Mechanical stability of proteins

A. M. Gabovich[1,a)] and Mai Suan Li[2,b)]
[1]*Institute of Physics, Nauka Avenue 46, Kiev 03680, Ukraine*
[2]*Institute of Physics, Polish Academy of Sciences, Al. Lotników 32/46, 02-668 Warsaw, Poland*

(Received 2 April 2009; accepted 12 June 2009; published online 14 July 2009)

A number of experiments and experimentally based simulations showed that β-proteins are mechanically more stable than α-proteins. However, the theory that might explain this evidence is still lacking. In this paper we have developed a simple elastic theory, which allows to estimate critical forces for stretching both kinds of proteins. It has been shown that unfolding of β-proteins does really require notably higher forces as compared to the stretching of α-proteins. © *2009 American Institute of Physics*. [DOI: 10.1063/1.3170940]

I. INTRODUCTION

The last 10 years witnessed an intense activity in single-molecule force spectroscopy experiments in detecting inter- and intramolecular forces of biological systems to understand their functions and structures. Much of the research has been focused on elastic properties of proteins, DNA, and RNA, i.e., their response to an external force, following the seminal papers by Rief *et al.*,[1] and Tskhovrebova *et al.*[2] The main advantage of this technique is its ability not only to decipher the unfolding free energy landscape of biomolecules[3–6] but also to probe their mechanical stability. Understanding the resistance of proteins to an external force is important because many processes in living systems, such as cell division, locomotion, and enzyme activity, depend critically on protein conformational changes and mechanical rigidity.[3] Knowledge gained from the integration of physical, biological, and chemical studies would be useful for potential applications in material design, nanotechnology, and medicine.[7]

Experiments and simulations[4,8] have shown that β-proteins are mechanically more stable than α-rich proteins. For demonstration, we present experimental results for unfolding forces f_u as a function of the contact order[9] (Fig. 1). Since the correlation level between these two quantities is high ($R=0.74$), α-proteins are clearly less stable compared to β-proteins. At room temperature, typical forces, needed to unfold β-proteins at pulling speeds $v_p \sim 100$ nm/s are $f_u \sim 100$ pN, whereas helix proteins unfold under $f_u \sim 10$ pN. In the equilibrium case ($v_p \to 0$), at $T=0$ the critical unfolding force may be estimated as $f_u^{\mathrm{eq}} \approx \epsilon_{\mathrm{H}}/x_u$, where ϵ_{H} and x_u are the hydrogen bond energy and a distance between the native state and transition state, respectively. Since β-proteins have x_u larger than that of α-proteins,[4] they are expected more stable. Taking $\epsilon_{\mathrm{H}}=1-5$ kcal/mol, for β-proteins which have $x_u \approx 3$ Å,[4,10] we have $f_u^{\mathrm{eq}} \approx 25-125$ pN. For helix proteins such as the α-spectrin with $x_u=15$ Å,[29] $f_u^{\mathrm{eq}} \approx 4-20$ pN. One can estimate the effect of thermal fluctuations on reducing the critical force, but in this paper we deal with the $T=0$ case only. Overall, f_u^{eq} of β-proteins is about one order of magnitude larger than that of α-proteins.

To the best of our knowledge, a contansive theory of mechanical stability of β- and α-proteins is not yet available. Our goal is to develop a simple theory which is able to explain why the former are more stable. Our idea is to approximate a β-protein by a zigzag elastic model, while a α-protein, by a simple spring. Using the elastic theory, we derived the expression for f_c^{eq} and showed that β-proteins are more stable than α-proteins.

