

# Unfolding proteins by external forces and temperature: The importance of topology and energetics

Emanuele Paci\*<sup>†</sup> and Martin Karplus\*<sup>†‡§</sup>

\*Laboratoire de Chimie Biophysique, ISIS, Université Louis Pasteur, 4 rue Blaise Pascal, 67000 Strasbourg, France; <sup>†</sup>Oxford Centre for Molecular Sciences, New Chemistry Laboratory, University of Oxford, South Parks Road, Oxford OX1 3QT, United Kingdom; and <sup>‡</sup>Department of Chemistry and Biological Chemistry, Harvard University, 12 Oxford Street, Cambridge, MA 02138

Contributed by Martin Karplus, March 20, 2000

**Unfolding of proteins by forced stretching with atomic force microscopy or laser tweezer experiments complements more classical techniques using chemical denaturants or temperature. Forced unfolding is of particular interest for proteins that are under mechanical stress in their biological function. For  $\beta$ -sandwich proteins (a fibronectin type III and an immunoglobulin domain), both of which appear in the muscle protein titin, the results of stretching simulations show important differences from temperature-induced unfolding, but there are common features that point to the existence of folding cores. Intermediates detected by comparing unfolding with a biasing perturbation and a constant pulling force are not evident in temperature-induced unfolding. For an  $\alpha$ -helical domain ( $\alpha$ -spectrin), which forms part of the cytoskeleton, there is little commonality in the pathways from unfolding induced by stretching and temperature. Comparison of the forced unfolding of the two  $\beta$ -sandwich proteins and two  $\alpha$ -helical proteins (the  $\alpha$ -spectrin domain and an acyl-coenzyme A-binding protein) highlights important differences within and between protein classes that are related to the folding topologies and the relative stability of the various structural elements.**

The mechanism of protein folding and unfolding raises one of the fundamental questions of structural biology (1, 2). Recently, atomic force and laser tweezer experiments have focused attention on unfolding individual protein molecules with external forces (3–6). This approach provides single molecule information that has not been available previously. The results are ideal for analysis by molecular dynamics simulations, which also study the behavior of individual molecules (7, 8). Structural features of the forced unfolding pathways, which are not provided by presently available measurements, are obtained from biased molecular dynamic simulation for two  $\beta$ -sandwich proteins (Fig. 1A) and two  $\alpha$ -helical proteins (Fig. 1B). The differences in unfolding behavior can be correlated with the folding topology and stability of the secondary structural elements.

Because unfolding in solution leads to collapsed structures (tending to a random coil at high temperature), whereas forced unfolding, as clearly demonstrated by experiments and simulations, results in an extended chain, it is important to compare the pathways that occur with the two unfolding methods. This is of particular interest because an extrapolation to zero force of atomic force microscopy experiments for a  $\beta$ -sandwich protein has yielded results for the unfolding rate that correlate with the denaturation in solution by Gdm Cl (6). We have used high-temperature unfolding here, because it is better suited for simulations and has been shown to produce meaningful results when compared with denaturant-induced unfolding (9, 10). In temperature-induced unfolding, the effective energy of the relatively compact denatured state is of the order of tens of kcal/mol (1 cal = 4.184 J) above the native state, whereas it is found to be more than 300 kcal/mol higher in the extended conformation induced by simulated stretching. Forced unfolding

simulations (i.e., pulling the N- and C-terminal ends apart) and more classical temperature-induced unfolding simulations are reported for two  $\beta$ -sandwich proteins and two  $\alpha$ -helical proteins. Comparisons of the unfolding pathways show significant differences, which can be related to the topology and energetics.

