Biophysical Journal Volume 93 October 2007 2644–2654

Secondary Structure, Mechanical Stability, and Location of Transition State of Proteins

Mai Suan Li
Institute of Physics, Polish Academy of Sciences, Warsaw, Poland

ABSTRACT It is well known that the unfolding times of proteins, τ_u, scales with the external mechanical force f as $\tau_u = \tau_u^0 \exp(-fx_u/k_B T)$, where x_u is the location of the average transition state along the reaction coordinate given by the end-to-end distance. Using the off-lattice Go-like models, we have shown that in terms of x_u, proteins may be divided into two classes. The first class, which includes β- and β/α-proteins, has $x_u \approx 2$–5 Å whereas the second class of α-proteins has x_u about three times larger than that of the first class, $x_u \approx 7$–15 Å. These results are in good agreement with the experimental data. The secondary structure is found to play the key role in determining the shape of the free energy landscape. Namely, the distance between the native state and the transition state depends on the helix content linearly. It is shown that x_u has a strong correlation with mechanical stability of proteins. Defining the unfolding force, f_u, from the constant velocity pulling measurements as a measure of the mechanical stability, we predict that x_u decays with f_u by a power law, $x_u \sim f_u^{-\mu}$, where the exponent $\mu \approx 0.4$. We have demonstrated that the unfolding force correlates with the helix content of a protein. The contact order, which is a measure of fraction of local contacts, was found to strongly correlate with the mechanical stability and the distance between the transition state and native state. Our study reveals that x_u and f_u might be estimated using either the helicity or the contact order.

INTRODUCTION

Despite numerous advances in recent years (1,2), deciphering the free energy landscape (FEL) of biomolecules remains a challenge in molecular biology. The most detailed information on FEL may be gained from the all-atom simulations but this approach is restricted to rather short peptides and proteins (3,4) due to its high computational expenses. In this situation atomic force microscopy (AFM) or optical tweezers (5) have been proved to be very useful tools to probe the FEL of proteins. If the external mechanical force, f, is not very large, one can assume that it moves the FEL profile but leaves the distance between the native state (NS) and the transition state (TS), x_u, unchanged (Fig. 1). Then the unfolding barrier is reduced by $\Delta\Delta G_{TS-NS} = -fx_u$ and the dependence of unfolding time τ_u on f is given by the following Bell equation (6):

$$\tau_u = \tau_u^0 \exp(-fx_u/k_B T). \tag{1}$$

Here k_B is Boltzmann's constant and T is the temperature. Clearly, one can measure the unfolding time as a function of force to find the location of the TS. In the past decade the Bell approximation (Eq. 1) has been repeatedly refined using various approximations (7–10). However, because this approximation works pretty well for many experiments (11,12); we will follow it.

Accumulated experimental as well as simulation data (13–15) point to the important role of the secondary structure of the native topology in the mechanical resistance of folded biomolecules. The β- and α/β-proteins, e.g., can withstand higher force compared to the α-helix proteins (16). However, the effect the secondary structure on the distance between the NS and TS has not been studied systematically. It is believed (J. M. Fernandez, personal communication, 2006) that x_u is $\sim$3 Å, which is probably a quasiuniversal length scale of unfolding process. To check this point we have collected the experimental values of x_u for all 12 proteins studied so far (Table 1). The largest departure from the common tendency is the α-helix spectrin, which has considerably larger value of $x_u \approx 15$ Å (17). This result suggests that, contrary to the common belief, x_u is not quasiuniversal quantity but rather depends on the secondary structure of native topology, which is probably responsible not only for the mechanical resistance but also for the distance between NS and TS along the end-to-end reaction coordinate.

Recently Plaxco et al. (18) have introduced the so-called contact order (CO) parameter to probe the folding of proteins. It has been shown that the CO correlates with folding rates of small two-state proteins but not of three-state ones (19). Because the CO reflect the topology of native conformation, one can use it to study the mechanical unfolding for which the native topology may play the important role.

In this article we address the following questions: 1), Do the α-proteins have, indeed, markedly higher values of x_u compared to β- and α/β-ones? 2), How x_u depends on the content of secondary structure? 3), What is the correlation between x_u and the mechanical stability of proteins? 4), Are these two quantities correlated with the contact order (CO) (18), which reflects the topology of the NS?

Because theoretical estimation of x_u by all-atom models is beyond present computational facilities, we use coarse-grained continuum representation for proteins in which only

Submitted February 8, 2007, and accepted for publication June 11, 2007.
Address reprint requests to Mai Suan Li, E-mail: masli@ifpan.edu.pl.
Editor: Angel E. Garcia.

0006-3495/07/10/2644/11 $2.00
doi: 10.1529/biophysj.107.106138

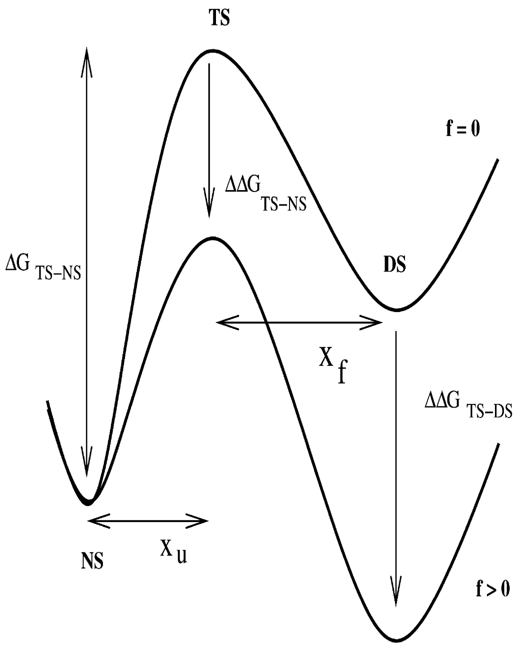

FIGURE 1 Schematic plot of the free energy profile of a two-state protein as a function of the end-to-end distance. In the absence of force (*upper curve*) the distance between NS and TS is x_{uf}; x_f refers to the distance between TS and denaturated state (*DS*). When the force is applied (*lower curve*) the unfolding barrier is lowered by amount of $\Delta\Delta G_{TS-NS} = -fx_u$ whereas the folding barrier increases by $\Delta\Delta G_{TS-DS} = fx_f$.

the positions of C_α-carbons are retained and the interactions between residues are assumed to be Go-like (20). Our choice is relied on the fact that the unfolding of biomolecules is mainly defined by the topology of the native conformation (15). Moreover, it has been shown (21) that the Go model (20) provides the reasonable estimate for x_u for ubiquitin and it is expected to work for other molecules.

To elucidate the role of the secondary structure, we have computed x_u for 23 proteins (Table 1). Our results are in good agreement with the experimental data not only for β- and α/β-proteins but also for the α-protein spectrin that has very high value of x_u (17). The careful analysis of the simulation and experimental sets reveals that, in term of x_u, the proteins can be divided into two classes. The first class consists of β- and α/β-proteins and the second class of α-helix proteins. The distance between the TS and NS of the first class is $x_u \approx 2$–5 Å, which is about three times lower than $x_u \approx 7$–15 Å of the second one. This is probably due to the fact that soft α-proteins are more sensitive to the external force compared to the hard ones. So the native topology is an important factor governing values of x_u.

We have shown that, x_u scales with the helix content linearly. This important prediction allows one to estimate x_u for α- and α/β-proteins using only the native conformation without performing any simulations or experiments. The correlation with the β-content is not pronounced.

To study the correlation between the mechanical stability and the distance between the NS and TS, we have carried out the constant force pulling simulations. One of the most interesting findings is that x_u decays with unfolding force, f_u (see Materials and Methods for the definition of f_u), by a power law, $x_u \sim f_u^{-\mu}$, where the exponent $\mu \approx 0.4$. However, the linear dependence between these quantities is not excluded.

As far as the correlation obtained for β-rich proteins is low, we use the CO (18) to probe the mechanical stability and the distance between the TS and NS. It turns out that CO is strongly correlated with x_u and f_u and it can, therefore, serve as a more general parameter to describe the mechanical stability of proteins compared to the secondary structure content.