II. ELASTIC THEORY FOR β-PROTEINS

An example of β-proteins is displayed in Fig. 2(a), where β-strands are represented by arrows. In our model, β-strands are replaced by tubes (blue lines). Middle points of turns are the ends of these tubes. One end of the first and last

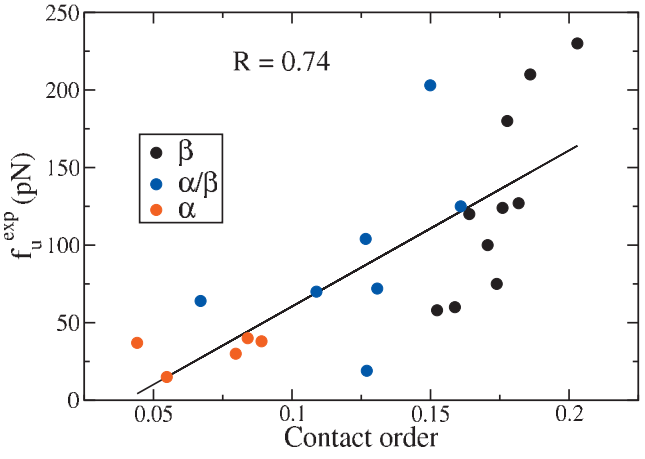

FIG. 1. Dependence of experimental values of unfolding force f_u^{exp} on the contact order for 22 proteins. The unfolding force is define as the highest peak in the force-extension curve (Ref. 4). Results are shown for 10 β-(PDB ID: 1TIT, (Ref. 10) 1G1C, (Ref. 11) 1WIT, (Ref. 12) 1TEN, (Ref. 13) 1FNF, (Refs. 14 and 15) 1FNH, (Ref. 15) 1KSR, (Refs. 16–18) 1RSY, (Ref. 19) 1NCT, (Refs. 20 and 21) and 1OWW (Ref. 15)), 7 α/β-(1HZ6, (Ref. 22) 1UBQ, (Ref. 23) 1BNI, (Ref. 24) 1HFR, (Ref. 25) 1B9C, (Ref. 26) 1B6I, (Ref. 27) and 1RNH (Ref. 28)), and 5 α-proteins (1AJ3, (Ref. 29) 1HCI, (Refs. 29 and 30) 1CFC, (Ref. 19) 1VCS, (Ref. 31) and 1N11 (Refs. 32 and 33)). The correlation level between f_u and the contact order is $R=0.74$.

a)Electronic mail: alexander.gabovich@gmail.com.
b)Electronic mail: masli@ifpan.edu.pl.

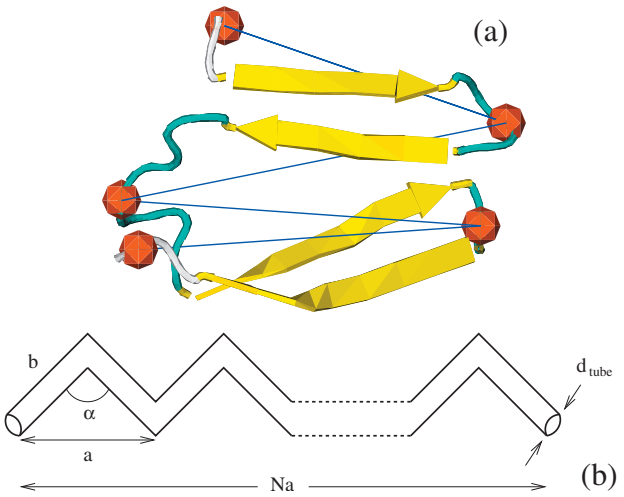

FIG. 2. (a) Fragment of the titin domain I27 (PDB ID: 1TIT). In our zigzag model each β-strand is replaced by a tube (in blue). (b) Schematic plot for the zigzag model.

tubes is the first and last amino acid, respectively. Thus, we simulate a β-strand structure by a zigzag planar chain consisting of identical hinged rodlike segments of the length b. For simplicity, the twist of the β-sheet is neglected. The opening angles α between adjacent segments are considered equal [Fig. 2(b)]. The distance between neighboring hinges is $a=2b\sin(\alpha/2)$. We assume that α is determined by interstrand hydrogen bonds[34] modeled by elastic tubes; connecting adjacent hinges, the true internal structure of which does not need to be specified in our approach. This viewpoint is in agreement, e.g., with the molecular dynamics simulations of force-induced titin unfolding, which show that interstrand hydrogen bonds are broken during the process.[35] One should note that a similar model has been applied earlier[36] to fit nonlinear empirical force-extension curves obtained in atomic-force-microscopy experiments for bovine carbonic anhydrase II, Q253C.