## Methods

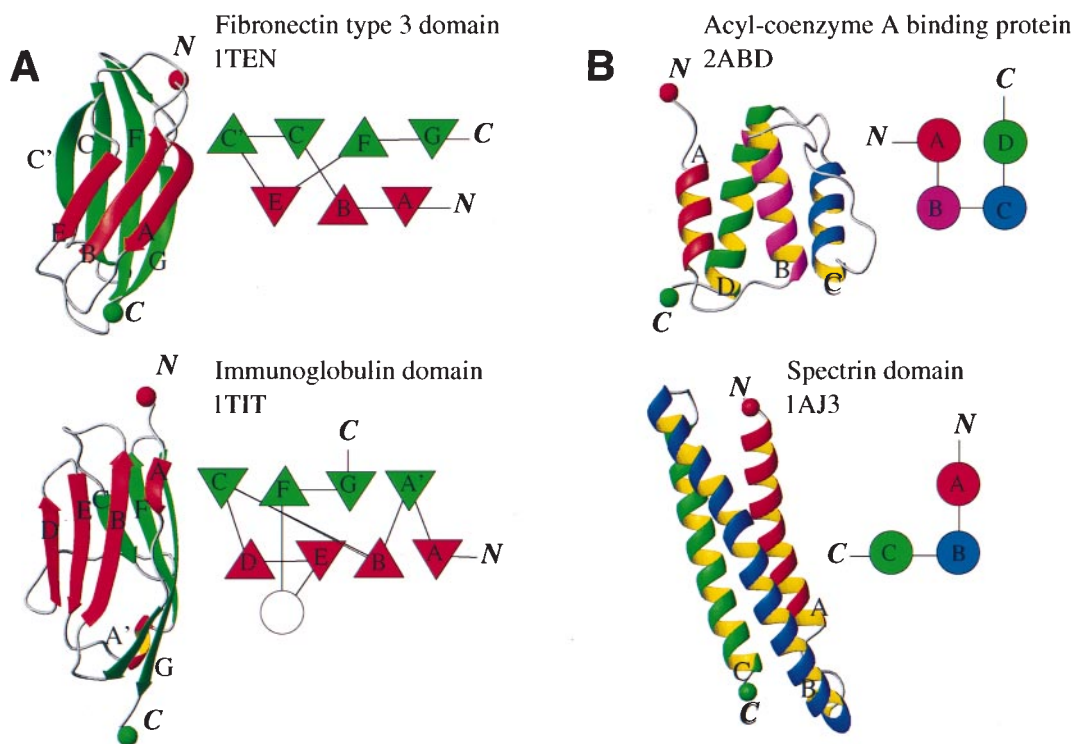
Molecular dynamics simulations were performed with a version of the CHARMM program (11) modified to include an external the bias force on the protein to accelerate unfolding (for details on the simulation methodology, see ref 8). A polar hydrogen model was used for the protein (12) and an implicit Gaussian model for the solvent (13). The implicit solvent provides an adiabatic solvent response that is important for the artificially fast unfolding simulated here; i.e., the simulated unfolding takes place on a nanosecond timescale, whereas most experiments require milliseconds or longer. If water relaxation times in the nanosecond to microsecond range are involved, explicit solvent treatments may give physically less meaningful results in nanosecond simulations that mimic millisecond experiments (3, 6) than the implicit solvent method. A reaction coordinate  $\rho(t)$  leading from the initial to the final state is chosen to have the form  $\rho(t) = r_{NC}^2(t)$ , where  $r_{NC}(t)$  is the distance between the N- and C-terminal atoms at time  $t$  (8). This reaction coordinate corresponds closely to that used in the forced unfolding experiments (3–6). The time-dependent bias force is equal, in absolute value, to  $\gamma(\rho - \rho_a) \sqrt{\rho}$ , if  $\rho(t) < \rho_a$  and to zero if  $\rho(t) \geq \rho_a$ . The quantity  $\rho_a(t)$  is the maximum value of  $\rho$  reached by the reaction coordinate at times less than or equal to  $t$ . This force is parallel to the vector  $\mathbf{r}_{NC}$  and pulls the N- and C-terminal atoms away from each other. The method accelerates the reaction while remaining close to the minimum energy surface with normal equilibrium fluctuations; the force has little effect on the short time dynamics. The force is nonzero only when spontaneous thermal fluctuations tend to decrease the reaction coordinate; it is large when the system is crossing a barrier and small otherwise. The parameter  $\gamma$  is adjusted so that unfolding occurs over a nanosecond timescale ( $\gamma$  is a scaling factor for the perturbation, with  $\gamma = 1$  corresponding to 0.0012 kcal·mol<sup>-1</sup> Å<sup>-4</sup>). Different values were tried, and the general behavior was not sensitive to the choice of  $\gamma$  within a certain range. Each starting configuration was energy minimized, then heated for 200 ps and equilibrated at  $T = 300$  K for 100 ps. An equilibrium simulation at least 1 ns in length was performed for each protein. Initial conditions for biased stretching simulations were taken from the equilibrium trajectory and

Abbreviation: PDB, Protein Data Bank.

<sup>§</sup>To whom reprint requests should be addressed. E-mail: marci@brel.u-strasbg.fr.

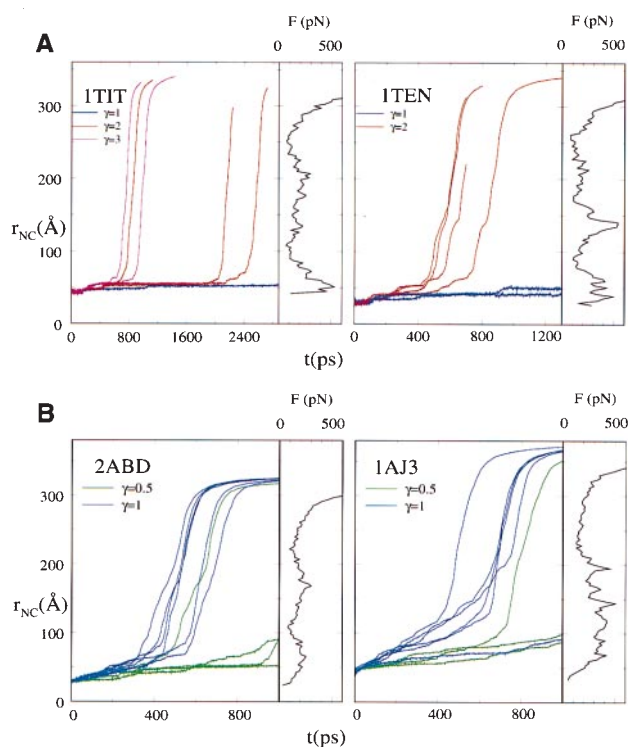
The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Article published online before print: *Proc. Natl. Acad. Sci. USA*, 10.1073/pnas.100124597. Article and publication date are at [www.pnas.org/cgi/doi/10.1073/pnas.100124597](http://www.pnas.org/cgi/doi/10.1073/pnas.100124597)