MATERIALS AND METHODS

The energy of the Go-like model has the following form (20):

$$E = \sum_{\text{bonds}} K_r(r_i - r_{0i})^2 + \sum_{\text{angles}} K_\theta(\theta_i - \theta_{0i})^2 + \sum_{\text{dihedral}} \{K_\phi^{(1)}[1-\cos(\phi_i - \phi_{0i})] + K_\phi^{(3)}[1-\cos 3(\phi_i - \phi_{0i})]\} + \sum_{i>j-3}^{\text{NC}} \epsilon_H \left[5\left(\frac{r_{0ij}}{r_{ij}}\right)^{12} - 6\left(\frac{r_{0ij}}{r_{ij}}\right)^{10}\right] + \sum_{i>j-3}^{\text{NNC}} \epsilon_H \left(\frac{C}{r_{ij}}\right)^{12}. \tag{2}$$

Here $\Delta\phi_i = \phi_i - \phi_{0i}$, $r_{i,\,i+1}$ is the distance between beads i and $i+1$, θ_i is the bond angle between bonds $(i-1)$ and i, ϕ_i is the dihedral angle around the ith bond, and r_{ij} is the distance between the ith and jth residues. Subscripts ''0'', ''NC'', and ''NNC'' refer to the native conformation, native contacts, and nonnative contacts, respectively. Residues i and j are in native contact if r_{0ij} is less than a cutoff distance d_c taken to be $d_c = 6.5$ Å, where r_{0ij} is the distance between the residues in the native conformation.

The first harmonic term in Eq. 2 accounts for chain connectivity and the second term represents the bond angle potential. The potential for the dihedral angle degrees of freedom is given by the third term in Eq. 2. The interaction energy between residues that are separated by at least three beads is given by 10–12 Lennard-Jones potential. A soft sphere repulsive potential (the fourth term in Eq. 2) disfavors the formation of nonnative contacts. We choose $K_r = 100\epsilon_H/\text{Å}^2$, $K_\theta = 20\epsilon_H/rad^2$, $K_\phi^{(1)} = \epsilon_H$, and $K_\phi^{(3)} = 0.5\epsilon_H$, where ϵ_H is the characteristic hydrogen bond energy and $C = 4$ Å. As in the case of ubiquitin (21) we set $\epsilon_H = 0.98$ kcal/mol. Then temperature $T = 285$ K corresponds to $0.53\epsilon_H/k_B$. The computation has been performed at this temperature for all proteins. The force unit $[f] = \epsilon_H/\text{Å} = 68$ pN (21).

We assume that the dynamics of the polypeptide chain obeys the Langevin equation. The equations of motion (see Kouza et al. (22) for more details) were integrated using the velocity form of the Verlet algorithm (23) with the time step $\Delta t = 0.005\,\tau_L$, here $\tau_L = (ma^2/\epsilon_H)^{1/2} \approx 3$ ps, m is the typical mass of amino acids, and a is the distance between two neighboring beads.

In the constant force simulations we add an energy $-|\vec{f}.\vec{r}|$ to the total energy of the system, where $\vec{r}$ is the end-to-end vector and $\vec{f}$ is the force applied to the both termini. We define the unfolding time, τ_u, as the average of first passage times to reach a conformation that has no native contacts. Different trajectories start from the same native conformation but with different random number seeds. To get the reasonable estimate for τ_u, for each value of f we have generated 30–50 trajectories depending on proteins.

In the constant velocity force simulation we fix the N-terminal and pull the C-terminal by force $f = K_r(\nu t - x)$, where x is the displacement of the pulled atom from its original position (24) and the spring constant K_r is set to be the same as in Eq. 2. The pulling direction was chosen along the vector from fixed atom to pulled atom. The pulling speed is set equal $\nu = 3.6 \times 10^7$ nm/s for all proteins. The unfolding force, which characterizes mechanical

TABLE 1 The dependence of simulated (23 proteins) and experimental (12 proteins) values of x_u on the secondary structure content

Protein	PDB code	N	SCOP Class	α (%)	β (%)	CO[†]	CO[‡]	f_u(pN) Simulated	f_u(pN) Experimental	x_u (Å) Simulated	x_u (Å) Experimental	Reference
Protein L	1HZ6	62	α/β	23.8	44.4	0.222	0.161	254.3	125	2.0	2.2	(16)
Ubiquitin	1UBQ	76	α/β	21.1	43.4	0.219	0.150	163.9	203	2.4	1.4–2.5	(11,36,37)
Barnase	1BNI	108	α/β	23.2	21.3	0.096	0.109	44.2	70	3.3	3.3	(33)
Human DHFR	1HFR	186	α/β	26.3	24.7	0.175	0.131	50.3	72	3.2	3.7	(26)
GFP	1B9C	224	α/β	9.8	54.9	0.156	0.127	203.7	104	5.38	5.5	(38)
E2lip3	1QJO	80	all β	0	53.8	0.319	0.211	57.8	15	3.0	2.9	(39)
I27	1TIT	89	all β	0	59.6	0.296	0.178	248.2	180	3.2	3.3	(40)
Titin I1	1G1C	100	all β	0	64.3	0.286	0.182	272.7	127	4.4	3.5	(41)
Tenascin (TNfn3)	1TEN	89	all β	0	80.9	0.245	0.171	80.2	100	3.9	2.0–4.8	(42,43)
Fibronectin M10	1FNF	94	all β	0	44.7	0.216	0.174	76.2	75	3.2	3.8	(44,45)
ddFLN4	1KSR	100	all β	0	39.0	0.217	0.152	121.0	58	4.9	4.0–5.0	(9,46,47)
Domain C2A	1RSY	126	all β	0	43.7	0.208	0.159	59.8	60	3.9		(13)
spectrin	1AJ3	98	all α	87.8	0	0.035	0.080	6.8	30	14.2	15.0	(17)
Calmodulin	1CFC	148	all α	56.8	0	0.044	0.055	10.9	15	7.4	–	(13)
Ankyrin six-repeats	1N11	198	all α	63.6	0	0.042	0.044	19.0	50	12.6	–	(25)
*	2PDD	43	all α	48.8	0	0.114	0.110	6.9	–	7.00	–	(48)[§]
IM9	1IMQ	86	all α	53.5	0	0.116	0.118	33.8	–	8.3	–	(49)[§]
Cytochrome *c*	1YCC	103	all α	51.5	0	0.119	0.115	55.8	–	8.85	–	(50)[§]
Cytochrome *c*	1HRC	104	all α	44.2	0	0.121	0.111	53.7	–	7.04	–	(51)[§]
Acyl-coenzyme A	2ABD	86	all α	68.6	0	0.118	0.137	31.1	–	10.70	–	(52)[§]
Cytochrome B562	256B	106	all α	78.3	0	0.054	0.073	5.4	–	14.10	–	(53)[§]
λ-repressor	1LMB	80	all α	66.3	0	0.070	0.080	6.8	–	11.30	–	(54)[§]
Myoglobin	1F63	154	all α	87.0	0	0.050	0.082	7.5	–	15.77	–	(55)[§]

*The molecule full name is dihydrolipoamide acetyltransferase.
†The backbone CO.
‡The side chain CO.
§References, taken from PDB, refer to structures of proteins whose mechanical properties have not been studied yet.

stability of proteins, is identified as the maximum force, f_{max}, in the force-extension profile ($f_u \equiv f_{max}$).If this profile has several local maxima, then we choose the largest one.

To our best knowledge, experimentally x_u was determined for only 12 proteins (Table 1). In addition to this set we performed simulations for 11 more proteins. Among them domain C2A, calmodulin, and six-repeat ankyrin were studied by AFM but their x_u is not available. Mechanical properties of other α-proteins have not been studied yet but our choice is aimed at illuminating the difference between α-proteins and other proteins.

The native structures for Go modeling were obtained using the Protein Data Bank (PDB) code given in Table 1, but the following comments are in order. Following Li et al. (25), the native structure of the six-repeat ankyrin was taken from domains 2–7 of the PDB structure (code, 1N11). Ainavaparu et al. (26) performed the experiment for Chinese hamster ovary DHFR (CHO-DHFR), which consists of 186 amino acids. The structure of this protein has not been solved yet but the closest resembling DHFR is from *Homo sapiens* (human DHFR), having exactly 186 amino acids and sharing a very high sequence homology (>90%) with CHO-DHFR. Although the structure of human DHFR has been solved only in the presence of ligands we use its PDB structure (code, 1HFR) in our simulations.

