressure-induced unfolding and refolding of ubiquitin: a static and kinetic Fourier transform infrared spectroscopy study. *Biochemistry.* 41:2396–2401.
4. Ibarra-Molero, B., V. V. Loladze, G. I. Makhatadze, and J. M. Sanchez-Ruiz. 1999. Thermal versus guanidine-induced unfolding of

ubiquitin. An analysis in terms of the contributions from charge–charge interactions to protein stability. *Biochemistry*. 38:8138–8149.

5. Krantz, B. A., R. S. Dothager, and T. R. Sosnick. 2004. Discerning the structure and energy of multiple transition states in protein folding using Ψ-analysis. *J. Mol. Biol.* 337:463–475.
6. Larios, E., J. S. Li, K. Schulten, H. Kihara, and M. Gruebele. 2004. Multiple probes reveal a native-like intermediate during low-temperature refolding of ubiquitin. *J. Mol. Biol.* 340:115–125.
7. Carrion-Vazquez, M., P. E. Marszalek, A. F. Oberhauser, and J. M. Fernandez. 1999. Atomic force microscopy captures length phenotypes in single proteins. *Proc. Natl. Acad. Sci. USA*. 96:11288–11292.
8. Carrion-Vazquez, M., H. Li, H. Lu, P. E. Marszalek, A. F. Oberhauser, and J. M. Fernandez. 2003. The mechanical stability of ubiquitin is linkage dependent. *Nat. Struct. Biol.* 10:738–743.
9. Li, H., A. F. Oberhauser, S. B. Fowler, J. Clarke, and J. M. Fernandez. 2000. Atomic force microscopy reveals the mechanical design of a modular protein. *Proc. Natl. Acad. Sci. USA*. 97:6527–6531.
10. Li, H., W. A. Linke, A. F. Oberhauser, M. Carrion-Vazquez, J. G. Kerkvliet, H. Lu, P. E. Marszalek, and J. M. Fernandez. 2002. Reverse engineering of the giant muscle protein titin. *Nature*. 418:998–1002.
11. Marszalek, P. E., H. Lu, H. Li, M. Carrion-Vazquez, A. F. Oberhauser, K. Schulten, and J. M. Fernandez. 1999. Mechanical unfolding intermediates in titin molecules. *Nature*. 402:100–103.
12. Oberhauser, A. F., P. E. Marszalek, H. P. Erickson, and J. M. Fernandez. 1998. The molecular elasticity of the extracellular matrix protein tenascin. *Nature*. 393:181–185.
13. Rief, M., M. Gautel, F. Oesterhelt, J. M. Fernandez, and H. E. Gaub. 1997. Reversible unfolding of individual titin immunoglobulin domains by AFM. *Science*. 276:1109–1112.
14. Rief, M., M. Gautel, A. Schemmel, and H. E. Gaub. 1998. The mechanical stability of immunoglobulin and fibronectin III domains in the muscle protein titin measured by atomic force microscopy. *Biophys. J.* 75:3008–3014.
15. Oberhauser, A. F., P. K. Hansma, M. Carrion-Vazquez, and J. M. Fernandez. 2001. Stepwise unfolding of titin under force-clamp atomic force microscopy. *Proc. Natl. Acad. Sci. USA*. 98:468–472.
16. Fernandez, J. M., and H. Li. 2004. Force-clamp spectroscopy monitors the folding trajectory of a single protein. *Science*. 303:1674–1678.
17. Schlierf, M., H. Li, and J. M. Fernandez. 2004. The unfolding kinetics of ubiquitin captured with single-molecule force-clamp techniques. *Proc. Natl. Acad. Sci. USA*. 101:7299–7304.
18. Gao, M., H. Lu, and K. Schulten. 2002. Unfolding of titin domains studied by molecular dynamics simulations. *J. Muscle Res. Cell Motil.* 23:513–521.
19. Li, P., and D. E. Makarov. 2003. 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.
20. Li, P., and D. E. Makarov. 2004. Simulation of the mechanical unfolding of ubiquitin: probing different unfolding reaction coordinates by changing the pulling geometry. *J. Chem. Phys*. 121:4826–4832.
21. Lu, H., and K. Schulten. 1999. Steered molecular dynamics simulations of force-induced protein domain unfolding. *Proteins*. 35:453–463.
22. Makarov, D. E., P. K. Hansma, and H. Metiu. 2001. Kinetic Monte Carlo simulation of titin unfolding. *J. Chem. Phys.* 114:9663–9673.
23. Szymczak, P., and M. Cieplak. 2006. Stretching of proteins in a force-clamp. *J. Phys. Condens. Matter*. 18:L21–L28.
24. Brockwell, D. J., G. S. Beddard, E. Paci, D. K. West, P. D. Olmsted, D. A. Smith, and S. E. Radford. 2005. Mechanically unfolding the small, topologically simple protein L. *Biophys. J.* 89:506–519.
25. Brockwell, D. J., E. Paci, R. C. Zinober, G. S. Beddard, P. D. Olmsted, D. A. Smith, R. N. Perham, and S. E. Radford. 2003. Pulling geometry defines the mechanical resistance of a β-sheet protein. *Nat. Struct. Biol.* 10:731–737.
26. Ng, S. P., R. W. S. Rounsevell, A. Steward, C. D. Geierhaas, P. M. Williams, E. Paci, and J. Clarke. 2005. Mechanical unfolding of TNfn3: the unfolding pathway of a fnIII domain probed by protein engineering, AFM, and MD simulation. *J. Mol. Biol.* 350:776–789.
27. West, D. K., D. J. Brockwell, P. D. Olmsted, S. E. Radford, and E. Paci. 2006. Mechanical resistance of proteins explained using simple molecular models. *Biophys. J.* 90:287–297.
28. West, D. K., P. D. Olmsted, and E. Paci. 2006. Mechanical unfolding revisited through a simple but realistic model. *J. Chem. Phys.* 124: 154909.
29. Irbäck, A., S. Mitternacht, and S. Mohanty. 2005. Dissecting the mechanical unfolding of ubiquitin. *Proc. Natl. Acad. Sci. USA*. 102:13427–13432.
30. Shimada, J., E. L. Kussell, and E. I. Shakhnovich. 2001. The folding thermodynamics and kinetics of crambin using an all-atom Monte Carlo simulation. *J. Mol. Biol.* 308:79–95.
31. Shimada, J., and E. I. Shakhnovich. 2002. The ensemble folding kinetics of protein G from an all-atom Monte Carlo simulation. *Proc. Natl. Acad. Sci. USA*. 99:11175–11180.
32. Vijay-Kumar, S., C. E. Bugg, and W. J. Cook. 1987. Structure of ubiquitin refined at 1.8 Å resolution. *J. Mol. Biol.* 194:531–544.
33. Gō, N., and H. Abe. 1981. Noninteracting local-structure model of folding and unfolding transition in globular proteins. I. Formulation. *Biopolymers*. 20:991–1011.
34. Muñoz, V., and L. Serrano. 1994. Elucidating the folding problem of helical peptides using empirical parameters. *Nat. Struct. Biol.* 1:399–409.