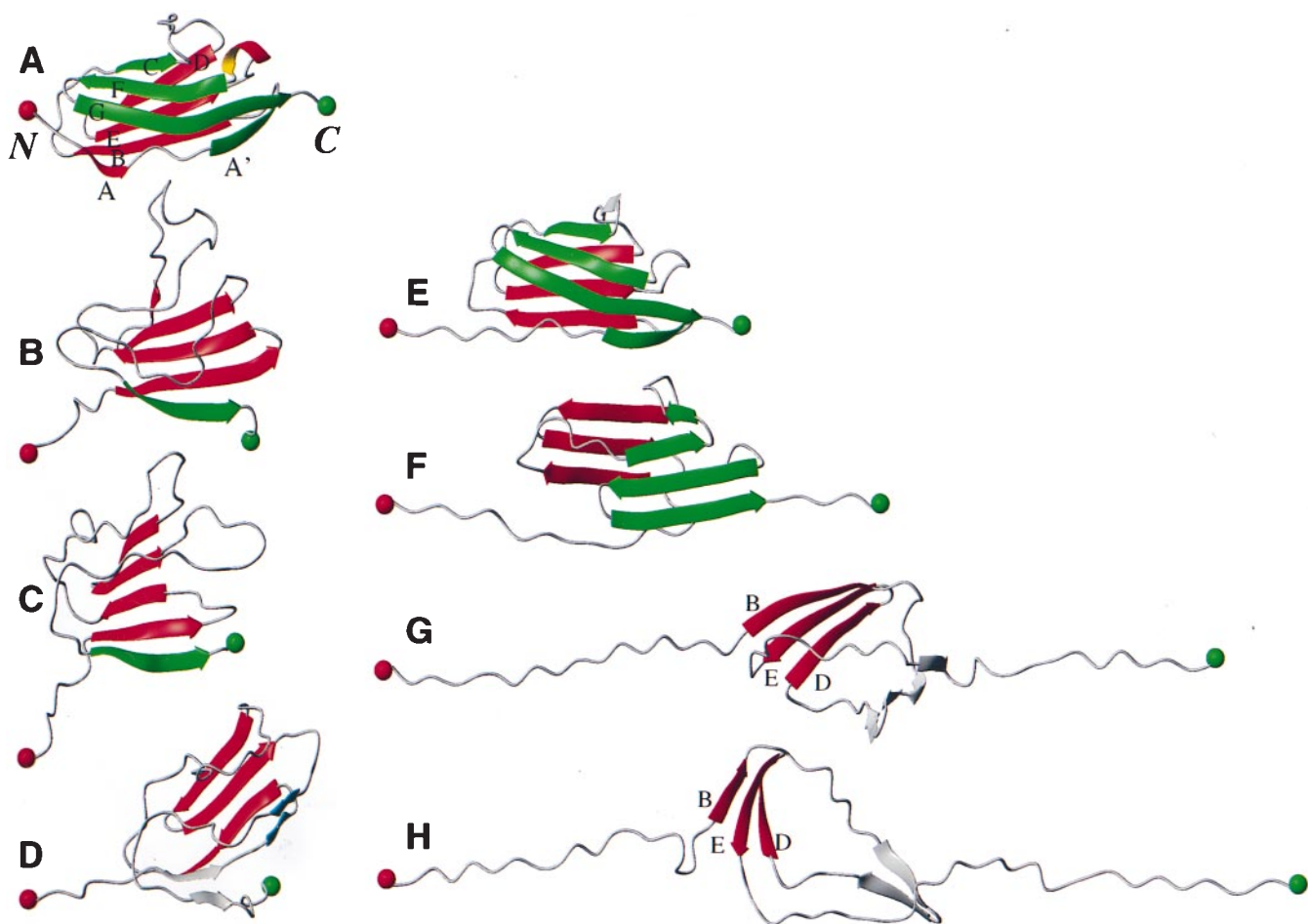


**Fig. 1.** Structures and topology of the protein studied. (A) Fn3 domain (1TEN) (32) and an Ig domain (1TIT) (33). (B)  $\alpha$ -Spectrin domain (1AJ3) (34) and the four-helix bundle protein acyl-coenzyme A-binding protein (2ABD) (35). The N- and C-terminal atoms are represented by red and green spheres, respectively; the same colors are used to represent  $\beta$ -strands belonging to the N- and C-terminal sheets of the  $\beta$ -sandwich proteins. In the schematic topology, triangles represent  $\beta$ -strands, and circles represent  $\alpha$ -helices (and also a turn of  $3_{10}$ -helix in the Ig domain). The experimental structures are drawn with the program MolMO (36) and the folding topology, with the symbols used in the program (37).

spaced at least 100 ps apart to ensure statistical independence. Multiple simulations were calculated, as required by the chaotic nature of protein dynamics (14), to determine the common elements of the pathways for each system. All simulations were performed in the canonical ensemble using the Nosé-Hoover (15, 16) thermostat; the “mass” of the thermostat is chosen to be large enough so that local heating is negligible. A time step of 2 fs was used, and coordinates were saved every 500 steps (1 ps). During the equilibration period, all the simulated systems are relatively stable. The average  $C_{\alpha}$  rms deviation from the experimental structure during the last 1 ns of equilibration was 2.9 Å for the spectrin domain [Protein Data Bank (PDB) ID 1AJ3], 3.0 Å for the acyl-coenzyme A-binding domain (PDB ID 2ABD), 2.4 Å for the immunoglobulin domain (PDB ID 1TIT), and 1.9 Å for the fibronectin type III domain (PDB ID 1TEN). Forced unfolding simulations have been made for three other Fn3 domains, the 9th and 10th repeats from human fibronectin (PDB ID 1FNF) (17), the glycosidase inhibitor (PDB ID 1HOE) (18), and two Ig domains, the titin Ig domain from human cardiac muscle (5th in the M-line region) (PDB ID 1TNM) (19) and the twitchin Ig domain from nematode muscle (18th in the A-band region) (PDB ID 1WIT) (20); they give corresponding results (ref. 8 and unpublished data). To complement these results, unfolding with a constant pulling force (8) was also studied. A constant force alters the free energy surface so that the completely stretched conformation becomes the most stable one. It accelerates unfolding by lowering the barriers between the native and the completely stretched state (21, 22). A constant force large enough that the protein unfolds on a computationally accessible timescale ( $\sim 1$  ns) is likely to lead to pathways that differ more from those followed on the experimental timescale than the results obtained with the variable biasing force.