The β-strand and α-helix contents are calculated using the information available in the PDB.

The CO is defined as follows (18)

$$\mathrm{CO} = \frac{\sum_{ij} \Delta_{ij} |i-j|}{N \sum_{ij} \Delta_{ij}}, \tag{3}$$

where N is the number of residues, and $\Delta_{ij} = 1$ if amino acids form contact and $\Delta_{ij} = 0$ otherwise. The values of CO depend on how a contact is defined. Here we use two definitions. In the first definition, which is based on the backbone, two residues i and j form a contact if they are not nearest neighbors ($|i-j| \geq 2$) and the distance between the C_α-carbons is smaller than $d_c = 6.5$ Å (the same cutoff as in our simulations). The second definition is based on the position of side chains (19). In this case a contact between any two amino acids ($|i-j| \geq 1$) is said to be formed if the distance between the centers of mass of side chains $d_{ij} \leq 6.0$ Å. The CO defined by two definitions will be referred to as the backbone and side chain CO. We have written our own code to compute the side chain CO but one can also use the program developed by Baker and co-workers (see http://depts.washington.edu/bakerpg/contact_order/). Our code gives essentially the same results. The values of the backbone and side chain CO for 23 proteins are listed in Table 1. The correlation level between these two sets is very high ($R = 0.96$). We have also tried other definitions for CO using different values of the cutoff $6 \leq d_c \leq 7.5$ Å and choices for pairs forming the contact ($|i-j| \geq m$, $m = 1, 2, 3$, and 4). Because the main conclusions are almost independent of definitions, we will present results for two cases mentioned above.

RESULTS

Estimation of *x*

Fig. 2 shows the dependence of the unfolding time τ_u on the constant force f for barnase, ddFLN4, and spectrin at $T = 285$ K. Similar to the case of ubiquitin (21), there exists the critical force f_c separating the low force and high force regime. The value of f_c depends on proteins and it is roughly equal 65 pN for barnase and 100 pN for ddFLN4 and spectrin. In the high force regime the unfolding barrier

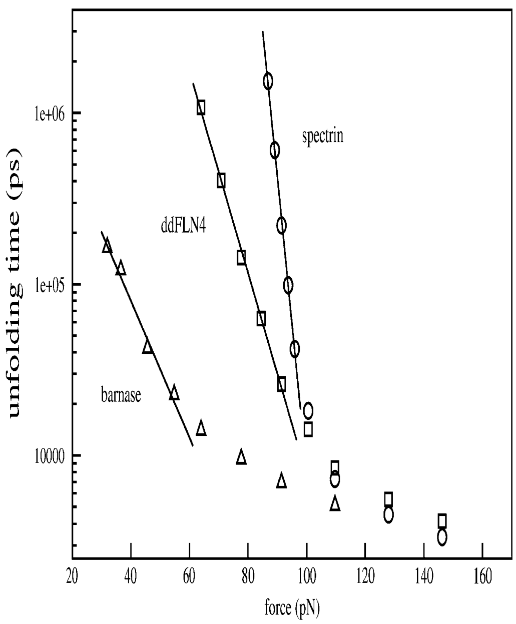

FIGURE 2 Dependence of τ_u on f for barnase (*triangles*), ddFLN4 (*squares*), and spectrin (*circles*). The straight lines correspond to the fit $y = 14.963–0.084x$ (barnase), $y = 22.516–0.124x$ (ddFLN4), and $y = 48.41–0.36x$ (spectrin). Using the fitting slope, γ, we have $x_u = \gamma k_B T/1$ pN, where $T = 285$ K. The corresponding values of x_u are shown in Table 1. One can show that at high forces τ_u depends on f linearly.

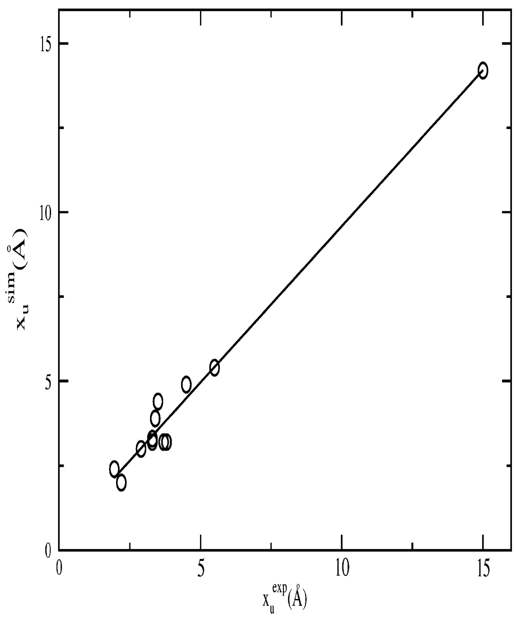

FIGURE 3 The simulation values of x_u are plotted versus the experimental ones. In the case when the experiments provide different values of x_u (Table 1), we took their average. The linear fit has the correlation level $R = 0.93$.

disappears and τ_u depends on f linearly (fit lines are not shown) as predicted theoretically by Evans and Ritchie (27).

In the low-force regime (Fig. 2) the dependence of τ_u on f becomes exponential (Eq. 1). Using slopes of linear fits we obtain x_u listed in Table 1 together with the results for other proteins. Using a more sophisticated version of Go model (28) for protein L, West et al. (29) have obtained $x_u \approx 1.9$ Å, which is consistent with our estimate. Thus, the value of x_u is not model-specific at least within the framework of the Go modeling. The question of to what extent the nonnative interactions change x_u remains to be elucidated. However, as evident from Table 1, this Go-like model gives acceptable agreement with the experiments. It is also clearly demonstrated in Fig. 3 where we plot the theoretical values of x_u against the experimental ones. The high correlation level ($R = 0.93$) between the theory and experiments justifies the use of Go modeling for computation of x_u. Our result is very appealing because this simple model provides even the quantitative agreement with experiments.

Dependence of x_u on the secondary structure of proteins

One of the most exciting results (Table 1) is that, in agreement with the experiments (17), we obtained the large value $x_u \approx 14.2$ Å for spectrin that contain only helices. The mechanical resistance to the external force of two other α-proteins, six-repeat ankyrin and calmodulin, have been studied experimentally (13,25) but their values of x_u were not reported. We predict that these and the other eight α-proteins (the mechanical unfolding of these proteins has not been investigated yet) have large distances between the NS and TS (Table 1). It would be very interesting to check our prediction experimentally.

Based on the results from Table 1, one can divide proteins into two main classes. One of them consists of α- and α/β-proteins that have $x_u \approx 2–5$ Å. The another one is of purely α-proteins with markedly higher values $x_u \approx 7–15$ Å. In terms of x_u, neither our simulations nor the experiments can distinguish clearly the β-proteins from the mixed ones. The larger statistics is required to solve this delicate issue. Nevertheless, it is obvious that the secondary structure of the native conformation plays the decisive role in determination of the distance between the TS and NS. To demonstrate this better, we plot x_u as a function of the helix content for α- and α/β-proteins (Fig. 4). The linear regression for the experimental and simulation data gives

$$x_u = \begin{cases} -0.431 + 0.174\alpha, & \text{from simulations} \\ 0.159 + 0.16\alpha, & \text{from experiments,} \end{cases} \tag{4}$$

where x_u is measured in angstroms and the helix content α is measured in percent. The correlation level between x_u and the helix content is equal to 0.91 and 0.94 for the experimental and the simulation set, respectively. Such high quality of fitting unambiguously shows the strong correlation between these two quantities: the higher helix content, the

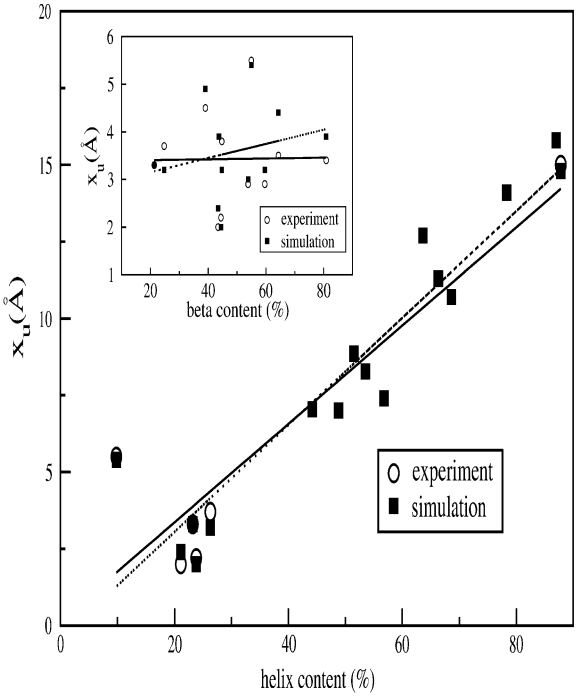

FIGURE 4 Dependence of x_u on the helix content (%) for α- and α/β-proteins. The circles and squares refer to the experimental and simulation results, respectively. The linear fits to simulation (*dotted line*, $y = -0.431 + 0.174x$) and experimental (*solid line*, $y = 0.159 + 0.16x$) data have the correlation level $R = 0.94$ and 0.91, respectively. The inset shows the dependence of x_u on the β-content (%) for the set of the all β- and α/β-proteins. For the simulation set the linear fit ($y = 2.854 + 1.488x$) gives $R = 0.25$. In the case of experimental data we have $R = 0.02$ ($y = 3.385 + 0.001x$).

larger is x_u. It should be stressed that Eq. 4 is useful for estimating x_u based solely on the topology of the native state.