The strain $\Delta a/a$ in the tube is proportional to the stress $\sigma=F/S$, according to the Hooke's law

$$\frac{\Delta a}{a}=\frac{F}{SE}=\frac{4F}{d_{\text{tube}}^2E}. \tag{1}$$

Here E is the Young modulus (modulus of elongation), F is the pulling force, $S=\pi d_{\text{tube}}^2/4$ is the effective cross section, and d_{tube} is the effective diameter of the tube. The last quantity may be estimated as several angstroms. A distance between outermost ends of the segments is $L=Na=2bN\sin(\alpha/2)$, where N is the number of segments. Naturally, the overall strain $\Delta L/L$ equals to $\Delta a/a$. After the pulling process is finished, the final length of the protein chain becomes $L_{\max}=2bN$. The difference between L and $L_{\max}$ is the maximal overall elongation $\Delta L_{\lim}=L_{\max}-L=2bN[1-\sin(\alpha/2)]$.

Therefore, the limiting strain value $\Delta L_{\lim}/L$ for the β-protein stretching is connected to the limiting unfolding force f_u^{eq}, namely, $\Delta L_{\lim}/L=f_u^{\text{eq}}/SE$, so that

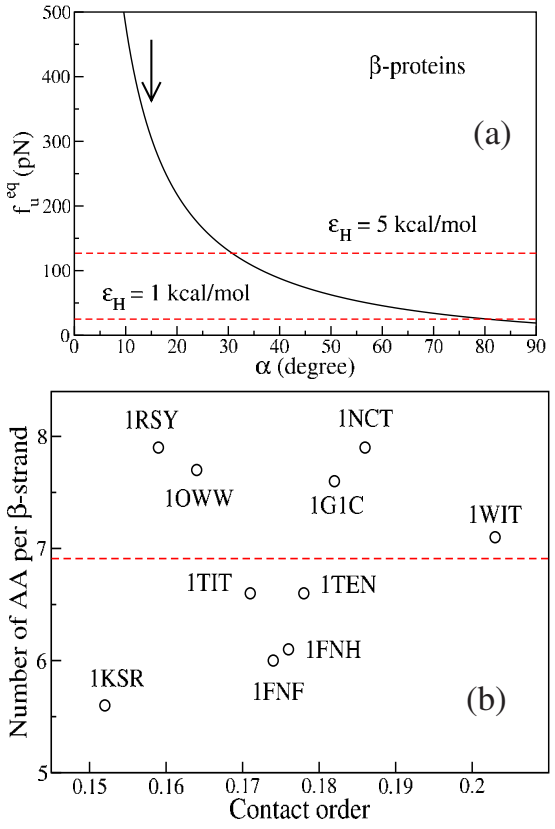

FIG. 3. (a) Dependence of f_u^{eq} on angle α for β-proteins, where we used $E=200$ MPa and $d_{\text{tube}}=5.4$ Å. Arrow refers to the most probable value of $\alpha=15°$ which corresponds to $f_u^{\text{eq}}\approx 300$ pN. The red lines denote $f_u^{\text{eq}}=25$ and 125 pN obtained from the free energy landscape picture using $\epsilon_{\text{H}}=1$ and 5 kcal/mol, respectively. (b) The averaged number of AA, per β-strand plotted as a function of contact order for ten β-proteins. Their PDB IDs are shown next to the data points. The horizontal red line corresponds to the averaged (over all proteins) value which is about 7.