**Fig. 2.** Unfolding reaction coordinate  $r_{NC}$  (in Å) as a function of time (in picoseconds). Summary of the results for (A) the two  $\beta$ -sandwich domains and (B) the two  $\alpha$ -helical domains. For a definition of the parameter  $\gamma$ , see *Methods*. The external force applied to unfold the domains is also plotted as a function of  $r_{NC}$ ; the force values are averaged over all trajectories and over 3-Å intervals of  $r_{NC}$ .



**Fig. 3.** Pathway for the unfolding of the 1TIT domain. (A) X-ray structure,  $r_{NC} = 45 \text{ \AA}$ ; B–D temperature-induced unfolding pathway; (B) 0.6 ns,  $r_{NC} = 34 \text{ \AA}$ , (C) 1.5 ns,  $r_{NC} = 37 \text{ \AA}$ ; (D) 4.8 ns,  $r_{NC} = 38 \text{ \AA}$ . E–H forced unfolding pathway; (E)  $r_{NC} = 54 \text{ \AA}$ ; (F)  $r_{NC} = 75 \text{ \AA}$ ; (G)  $r_{NC} = 126 \text{ \AA}$ ; (H)  $r_{NC} = 138 \text{ \AA}$ . Native elements of secondary structure are shown by using the same color code as in Fig. 1; nonnative are in grey.

Nevertheless, a constant force can be a useful way of isolating (kinetic) intermediates.

To compare with the results obtained by an external force directed along the reaction coordinate, high-temperature (400 K), simulations were used to obtain unfolding on the nanosecond timescale. They were started from native conformations equilibrated at room temperature by scaling the velocities and setting the Nosé–Hoover thermostat to the desired temperature. The same potential function (without the biasing term) as for the biased unfolding at 300 K was used. On a 6-ns timescale, only partial unfolding occurs at 400 K, and the unfolding is very fast (less than 500 ps) at 500 K; we report results for the 450 K simulations.