The dependence of x_u on the β-content for β- and α/β-proteins is presented in the inset of Fig. 4. The correlation level is very low for both the experimental ($R = 0.02$) and simulation ($R = 0.25$) sets. If one considers only β-proteins then for the experimental set x_u drops with the β-content with substantially higher correlation ($R = 0.53$). Unfortunately, this result is not robust due to the small data set of only six proteins. If the protein ddFLN4 were removed, e.g., then the correlation reduced to $R = 0.1$. So, to establish if there is any pronounced relation between x_u and the β-content one has to carry out more experiments to generate better statistics.

Correlation between x_u and mechanical stability

To define the mechanical stability quantitatively we perform the pulling simulations with the constant velocity force (see Materials and Methods). The unfolding force, f_u, which corresponds to the local maximum force in the force-extension profile, is considered as a measure of mechanical stability. We choose the pulling speed $\nu = 3.6 \times 10^7$ nm/s, which is about five orders of magnitude faster than typical experimental values but is about two to three orders of magnitude slower than those used in all-atom steered molecular dynamics simulations (30).

Fig. 5 shows the dependence of the force on the extension for spectrin, ddFLN4, and barnase, which are representative for three types of proteins. From this plot one can obtain f_u for the given loading rate. The unfolding forces at other pulling speeds may be estimated using the logarithmic dependence $f_u \sim \ln\nu$ (27). The results presented in Fig. 6 ascertain this dependence for our Go model. The similar behavior has been also found for ubiquitin molecules using a different version of Go models (31).

For spectrin Rief et al. (17), for example, reported $f_u \approx 30$ pN at the pulling speed $\nu = 800$ nm/s. Applying the formula $f_u \sim \ln \nu$ (27), we obtain $f_u \approx 321$ pN for the speed used in our simulation. This value is much higher than our estimate $f_u \approx 7$ pN (*inset* in Fig. 5). On the other hand, as seen from Fig. 5, in agreement with the experiments (32,33), ddFLN4 and barnase can withstand higher force than spectrin. Thus, although the Go model is not able to reproduce experimental values for f_u for individual proteins, it remains useful to predict their relative values. This becomes much more evident if we plot the simulated values of f_u against the experimental ones (Fig. 7). The high correlation level ($R = 0.78$) between these two sets gives us confidence to use the

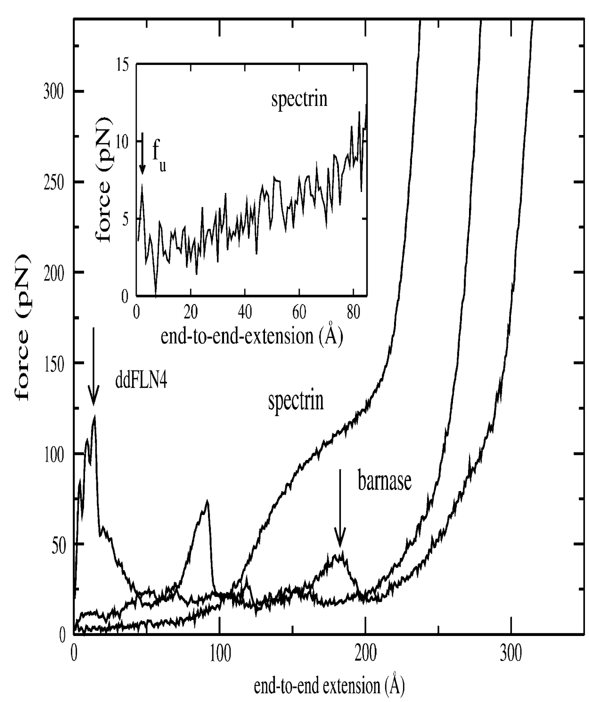

FIGURE 5 The force-extension profiles obtained by the pulling simulations with constant velocity force for ddFLN4, barnase, and spectrin. $\nu = 3.6 \times 10^7$ nm/s and $T = 285$ K. The results are averaged over 50 trajectories for ddFLN4 and barnase and 100 trajectories for spectrin. The arrows refer to the unfolding force. The inset shows the results for spectrin at small extensions. For this protein we have $f_u = f_{max} = 7 \pm 1$ pN.

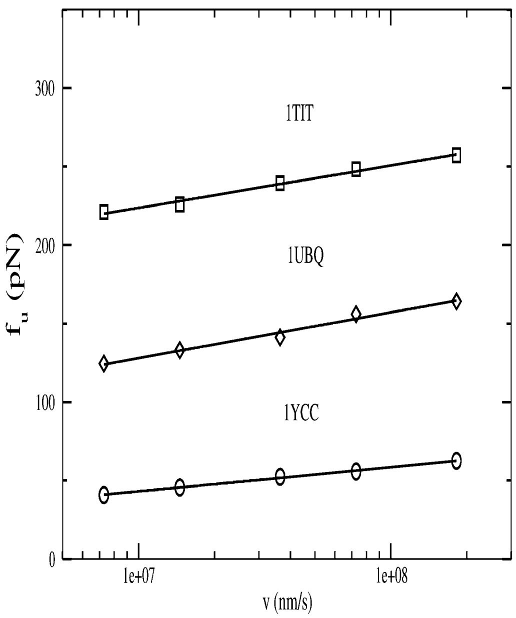

FIGURE 6 The dependence of f_u (pN) on the loading rate ν (nm/s) for three proteins. Straight lines are linear fits $y = 34.0 + 11.7x$, $y = -75.73 + 12.64x$, and $y = -64.7 + 6.69x$ for the titin domain I27, ubiquitin, and cytochrome *c* (1YCC), respectively.

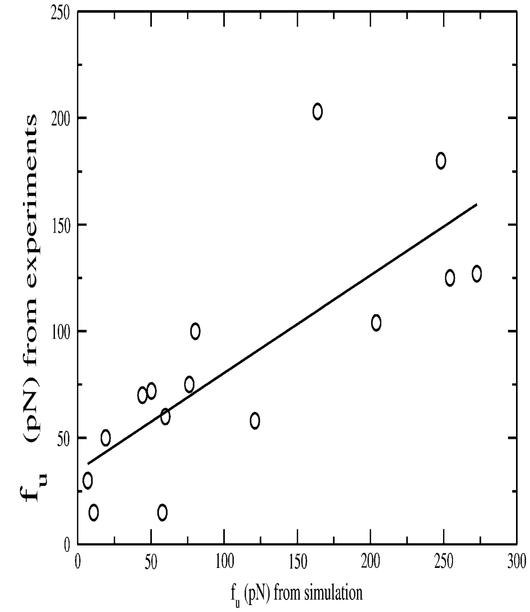

FIGURE 7 The experimental values of f_u is plotted as a function of f_u obtained from our simulations at $\nu = 3.6 \times 10^7$ nm/s. We have collected the experimental values of f_u at $\nu = 300$ nm/s for all proteins. If f_u is not available at this speed we used the formula $f_u \sim \ln\nu$ (27) to extract it from data obtained for other speeds. The references to experimental works are given either in Table 1 of this article or in Table 1 of Brockwell et al. (16). The straight line refers to the linear fit ($y = 34.637 + 0.458x$) with the correlation level of 0.78.