$$f_u^{\text{eq}}=\frac{\pi d_{\text{tube}}^2E}{4}\left[\frac{1-\sin(\alpha/2)}{\sin(\alpha/2)}\right]. \tag{2}$$

Experimental values of E are not known for proteins. For quantitative estimation of f_u^{eq}, we take a typical theoretical value $E\approx 0.2$ GPa.[37] Previous analyses[38] of the native structures of proteins have shown that a protein backbone may be thought of approximately as a tube of diameter $d_{\text{tube}}\approx 5.4$ Å. Using these parameters we plot f_u^{eq} as a function of the angle α in Fig. 3(a). f_u^{eq} decreases with α because the increase of this angle lowers the effective number of native contacts or the interaction between two β-strands in the native conformation. One can estimate the most probable value of α for proteins as follows. The detailed analysis of native conformations of ten β-proteins presented in Fig. 1 shows that, on the average, one strands contains seven amino acids [Fig. 3(b)]. Assuming that a turn has four amino acids, each segment in our zigzag model is then consisted of nine amino acids (seven from a strand and two from a turn). Therefore its length $b\approx 8c\approx 30.4$ Å, where $c=3.8$ Å, is the distance between two nearest neighboring C_α. Suppose that the native contact exists even between two ends of neighboring segments, then the distance a [Fig. 2(b)] should be less than or equal to some cutoff distance d_c for native contacts. For rough estimation we take $a=d_c=8$ Å. Then $\sin(\alpha/2)$

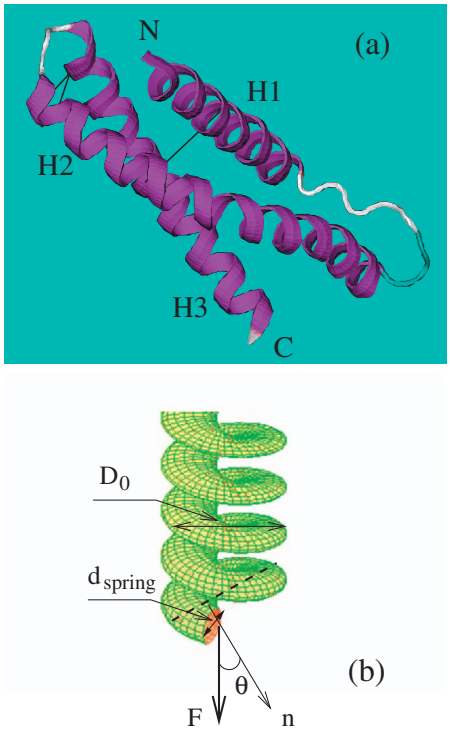

FIG. 4. (a) Shown is 98-residue protein α-spectrin (PDB ID: 1AJ3) which has three helices H1 (2–23), H2 (29–67), and H3 (72–96). For the cutoff distance d_c=6.5 Å, there are one and two native contacts, denoted by black lines, between H1 and H3, and H2 and H3, respectively. The total number of native contacts Q=72. (b) Schematic plot of our model for α-proteins. Meaning of D_0 and d_{spring} is given in the text. $\vec{n}$ is a vector perpendicular to the plane of coils which is schematically presented by a dashed line. θ is the angle between $\vec{F}$ and $\vec{n}$.

$=a/2b=0.132$ and $\alpha \approx 15°$. This angle corresponds to $f_u^{\text{eq}} \approx 300$ pN [Fig. 3(a)], which has the same order of magnitude as the estimation obtained from the free energy landscape picture with $\epsilon_{\text{H}}=4-5$ kcal/mol (see Sec. I). Nevertheless, our model may overestimate f_u^{eq} and this might be partially due to uncertainties in theoretical values of E.[37] Using experimental data and a different theoretical model, Keten and Buehler[39] have shown that β-proteins cannot exhibit rupture forces larger than ≈200 pN in the vanishing pulling speed limit. Their result is also consistent with ours.