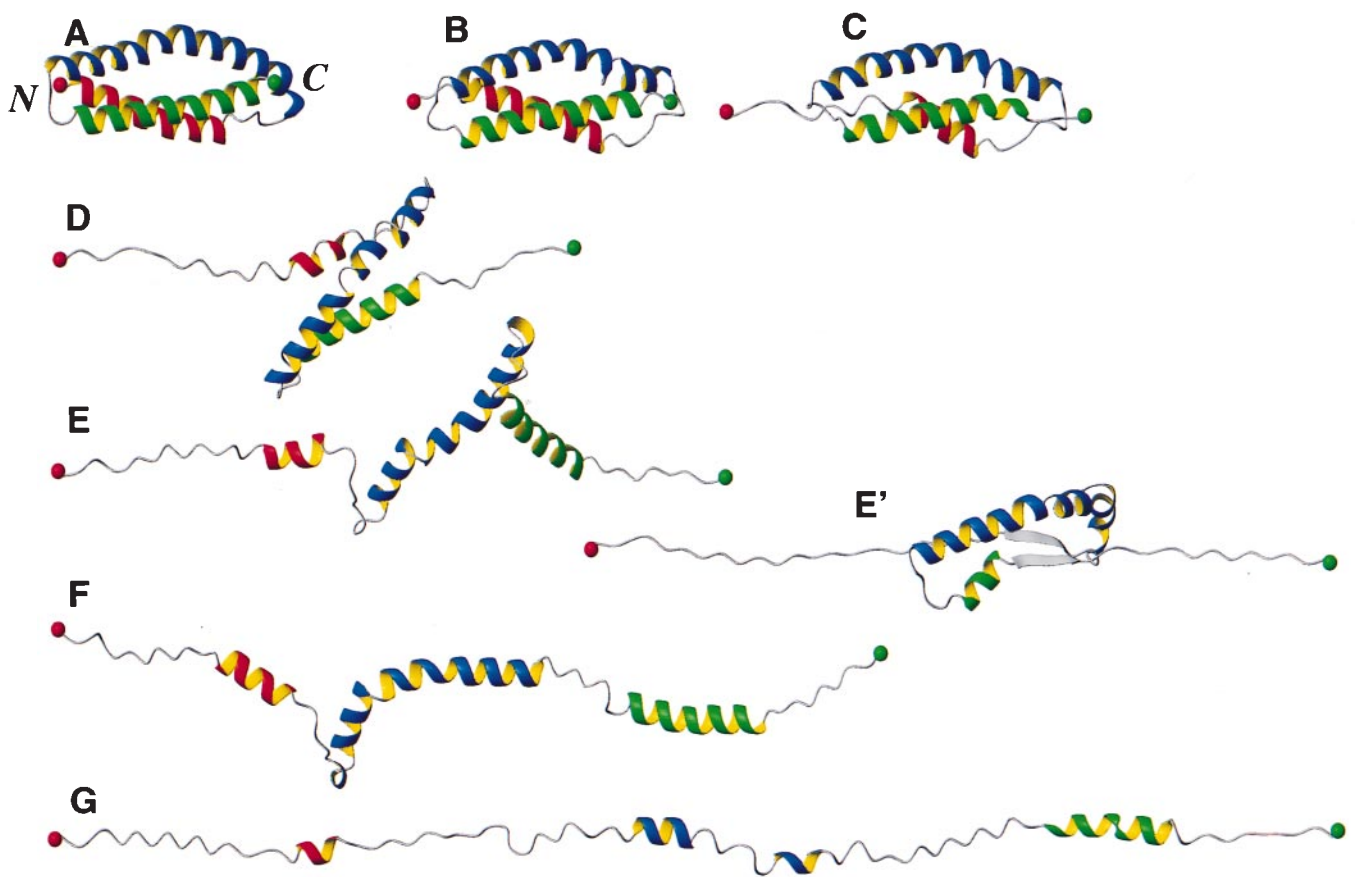
### Results and Discussion

Fig. 2 shows the end-to-end distance,  $r_{NC}$ , as a function of simulation time for the four proteins with the time-dependent external perturbation. Although individual trajectories for each protein vary as to the time of onset of unfolding, as expected, the general characteristics are preserved. Changes in the slope of the end-to-end distance as a function of time reveal bottlenecks in the induced unfolding pathway that may correspond to intermediates (see below). The two  $\beta$ -sandwich proteins show significant differences. The fibronectin type III (Fn3) domain (1TEN) unfolds in two major steps, whereas the Ig domain (1TIT) has only a single critical step close to the native state, although the early unfolding behavior is rather complex for both

systems; a single barrier was found for both systems with a somewhat different model (7, 23). The average unfolding force as a function of  $r_{NC}$  is shown on the right-hand side of the plots; the force peaks confirm that there are two dominant barriers for the Fn3 domain and only one for the Ig domain. The maximum local force found in the simulations is of the order of 400 pN, somewhat larger than that used in the experiments because the unfolding simulations are done on a shorter timescale (8). Results of constant force simulations indicate that forces less than 250 pN are not enough to unfold the domains on the nanosecond timescale; a larger force is required to unfold the Ig domain than the Fn3 domain, in qualitative agreement with experiments (24).

For the Ig domain, the single large bottleneck to unfolding occurs at rather small ( $\approx 6 \text{ \AA}$ ) extensions (Fig. 3), somewhat larger than the original experimental estimate of 2.5  $\text{\AA}$  (6) for the position of the transition state for unfolding; the simulation value is close to that suggested by more recent experiments (ref. 25; see also ref. 26). The critical event appears to be the breaking of the “seal” formed by the hydrogen bonds between strands A and B, and A’ and G of the two sheets (see Fig. 1); it involves pulling the N-terminal portion (strands A and A’) away from B and G (Fig. 3 E and F). After this, the domain unfolds along an essentially unique pathway with no additional bottlenecks. The two sheets now are able to slip relative to each other (Fig. 3G) and come apart, with B, E, and D lasting longer than C and F; interestingly, a nonnative hairpin forms in the late stage of





**Fig. 4.** Pathway for the unfolding of the 1AJ3 domain. (A) X-ray structure,  $r_{NC} = 42 \text{ \AA}$ ; (B)  $r_{NC} = 50 \text{ \AA}$ ; (C)  $r_{NC} = 70 \text{ \AA}$ ; (D)  $r_{NC} = 100 \text{ \AA}$ ; (E)  $r_{NC} = 130 \text{ \AA}$ ; (E')  $r_{NC} = 142 \text{ \AA}$ ; (F)  $r_{NC} = 160 \text{ \AA}$ ; (G)  $r_{NC} = 250 \text{ \AA}$ . E is a stable intermediate found in one set of unfolding pathways.

unfolding (Fig. 3H) in what was the C-terminal sheet. For the Fn3 domains, the first barrier is associated with the relative “slipping” of the two sheets with respect to each other and partial disruption of the hydrophobic core (see figure 9 of ref. 8, which presents the unfolding pathways of the ninth Fn3 domain of fibronectin that are very similar to those obtained here for 1TEN). The native secondary structure is generally well preserved up to  $80 \text{ \AA}$ , 2.5 times the native “length.” In the region corresponding to the second bottleneck at  $r_{NC} \approx 140 \text{ \AA}$ , where the molecule is stretched to approximately four times its “length,” there are two pathways (8): in one, strands A and B of the N-terminal sheet stretch completely, whereas there is little change in the C-terminal sheet; in the other, both ends (strands A and G) stretch. The bottleneck at  $r_{NC} \sim 140 \text{ \AA}$  is related to the simultaneous breaking of the hydrogen bonds between  $\beta$ -strands F and G and between C and F. The CC' hairpin is always the last secondary structural element to disappear. Because the second barrier is significantly higher for one pathway than for the other, it is possible that partially unfolded intermediates ( $r_{NC} \approx 140 \text{ \AA}$ ) would be observed if a low force were applied, as in constant force simulations (see Fig. 8 of ref. 8). The difference in unfolding behavior between the Fn3 and Ig domains can be understood from their folding topologies (Fig. 1A); that is, in Fn3, the N and C portions of the two  $\beta$ -sheets are separated, whereas they are intertwined in Ig.