Go model to study the mechanical stability of proteins. It should be stressed again that despite very different loading rates the values of f_u obtained from the Go modeling and experiments are in the same order of magnitude. This is probably an artifact of the simple Go modeling where non-native interactions and the effect of environment are not taken into account.

Fig. 8 *a* shows the dependence of x_u on the unfolding force f_u. The linear fitting to the simulation data gives

$$x_u = 9.456 - 0.0304 f_u, \tag{5}$$

where the correlation $R = 0.61$. It becomes worse ($R = 0.47$) in the case of the experimental set.

Because the correlation is not high in both cases, we try to see if the nonlinear fitting can get it enhanced. From the log-log plot (Fig. 8 *b*) for the simulation set it follows that x_u decays with f_u by a power law,

$$x_u = c f_u^{-\mu}, \tag{6}$$

where the constant $c \approx 24.6$, the exponent $\mu = 0.39 \pm 0.07$, and f_u is measured in pN. To check the robustness of the power law (Eq. 6) we removed one protein from the whole set and calculated μ and R for the subset of 22 proteins. Repeating this procedure 23 times we see that the exponent and the correlation vary between $0.361 \leq \mu \leq 0.415$ and $0.76 \leq R \leq 0.81$. For example, if the protein calmodulin with $f_u^{sim} = 10.9$ pN were removed, then $\mu = 0.397 \pm 0.069$ and $R = 0.79$. Thus the ‘‘scaling’’ law given by Eq. 6 is robust, and mechanically more stable biomolecules would have lower values of x_u because they are less sensitive to external perturbation. It should be noted that Eq. 6 has been derived for the pulling speed $\nu = 3.6 \times 10^7$ nm/s. For the other speed ν' the power law behavior is still valid but with the different constant c'. Using $f_u \sim \ln \nu$ (27), one can show that $c' = (\ln\nu'/\ln\nu)^{\mu} c$.

Having much higher correlation level ($R = 0.78$) the power law should work better than the linear fit and it strongly supports the nonlinear correlation between the distance between the NS and TS along the end-to-end distance reaction coordinate and the mechanical resistance. However it does not necessarily exclude the possibility that x_u depends on f_u linearly (Eq. 5) because the lower correlation given by the linear procedure may be merely a result of insufficient statistics. This question is left for future investigation.

From the experimental data set (Fig. 8 *b*) we obtain the exponent $\mu = 0.35 \pm 0.17$. Within error bars this value of μ coincides with the simulation estimate but the correlation level $R = 0.51$ is notably lower. There are several possible reasons why the correlation is not so high. First, the data set of 12 proteins is not large enough. Second, for some proteins the experimental values of f_u were not obtained at the speed $\nu = 300$ nm/s and the values shown in Fig. 8 have been estimated from other speeds as described in the caption to Fig. 7. This interpolation procedure may cause some inaccuracies. Thus, accurate measurements for a larger set of proteins at the same speed are needed to verify the power law given by Eq. 6.

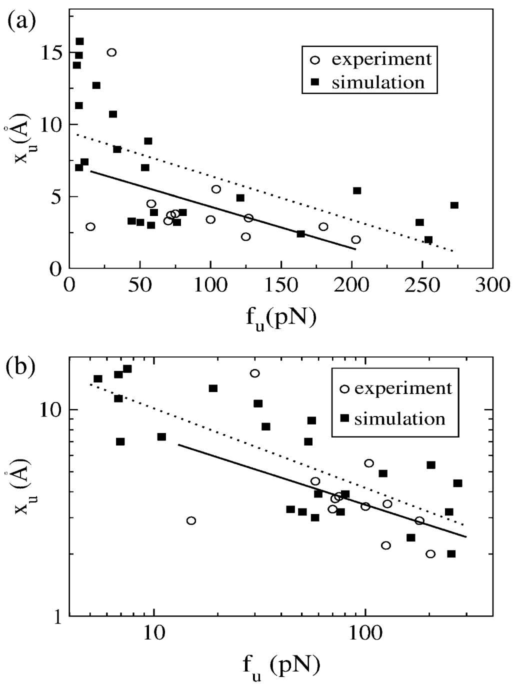

FIGURE 8 (*a*) The dependence of x_u on f_u. The circles and squares refer to the experimental and simulation results, respectively; f_u has been obtained at the same pulling speeds as in Fig. 7. The linear fit (*dotted line*, $y = 9.45 - 0.0304x$) to simulation data has the correlation level $R = 0.61$ whereas the experimental data give the lower value $R = 0.47$ (*solid line*, $y = 7.178 - 0.0289x$). (*b*) The same as in panel *a* but for the log-log plot. The linear fit (*dotted line*, $y = 3.203 - 0.385x$) to simulation data has the correlation level $R = 0.78$ whereas the experimental data give the lower value $R = 0.51$ (*solid line*, $y = 2.76 - 0.33x$).

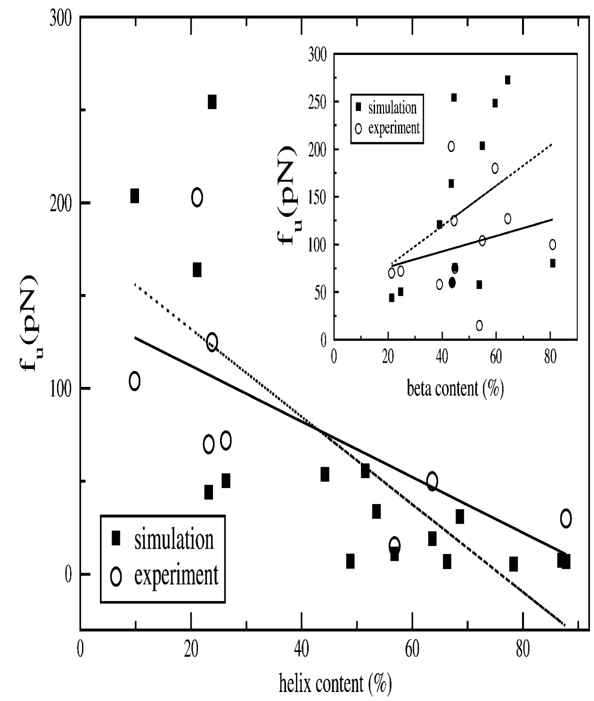

FIGURE 9 The dependence of f_u (pN) on the helix content (%) for α- and α/β-proteins. The linear fit to the simulation (*dotted line*, $y = 179.19 - 2.36x$) and experimental (*solid line*, $y = 142.09 - 1.497x$) data has the correlation level $R = 0.74$ and 0.67, respectively. The inset shows the dependence of x_u on the β-content (%) for the all β- and α/β-proteins. The linear fit to the simulation (*dotted line*, $y = 34.864 + 2.112x$) and experimental (*solid line*, $y = 59.924 + 0.818x$) data has the correlation $R = 0.38$ and 0.25, respectively.

Correlation between f_u and the secondary structure

Fig. 9 shows the dependence of f_u on the helix content for the set of α- and α/β-proteins. Using the linear fit we obtain

$$f_u = \begin{cases} 179.19 - 2.36\alpha, & \text{from simulations} \\ 142.09 - 1.497\alpha, & \text{from experiments,} \end{cases} \tag{7}$$

where f_u is measured in pN and the helix content α in percents. The correlation between the mechanical stability and the helicity is pronounced as the correlation level is rather high for the simulation ($R = 0.74$) as well as the experimental ($R = 0.67$) set. Thus Eq. 7 supports the experimental fact that α-rich proteins have low resistance to the external mechanical perturbation. It remains unclear if f_u correlates with the β-content because $R < 0.4$ (*inset* in Fig. 9). As in the case of x_u the poor correlation may be due to the small data set. This calls for further theoretical and experimental studies.

As follows from Eqs. 4 and 7, both of x_u and f_u depend on the helicity linearly. On the other hand, the connection between them may be better described by a power law (Eq. 6) than the linear one. Therefore it is interesting to try also the nonlinear fit for the dependence of f_u on the helix content. From Fig. 10 we obtain the following power law

$$f_u = a\alpha^{-\gamma}. \tag{8}$$

For the theoretical set constant $a \approx 19349.1$, the exponent $\gamma = 1.729 \pm 0.307$ and $R = 0.83$. In the case of the experimental set the correlation remains high ($R = 0.74$) but the exponent becomes much lower, $\gamma = 0.848 \pm 0.316$. So in terms of correlation level the power law works better than the linear fit but to decide what scenario is the really superior one needs much better statistics. Nevertheless we believe that f_u depends on the helicity and one can estimate the unfolding force of α- and α/β-proteins using only the native topology.