III. ELASTIC THEORY FOR HELIX PROTEINS

Figure 4(a) shows the native conformation of α-spectrin, which is a typical helix protein. It contains three helices H1, H2, and H3. Assuming that a contact between two residues is formed if the distance between two corresponding C_α atoms is less than or equal to the cutoff d_c=6.5 Å, we have the number of native contacts between H1 and H3 $Q_{13}=1$, and between H2 and H3 $Q_{23}=2$. We define an averaged relative contact number for a pair interacting secondary structures $RQ_{\text{av}}=\Sigma_{ij}Q_{ij}\Delta(i-j)/(Q\Sigma_{ij}\Delta(i-j))$, where the summation is taken over all pairs of interacting secondary structures, Q is the total number of native contacts, and $\Delta(i-j)$ is 1 if helices i and j interact and 0 otherwise. For α-spectrin with Q=72, we obtain $RQ_{\text{av}} \approx 2\%$. In addition to α-spectrin, the values of RQ_{av} were estimated for other ten proteins (1CFC, 1F63, 1HRC, 1IMQ, 1LMB, 1VSC, 1YCC, 256B, 2ABD, and

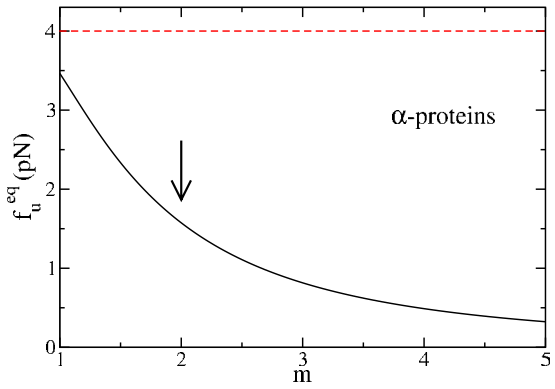

FIG. 5. Dependence of f_u^{eq} on ratio m for helical proteins. We used G =33 MPa and d_{spring}=5.4 Å. Arrow refers to the most probable value of m=2 which corresponds to $f_u^{\text{eq}} \approx 2$ pN. Horizontal red line refers to the typical value of f_u^{eq}=4 pN estimated, using the distance between the native and transition state and ϵ_{H}=1 kcal/mol.

2PPD), the mechanical properties of which were studied in detail previously (see Ref. 4 and references therein). The value of RQ_{av} averaged over the whole set of studied proteins is equal to $RQ_{\text{av}}=3.87\%$. Therefore, the interaction between helices, which is proportional to the number of native contacts between them, is weak. This conclusion remains valid for values of the cutoff d_c varying in the interval 5.5 Å $\leq d_c \leq$ 8 Å.

Neglecting the interaction between secondary structures, we propose to apply a mechanical spring model[40,41] with a single helix [see Fig. 4(b)] to describe mechanical properties of α-proteins. Thus, the unfolding in our model proceeds in such a way that secondary structures unravel in sequence. Let the initial spring diameter be D_0 and the current diameter under the load $D \equiv D(\theta)$, whereas the wire diameter is d_{spring}. Here θ is the current angle between the pulling direction and the perpendicular to the plane of coils [Fig. 4(b)]. The elongation under the action of a force F equals

$$\Delta L = \frac{8PD^3n}{Gd_{\text{spring}}^4} + \frac{4FDn}{Gd_{\text{spring}}^2}, \tag{3}$$

where G and n are the shear modulus and a number of coils, respectively. The wire length is $L=\pi Dn/\cos\theta$. During the protein stretching the helix straightens out with D varying from D_0 to d_{spring} and θ from a small value close to 0 to $\pi/2$. Taking into account the variability of D and θ, the maximal elongation takes the form

$$\Delta L_{\max} = \frac{2}{\pi}\left(\frac{16f_u^{\text{eq}}D_0^3n}{3Gd_{\text{spring}}^4} + \frac{4f_u^{\text{eq}}D_0n}{Gd_{\text{spring}}^2}\right). \tag{4}$$

At the same time, this quantity equals to $n(\pi D_0 - d_{\text{spring}})$, so that

$$f_u^{\text{eq}} = \frac{3\pi Gd_{\text{spring}}^2}{8m}\left(\frac{\pi m-1}{4m^2+3}\right), \tag{5}$$

where we have introduced the notation $m=D_0/d_{\text{spring}}$.