The nanosecond unfolding simulations of the  $\alpha$ -helical domains occur at a somewhat smaller force than do the  $\beta$ -sandwich domains (Fig. 2). Moreover, the  $\alpha$ -helical domains show more gradual unfolding. This suggests that even smaller forces would be required at the several orders of magnitude slower pulling

speed used experimentally (24). The results of Rief *et al.* (27) for  $\alpha$ -spectrin yield a sawtooth pattern in the force-extension curves corresponding to that observed for Ig and Fn3 domains but with smaller force peaks. As in the case of the Fn3 domains, a plateau at an extension between  $110 \text{ \AA}$  and  $150 \text{ \AA}$  is found in some trajectories when a constant force of  $200 \text{ pN}$  is applied, suggestive of an intermediate along one set of pathways; it consists of extended N- and C-terminal strands with all of helix B and part of helix C locked together by a nonnative  $\beta$ -hairpin (Fig. 4E'). Figs. 4 and 5 show representative unfolding pathways for the two  $\alpha$ -helical domains (Fig. 1B). For the 1AJ3 domain, the initial effect of the stretching is limited to the N- and C-terminal parts of helices A and C, respectively; they tend to become longer  $3_{10}$  helices, with the N-terminal helix unfolding first (Fig. 4B and C). The increase in the slope of  $r_{NC}$  in the region with  $r_{NC}$  between  $110 \text{ \AA}$  and  $150 \text{ \AA}$  (Fig. 2B) occurs when the tertiary arrangement of the helices is disrupted (Fig. 4D to E), even though a significant fraction of the  $\alpha$ -helical content is preserved; e.g., the conformation in Fig. 4F, which is almost four times longer than the native state, still has 59%  $\alpha$ -helix relative to 88% in the native structure. The last helical element to unfold involves residues 82 to 89 in helix C, which, according to two different secondary structure prediction algorithms (28, 29), has the highest helical propensity. The unfolding pathway of the 2ABD is different from that of the 1AJ3 domain. It shows a large number of metastable states with  $r_{NC}$  in the range  $70$ – $250 \text{ \AA}$  when a low ( $150$ -pN) constant force is applied. Although there is some unfolding of the N-terminal helix, the initial stages (Fig. 5A–C) involve primarily small tertiary rearrangements of the helices. The rate of unfolding increases when the three C-terminal