For the theoretical data, the correlation between f_u and the β-content gets better compared to the linear fitting (*inset* in Fig. 10) but it remains rather low ($R = 0.51$). As to the experimental set it becomes even worse ($R = 0.16$). Therefore the power law fit in Equation 8 does not improve the correlation between f_u and the β-content.

As has been shown above, it is not clear if one can estimate x_u and f_u using the β-content of strands of a protein. In the other words, the problem remains unsolved for β-rich proteins. In the next sections we will show how to use the CO to overcome this difficulty.

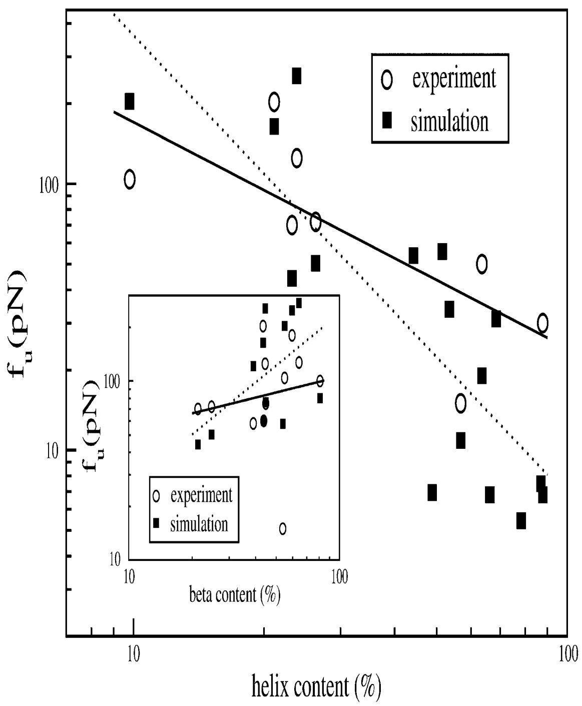

FIGURE 10 The same as in Fig. 9 but the log-log plot is used. The linear fit to the simulation (*dotted line*, $y = 9.87 - 1.729x$) and experimental (*solid line*, $y = 7.09 - 0.848x$) data has the correlation $R = 0.83$ and $R = 0.74$, respectively. The inset shows the dependence of f_u on the β-content. The linear fit to the simulation (*dotted line*, $y = 1.031 + 0.964x$) and experimental (*solid line*, $y = 3.321 + 0.291x$) set gives $R = 0.53$ and $R = 0.16$, respectively.

Correlation between x_u and CO

To better understand why one can use the CO to study the mechanical stability, we first plot the secondary structure content as a function of the backbone CO in Fig. 11 (similar results for the side chain CO case not shown). In general, helix-rich proteins have more local contacts compared to β-rich proteins and consequently the helicity falls with increasing CO whereas the percentage of β-structures increases. The correlation is rather high even for small sets of proteins ($R = 0.81$ and 0.66 for the helix and β-content, respectively).

The dependence of x_u on the CO is shown in Fig. 12 where the correlation levels are almost the same for two definitions of CO. The quality of the linear fitting to the experimental data is lower but one can expect its improvement for larger statistics. The decrease of x_u with CO is consistent with the fact that x_u depends on the helicity (Fig. 4) whereas the latter is anticorrelated with CO (Fig. 11). Using the linear fits from Fig. 12, *a* and *c*, we obtain, for example, the following dependence of x_u on the backbone CO:

$$x_u = \begin{cases} 13.047 - 39.51 \times CO, & \text{from simulations} \\ 10.478 - 29.558 \times CO, & \text{from experiments.} \end{cases} \quad (9)$$

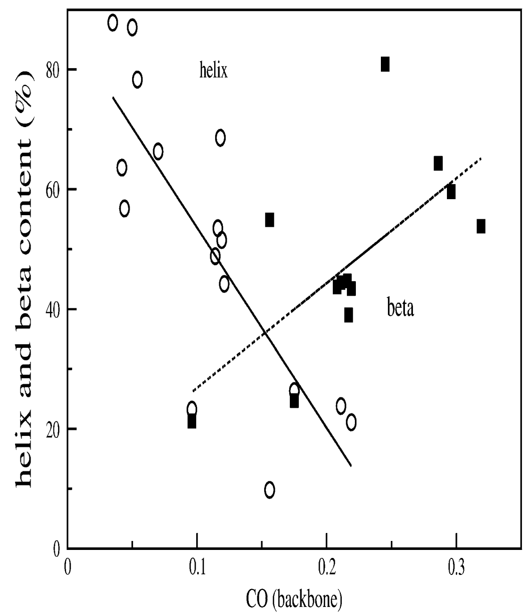

FIGURE 11 Dependence of the helix (*circles*) and β-content (*solid squares*) on the backbone CO. The straight lines refer to the linear fit $y = 87.047 - 334.57x$ ($R = 0.81$) and $y = 9.345 + 174.95x$ ($R = 0.66$) for the helix and β-content, respectively.

A similar equation may be written down for the side chain CO dependence using the information given in the caption to Fig. 12.

Thus we obtain very important results that the CO might be used to estimate the distance between the TS and NS of globular proteins regardless of whether they are helix- or β-rich. This parameter is more universal than the helicity because the latter can be applied to the α- and α/β-proteins only.

Correlation between f_u and CO

Fig. 13 shows the dependence of f_u on the CO. Because helix-rich proteins are more mechanically stable compared to helix-poor proteins, f_u grows with the CO (Fig. 13). This conclusion may be partially understood using the results shown in Figs. 9 and 11. The correlation for the theoretical set is acceptable and the linear fitting gives

$$f_u = \begin{cases} -24.749 + 691.628 \times CO, & \text{backbone CO} \\ -72.481 + 1215.704 \times CO, & \text{side chain CO.} \end{cases} \quad (10)$$

In addition to Eqs. 7 and 8 one can use this equation to estimate the mechanical stability from the native topology. It would be even better to employ all of these equations to find an optimal value of f_u that is closest to the experimental result.

For the experimental set the correlation is low and not robust. In the backbone CO case (Fig. 13 *c*), for example, it

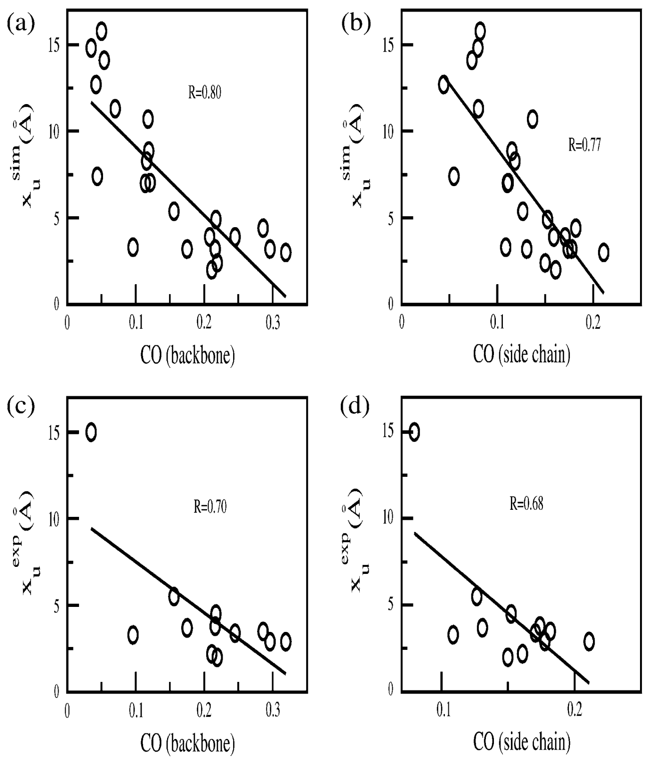

FIGURE 12 (*a*) Dependence of theoretical values of x_u on the backbone CO (linear fit $y = 13.047 - 39.51x$). (*b*) The same as in panel *a* but for the side chain CO (linear fit $y = 16.49 - 75.111x$). (*c*) The same as in panel *a* but the experimental values of x_u (linear fit $y = 10.478 - 29.558x$). The same as in panel *b* but for the experimental values of x_u (linear fit $y = 14.427 - 66.025x$). The correlation level of the linear fitting is shown on the plot.