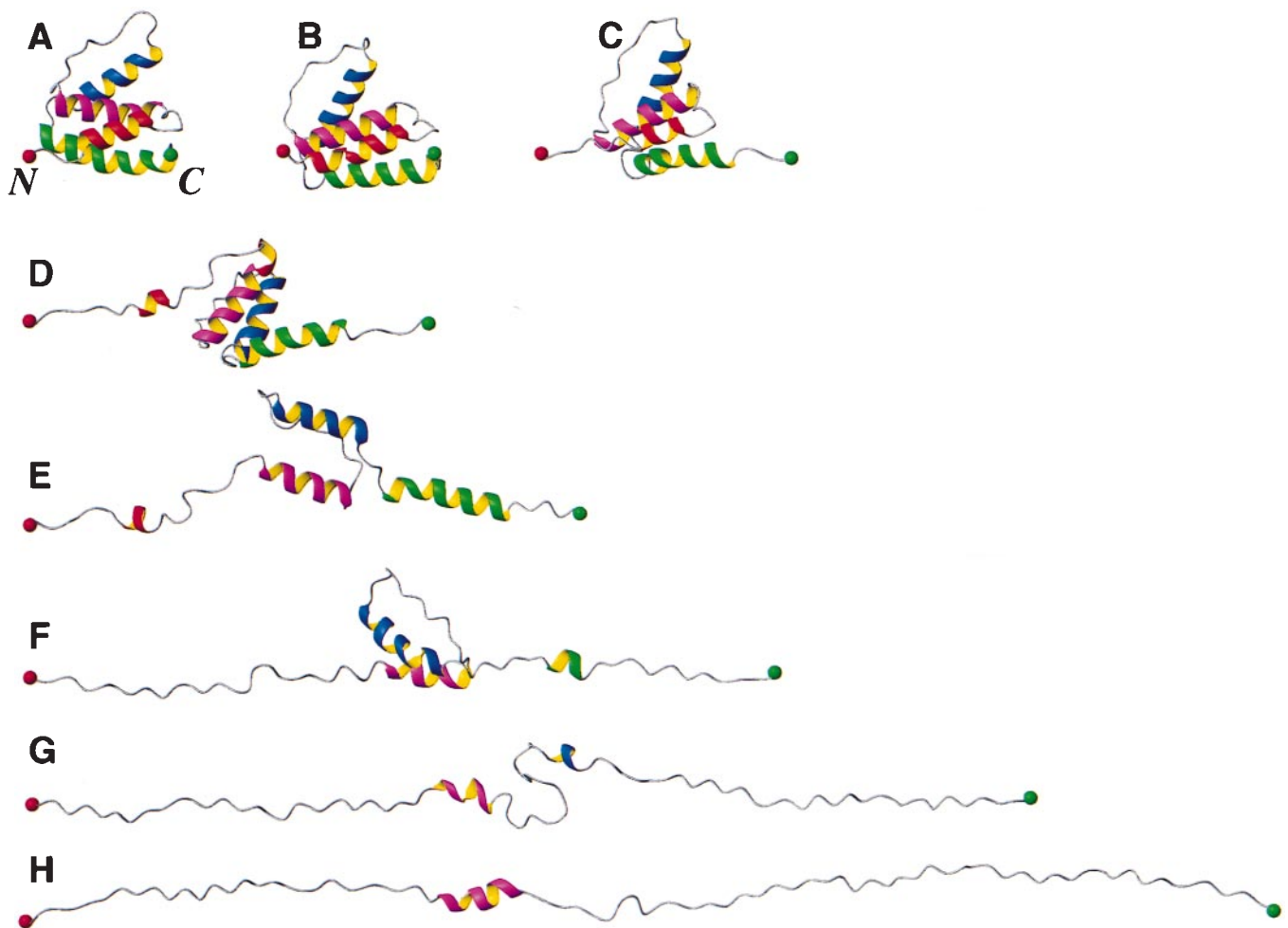


Fig. 5. Pathway for the unfolding of 2ABD. (A) X-ray structure,  $r_{NC} = 28.4 \text{ \AA}$ ; (B)  $r_{NC} = 30 \text{ \AA}$ ; (C)  $r_{NC} = 50 \text{ \AA}$ ; (D)  $r_{NC} = 80 \text{ \AA}$ ; (E)  $r_{NC} = 110 \text{ \AA}$ ; (F)  $r_{NC} = 150 \text{ \AA}$ ; (G)  $r_{NC} = 200 \text{ \AA}$ ; (H)  $r_{NC} = 250 \text{ \AA}$ .

helices have come apart and begin to unravel ( $r_{NC} \geq 80 \text{ \AA}$ ). A core formed by residues 32 to 54, which includes part of the B and C helices and the long loop joining them, resists most. This region has a lower helical propensity than helices A and D (29), suggesting that tertiary interactions are important in its stabilization; e.g., helices B and C are most buried (more than 80% of the surface area) in the native state and only become as exposed as helices A and D at  $r_{NC} > 200 \text{ \AA}$ . Kragelund *et al.* (30) have concluded from solution folding studies that the interactions of helices A and D are important in the transition state. Here they separate first, because the forces are applied at the N- and C-terminal ends of helices A and D, respectively. The results for 2ABD contrast with those for the 1AJ3 domain, where a large fraction of the secondary structure survives the disruption of the native tertiary interactions.

To determine the thermal unfolding pathways, we consider the 6-ns simulations performed for each of the proteins at 450 K. We use the same protein and solvent model as for the forced unfolding simulations and in previous temperature-induced unfolding studies (13). At 450 K, the domains unfold smoothly, and  $r_{NC}$  decreases from its native value, very different from the increase that occurs in  $r_{NC}$  during forced unfolding. An essential element of temperature-induced unfolding of both  $\beta$ -sheet domains is that the earliest events correspond to the destruction of certain  $\beta$ -strands and the formation of a coil region in parts of the sequence. This is the case for strands F

and C in the Ig domain (Fig. 3B) and strands A and B in the Fn3 domain. Unlike the forced unfolding, there is no tendency of the N- and C-terminal ends to become extended early and move apart. In fact, in the Ig domain, the breaking of the “seal” between the two sheets, an early event when stretching the protein, does not occur until most of the protein is unfolded at high temperature. However, in both types of unfolding simulations (i.e., with stretching and with temperature), the same  $\beta$ -sheet elements are unfolded last. They are the three-stranded (BED) sheet in the Ig domain and the four-stranded (C’CFG) sheet in the Fn3 domain. Moreover, in the latter, the hairpins formed by strands F and G, and, independently, C and C’, which are common to both forced unfolding pathways, are kinetically the most stable portions of the proteins during the temperature-induced unfolding.

For the  $\alpha$ -helical domains, the stretched and temperature-induced unfolding pathways are very different: the helices become frayed and fragmented and unfold much earlier at high temperature than in the stretching pathways, where the unfolding of helices is more cooperative, and, once helices are broken, they do not reform. In temperature-induced unfolding, there is a tendency to form transient nonnative  $\beta$ -sheets, which stabilize the collapsed form.

### Conclusion

Comparison of the forced unfolding of two protein classes (all- $\beta$ -sandwich proteins and all- $\alpha$ -helix bundle proteins) have shown

significantly different behavior, both within a protein class and for the different classes. Important differences are also found between the unfolding induced by high temperature and by external pulling forces. The result is interpreted in terms of the type of perturbation, the folding topologies, the nature of the secondary and tertiary interactions, and the relative stability of the various structural elements. The complex behavior observed in the simulations contrasts with the simple sawtooth profile observed in the first generation of atomic force microscopy (AFM) unfolding experiments; they revealed only the onset of the unfolding of individual domains in multidomain proteins. It is encouraging that recent forced unfolding experiments of an engineered polyprotein made of identical Ig domains (25) show a previously undetected component, which is attributed to an intermediate. Thus, it is likely that

improved AFM experiments will show more of the features observed in the simulation of four different proteins. When combined with protein engineering experiments of the same proteins (30) and simulations for the mutated systems, a detailed microscopic description of the unfolding reaction under different conditions will become available.

We thank Jane Clarke for comments on the manuscript. Support by the Ministère de l'Éducation Nationale de la Recherche et de la Technologie and by the Centre National de la Recherche Scientifique to ESA 7006 (Strasbourg) and by the National Science Foundation (Harvard University) is acknowledged. The computer time used for this work was provided by the Informatique et Calcul Parallèle de Strasbourg. E.P. was supported by the European Union through a Marie Curie fellowship.

- Dobson, C. M., Šali, A. & Karplus, M. (1998) *Angew. Chem. Int. Ed. Eng.* **37**, 868–893.
- Fersht, A. R. (1999) *Structure and Mechanism in Protein Science: A Guide to Enzyme Catalysis and Protein Folding* (Freeman, New York).
- Rief, M., Gautel, M., Oesterhelt, F., Fernandez, J. M. & Gaub, H. E. (1997) *Science* **276**, 1109–1112.
- Kellermayer, M. S., Smith, S. B., Granzier, H. L. & Bustamante, C. (1997) *Science* **276**, 1112–1116.
- Mehta, A. D., Rief, M., Spudich, J. A., Smith, D. A. & Simmons, R. M. (1999) *Science* **283**, 1689–1695.
- Carrion-Vasquez, M., Oberhauser, A. F., Fowler, S. B., Marszalek, P. E., Broedel, S. E., Clarke, J. & Fernandez, J. M. (1999). *Proc. Natl. Acad. Sci. USA* **96**, 3694–3699.
- Lu, H., Isralewitz, B., Krammer, A., Vogel, V. & Schulten, K. (1998) *Biophys. J.* **75**, 662–671.
- Paci, E. & Karplus, M. (1999) *J. Mol. Biol.* **288**, 441–459.
- Lazaridis, T. & Karplus, M. (1997). *Science* **278**, 1928–1931.
- Ladurner, A. G., Itzhaki, L. S. Daggett, V. & Fersht, A. R. (1998) *Proc. Natl. Acad. Sci. USA* **95**, 8473–8478.
- Brooks, B. R., Brucoleri, R. E. Olafson, B. D., States, D. J., Swaminathan, S. & Karplus, M. (1983) *J. Comp. Chem.* **4**, 187–217.
- Neria, E., Fischer, S. & Karplus, M. (1996). *J. Chem. Phys.* **105**, 1902–1921.
- Lazaridis, T. & Karplus, M. (1999) *Proteins* **35**, 133–152.
- Braxenthaler, M., Unger, R., Auerbach, D., Given, J. A. & Moul, J. (1997) *Proteins* **29**, 417–425.
- Nosé, S. (1984) *Mol. Phys.* **52**, 255–268.
- Hoover, W. G. (1985) *Phys. Rev. A.* **31**, 1695–1697.
- Leahy, D. J., Aukhil, I. & Erickson, H. P. (1996) *Cell* **84**, 155–164.
- Pflugrath, J. W., Wiegand, G., Huber, R. & Vertesy, L. (1986) *J. Mol. Biol.* **189**, 383–386.
- Pfuhl, M. & Pastore, A. (1995) *Structure (London)* **3**, 391–401.
- Fong, S., Hamill, S. J., Proctor, M., Freund, S. M., Benian, G. M., Chothia, C., Bycroft, M. & Clarke, J. (1996) *J. Mol. Biol.* **264**, 624–639.
- Evans, E. & Ritchie, K. (1997) *Biophys. J.* **72**, 1541–1555.
- Evans, E. (1999) *Faraday Discuss. Chem. Soc.* **111**, 1–16.
- Krammer, A., Lu, H., Isralewitz, B., Schulten, K. & Vogel, V. (1999) *Proc. Natl. Acad. Sci. USA* **96**, 1351–1356.
- Rief, M., Gautel, M., Schemmel, A. & Gaub, H. E. (1998) *Biophys. J.* **75**, 3008–3014.
- Marszalek, P. E., Lu, H., Li, H., Carrion-Vasquez, M., Oberhauser, A. F., Schulten, K. & Fernandez, J. M. (1999) *Nature (London)* **402**, 100–103.
- Lu, M. & Schulten, K. (1999) *Chem. Phys.* **247**, 141–153.
- Rief, M., Pascual, J., Saraste, M. & Gaub, H. E. (1999) *J. Mol. Biol.* **286**, 553–561.
- Muñoz, V. & Serrano, L. (1997) *Biopolymers* **41**, 495–509.
- Chandonia, J.-M. & Karplus, M. (1999) *Proteins* **35**, 293–306.
- Kragelund, B. B., Osmark, P., Neergaard, T. B., Schiødt, J., Kristiansen, K., Knudsen, J. & Poulsen, F. M. (1999). *Nat. Struct. Biol.* **6**, 594–601.
- Hamill, S. J., Steward, A. & Clarke, J. (2000) *J. Mol. Biol.* **297**, 165–178.
- Leahy, D. J., Hendrickson, W. A., Aukhil, I. & Erickson, H. P. (1992) *Science* **258**, 987–991.
- Improta, S., Politou, A. S. & Pastore, A. (1996) *Structure (London)* **4**, 323–337.
- Pascual, J., Pfuhl, M., Walther, D., Saraste, M. & Nilges, M. (1997) *J. Mol. Biol.* **273**, 740–751.
- Andersen, K. V. & Poulsen, F. M. (1993). *J. Biomol. NMR* **3**, 271–284.
- Koradi, R., Billeter, M. & Wüthrich, K. (1996) *J. Mol. Graphics* **14**, 51–55.
- Flores, T. P., Moss, D. S. & Thornton, J. M. (1997) *Protein Eng.* **7**, 31–37.