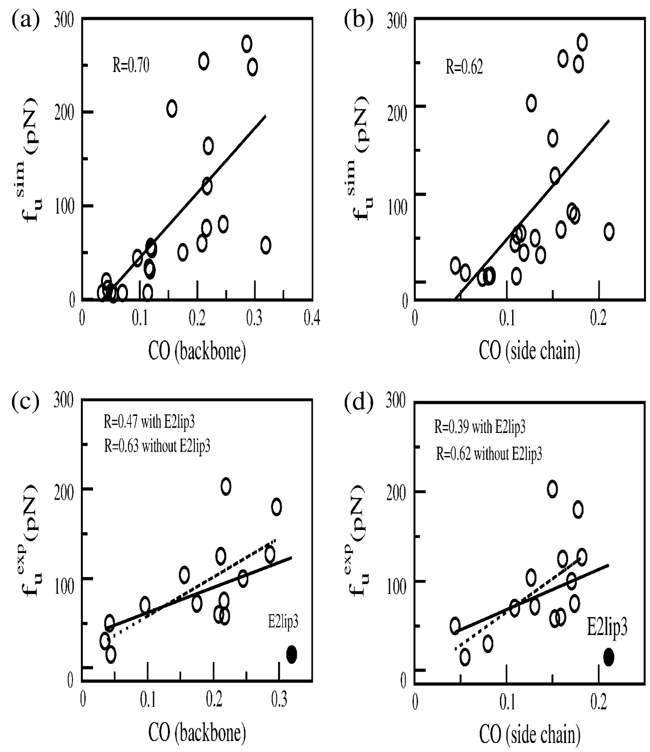

FIGURE 13 The same as in Fig. 12 but for f_u. The linear fits are $y = -24.749 + 691.628x$, $y = -72.481 + 1215.704x$, $y = 33.949 + 280.2x$, and $y = 23.181 + 449.78x$ for panels *a–d*, respectively. The correlation level of the linear fitting is shown on the plot.

arises from $R = 0.47$ to $R = 0.63$ if the protein E2lip3 were removed from the full set. It is also true when f_u is plotted against the side chain CO (Fig. 13 *d*). Thus, more experimental data are needed to ascertain the correlation between f_u and CO. Nevertheless we expect that they depend on each other as there exists the strong correlation between simulated and experimental values of f_u (Fig. 7).

CONCLUSION

Equation 1 is, in principle, valid for two-state proteins. In the case of three-state proteins, x_u may be considered as the sum of the distance between the first TS and NS and the distance between the second TS and the intermediate state. In force-clamp experiments, one end of proteins is kept fixed and the other end is pulled. In these simulations the external force is applied to both termini to accelerate unfolding about two times (21). Since fixing one end slows down unfolding but leaves x_u unchanged (21), it is not surprising that our results agree with the force-clamp data. The fact that the Go modeling provides the good agreement with experiments demonstrates that x_u largely depends on the topology of the native conformation but not on nonnative contacts.

Having performed simulations for 23 wild-type proteins, we have made a number of predictions. First, the distance x_u of α-proteins is about three times larger than that for β- and α/β-proteins. So far x_u was measured only for α-protein spectrin (17); more experiments on other proteins are needed to ascertain this prediction. Second, we have shown that for the α- and α/β-proteins the dependence of the distance between NS and TS on the helix content follows Eq. 4. This result is very useful because the estimation of x_u requires only the knowledge about the native topology. The question of if one can compute x_u of β-proteins using the β-content as a unique parameter, is left for further experimental as well as theoretical studies. However, we believe that this class of proteins has $2 \lesssim x_u \lesssim 5$ Å. Third, for the first time our simulations clearly show that the distance between NS and TS depends on the unfolding force by a power law with the exponent $\mu \approx 0.4$ although the linear dependence (Eq. 5) is also possible. Equation 6 can be used for estimating x_u provided f_u is known. The experimental data seem to support this prediction but additional measurements have to be performed at the same conditions (pulling speed, T, pH, etc.) for different proteins to confirm it on the quantitative level. Fourth, one could obtain the unfolding force of α- and mixed α/β-proteins using the helix content only (Eqs. 7 and 8). Fifth, the CO, which is a more universal parameter compared to percentages of secondary structures might be used to estimate x_u and f_u for any type of proteins. This parameter is successful in predicting the mechanical stability because it reflects the native topology that is believed to play the key role in the mechanical unfolding.

Using the data on Table 1, one can show that x_u has little correlation with the number of amino acids (the correlation level is below 0.5). Probably the size of proteins affects the prefactor in Eq. 1 but not the exponent itself. Recent experiments (34) and simulations (21,35) on refolding have shown that the force-clamp refolding technique can serve as a useful tool to probe the distance between the TS and DS, x_f (see Fig. 1). With the help of this technique one can demonstrate that x_f of domain I27 and ubiquitin is about three times larger than x_u. It would be highly useful to obtain x_f for other proteins and elucidate the impact of secondary structure and mechanical stability on it.

I thank M. Kouza, J. M. Fernandez, and S. R. K. Ainavarapu for very useful discussions and correspondence. I am extremely graceful to the referee for his/her important suggestion on the application of the CO to study the mechanical stability of a protein.

This work was supported by the Polish Komitet Badan Naukowych grant No. 1P03B01827.

REFERENCES

1. Onuchic, J. N., and P. G. Wolynes. 2004. Theory of protein folding. *Curr. Opin. Struct. Biol.* 14:70–75.
2. Shakhnovich, E. I. 2006. Protein folding thermodynamics and dynamics: where physics, chemistry and biology meet. *Chem. Rev.* 106:1559–1588.
3. Gnanakaran, S., H. Nymeyer, J. Portman, K. Y. Sanbonmatsu, and A. E. Garcia. 2003. Peptide folding simulations. *Curr. Opin. Struct. Biol.* 13:168–174.
4. Nguyen, P. H., G. Stock, E. Mittag, C. K. Hu, and M. S. Li. 2005. Free energy landscape and folding mechanism of β-hairpin in explicit water: a replica exchange molecular dynamics study. *Proteins. Structures, Functions, and Bioinformatics.* 61:795–808.
5. 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.
6. Bell, G. I. 1978. Models for the specific adhesion of cells to cells. *Science.* 100:618–627.
7. Bartolo, D., I. Derenyi, and A. Ajdari. 2002. Dynamic response of adhesion complexes: beyond the single-path picture. *Phys. Rev. E.* 65: 051910–051913.
8. Hummer, G., and A. Szabo. 2003. Kinetics from nonequilibrium single-molecule pulling experiments. *Biophys. J.* 85:5–15.
9. Schlierf, M., and M. Rief. 2006. Single-molecule unfolding force distributions reveal a funnel-shaped energy landscape. *Biophys. J.* 90:L33–L35.
10. Dudko, O. K., G. Hummer, and A. Szabo. 2006. Intrinsic rates and activation free energies from single-molecule pulling experiments. *Phys. Rev. Lett.* 96:108101–108104.
11. 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.
12. Brockwell, D. J., G. S. Beddard, J. Clarkson, R. C. Zinober, A. W. Blake, J. Trinick, P. D. Omsted, D. A. Smith, and S. E. Radford. 2002. The effect of core destabilization on the mechanical resistance of i27. *Biophys. J.* 83:458–472.
13. Carrion-Vazquez, M., A. Oberhauser, T. Fisher, P. Marszalek, H. Li, and J. Fernandez. 2000. Mechanical design of proteins studied by single-molecule force spectroscopy and protein engineering. *Prog. Biophys. Mol. Biol.* 74:63–91.
14. Ortiz, V., S. O. Nielsen, M. L. Klein, and D. E. Discher. 2005. Unfolding a linker between helical repeats. *J. Mol. Biol.* 349:638–647.
15. 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.
16. Brockwell, D. J., G. Beddard, E. Paci, D. West, P. Olmsted, D. Smith, and S. Radford. 2005. Mechanically unfolding the small, topologically simple protein l. *Biophys. J.* 89:506–519.
17. Rief, M., J. Pascual, M. Saraste, and H. Gaub. 1999. Single molecule force spectroscopy of spectrin repeats: low unfolding forces in helix bundles. *J. Mol. Biol.* 286:553–561.
18. Plaxco, K. W., K. T. Simon, and D. Baker. 1998. Contact order, transition state placement and the refolding rates of single domain proteins. *J. Mol. Biol.* 277:985–994.
19. Ivankov, D. N., S. O. Garbuzynskiy, E. Alm, K. W. Plaxco, D. Baker, and A. V. Finkelstein. 2003. Contact order revisited: influence of protein size on the folding rate. *Protein Sci.* 19:2057–2062.
20. Clementi, C., H. Nymeyer, and J. N. Onuchic. 2000. Topological and energetic factors: what determines the structural detail of transition state ensemble and en-route intermediates for protein folding? An investigation for small globular proteins. *J. Mol. Biol.* 298:937–953.
21. Li, M. S., M. Kouza, and C. K. Hu. 2007. Refolding upon force quench and pathways of mechanical and thermal unfolding of ubiquitin. *Biophys. J.* 91:547–551.
22. Kouza, M., C. F. Chang, S. Hayryan, T. H. Yu, M. S. Li, T. H. Huang, and C. K. Hu. 2005. Folding of the protein domain hbSBD. *Biophys. J.* 89:3353–3361.
23. Swope, W. C., H. C. Andersen, P. H. Berens, and K. R. Wilson. 1982. Computer simulation method for the calculation of equilibrium constants for the formation of physical clusters and molecules: application to small water clusters. *J. Chem. Phys.* 76:637–649.
24. Lu, H., B. Isralewitz, A. Krammer, V. Vogel, and K. Schulten. 1998. Unfolding of titin immunoglobulin domains by steered molecular dynamics simulation. *Biophys. J.* 75:662–671.
25. Li, L., S. Wetzel, A. Pluckthun, and J. Fernandez. 2006. Stepwise unfolding of ankyrin repeats in a single protein revealed by atomic force microscopy. *Biophys. J.* 90:L30–L32.
26. Ainavarapu, S., L. Li, and J. Fernandez. 2005. Ligand binding modulates the mechanical stability of dihydrofolate reductase. *Biophys. J.* 89:3337–3344.
27. Evans, E., and K. Ritchie. 1997. Dynamics strength of molecular adhesion bonds. *Biophys. J.* 72:1541–1555.
28. Karanicolas, J., and C. L. Brooks. 2002. The origins of asymmetry in the folding transition states of protein L and protein G. *Protein Sci.* 11: 2351–2361.
29. West, D., P. Olmsted, and E. Paci. 2006. Mechanical unfolding revisited through a simple but realistic model. *J. Chem. Phys.* 124:154909–154917.
30. Gao, M., M. Sotomayor, E. Villa, E. H. Lee, and K. Schulten. 2006. Molecular mechanisms of cellular mechanics. *Phys. Chem. Chem. Phys.* 8: 3692–3706.
31. Cieplak, M., and P. E. Marshalek. 2005. Mechanical unfolding of ubiquitin molecules. *J. Chem. Phys.* 123:194909–194915.
32. Schwaiger, I., A. Kardinal, M. Schleicher, A. A. Noegel, and M. Rief. 2004. A mechanical unfolding intermediate in an actin-crosslinking protein. *Nat. Struct. Mol. Biol.* 11:81–85.
33. Best, R., B. Li, A. Steward, V. Daggett, and J. Clarke. 2001. Can nonmechanical proteins withstand force? Stretching barnase by atomic force microscopy and molecular dynamics simulations. *Biophys. J.* 81:2344–2356.
34. Fernandez, J. M., and H. Li. 2004. Force-clamp spectroscopy monitors the folding trajectory of a single protein. *Science.* 303:1674–1678.

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

36. 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.

37. Chyan, C., F. Lin, H. Peng, J. Yuan, C. Chang, S. Lin, and G. Yang. 2004. Reversible mechanical unfolding of single ubiquitin molecules. *Biophys. J.* 87:3995–4006.

38. Dietz, H., and M. Rief. 2004. Exploring the energy landscape of GFP by single-molecule mechanical experiments. *Proc. Natl. Acad. Sci. USA.* 101:16192–16197.

39. Brockwell, D. J., E. Paci, R. Zinober, G. Beddard, P. Olmsted, D. Smith, R. Perham, and S. Radford. 2003. Pulling geometry defines the mechanical resistance of a β-sheet protein. *Nat. Struct. Biol.* 10:731–737.

40. Best, R., S. Fowler, J. Herrera, A. Steward, E. Paci, and J. Clarke. 2003. Mechanical unfolding of a titin Ig domain: structure of transition state revealed by combining atomic force microscopy, protein engineering and molecular dynamics simulations. *J. Mol. Biol.* 330:867–877.

41. Li, H., and J. Fernandez. 2003. Mechanical design of the first promixal Ig domain of human cardiac titin revealed by single molecules force spectroscopy. *J. Mol. Biol.* 334:75–86.

42. Ng, S., R. Rounsevell, A. Steward, C. Geierhass, P. 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 simulations. *J. Mol. Biol.* 330:776–789.

43. Wang, M., Y. Cao, and H. Li. 2006. The unfolding and folding of TNfnall probed by single molecule force-ramp spectroscopy. *Polym.* 47:2548–2554.

44. Oberhauser, A., C. Badilla-Fernandez, M. Carrion-Vazquez, and J. Fernandez. 2002. The mechanical hierarchies of fibronectin observed with single-molecule AFM. *J. Mol. Biol.* 319:433–447.

45. Li, L., H. Huang, C. L. Badilla, and J. Fernandez. 2005. Mechanical unfolding intermediates observed by single-molecule force spectroscopy in a fibronectin type III module. *J. Mol. Biol.* 345:817–826.

46. Schlierf, M., and M. Rief. 2005. Temperature softening of a protein in single-molecule experiments. *J. Mol. Biol.* 354:497–503.

47. Schwaiger, I., M. Schlierf, A. Noegel, and M. Rief. 2005. The folding pathways of a fast-folding immunoglobulin domain revealed by single-molecule mechanical experiments. *EMBO Rep.* 6:46–51.

48. Kalia, Y. N., S. M. Brocklehurst, D. S. Hipps, E. Appella, K. Sakaguchi, and R. N. Perham. 1993. The high-resolution structure of the peripheral subunit-binding domain of dihydrolipoamide acetyltransferase from the pyruvate dehydrogenase multienzyme complex of Bacillus stearothermophilus. *J. Mol. Biol.* 230:323–341.

49. Osborne, M. J., A. Breeze, L. Lian, A. Reilly, R. James, C. Kleanthous, and G. R. Moore. 1996. Three-dimensional solution structure and 13C nuclear magnetic resonance assignments of the colicin E9 immunity protein Im9. *Biochemistry.* 35:9505–9512.

50. Louie, G., and G. Brayer. 1990. High-resolution refinement of yeast iso-1-cytochrome c and comparisons with other eukaryotic cytochromes c. *J. Mol. Biol.* 214:527–555.

51. Bushnell, G., G. Louie, and G. Brayer. 1990. High-resolution three-dimensional structure of horse heart cytochrome c. *J. Mol. Biol.* 214: 585–595.

52. Andersen, K., and F. Poulsen. 1993. The three-dimensional structure of acyl-coenzyme A binding protein from bovine liver: structural refinement using heteronuclear multidimensional NMR spectroscopy. *J. Biomol. NMR.* 3:271–284.

53. Lederer, F., A. Glatigny, P. Bethge, H. Bellamy, and F. Matthew. 1981. Improvement of the 2.5 A resolution model of cytochrome b562 by redetermining the primary structure and using molecular graphics. *J. Mol. Biol.* 148:427–448.

54. Beamer, L., and C. Pabo. 1992. Refined 1.8 A crystal structure of the lambda repressor-operator complex. *J. Mol. Biol.* 227:177–196.

55. Brunori, M., F. Cutruzzola, C. Savino, C. Travaglini-Allocatelli, B. Vallone, and Q. Gibson. 1999. Structural dynamics of ligand diffusion in the protein matrix: a study on a new myoglobin mutant Y(B10) Q(E7) R(E10). *Biophys. J.* 76:1259–1269.