As far as we know, the shear modulus G for α-proteins shown in Fig. 1 is not known. To estimate f_u^{eq}, we take G^α =33 MPa (Ref. 42) measured for the collagen superhelix. The dependence of f_u^{eq} on m is shown in Fig. 5, where

d_{spring}=5.4 Å, was used. For superhelix, diameter D_0 varies between 5 (Ref. 41) and 10 Å.[43] Therefore, for α-proteins m takes values in the interval $1<m<2$ and f_u^{eq} is about 2–3 pN (Fig. 5). Using rupture forces measured at two pulling speeds[29] and the formula $f_u \approx f_u^{eq} \ln(v/v_0)$,[44] we obtain $f_u^{eq} = 2 \pm 2$ pN for α-spectrin. Thus, our result is in reasonable agreement with experiments and estimates from the free energy landscape picture which show that the critical force at vanishing pulling speed is of a few piconewtons for α-proteins.

As evident from Figs. 3(a) and 5, β-proteins are much more stable than α-proteins. This is our main qualitative result obtained in the framework of the proposed mechanical model of proteins. Note that there is some uncertainty in values of parameters used in our theory. Nevertheless, this cannot change our final conclusions, since we have a large safety margin of at least one order of magnitude.

IV. CONCLUSIONS

In conclusion, we have developed a simple elastic theory to understand the mechanical stability of proteins. Our theory contains two free parameters (angle α and E in the case of β-proteins, and m and G in the case of α-proteins) which may be estimated from either experiments or the Protein Data Bank. The success of our model in estimating f_u^{eq} for helix proteins is probably due to the weak interaction between their secondary structures. More accurate estimation of the mechanical stability threshold for individual β-proteins by our theory will be possible once experimental data for the modulus E are available. Although α/β-proteins have not been considered, one can infer from our theory that, in agreement with experiments and simulations, their mechanical stability is intermediate.

It is well know that mechanical stability of proteins depends on the pulling geometry,[45–48] i.e., on either the force direction or the linkage through which the force is applied to. For this reason, the protein C2A, e.g., is mechanically weak[19] although it consists of β-strands only. Our basic theory is developed for the case when the external force is applied to one end of a protein keeping another end fixed and the force direction is along the vector connecting two termini. Our theory may be applied to the case of different pulling geometries but then one has to determine the effective elastic moduli for corresponding geometries. This problem is left for further studies.

As follows from Fig. 3(a), rupture forces grow with the decrease in angle α. One of possibilities to obtain small values of this angle is to increase lengths of β-strands. This prediction may have interesting implications for designing mechanically strong biological materials.

ACKNOWLEDGMENTS

The work was supported by the Ministry of Science and Informatics in Poland (Grant No. 202-204-234) and Project N 23 of the 2009-2011 Scientific Cooperation Agreement between Poland and Ukraine. A.M.G. is grateful to Kasa im. Józefa Mianowskiego, Polski Koncern Naftowy ORLEN, and Fundacja Zygmunta Zaleskiego as well as to the 2008 and 2009 Visitors Programs of the Max Planck Institute for the Physics of Complex Systems (Dresden, Germany) for the support of his study.

[1] M. Rief, M. Gautel, F. Oesterhelt, J. M. Fernandez, and H. E. Gaub, Science **276**, 1109 (1997).
[2] L. Tskhovrebova, K. Trinick, J. A. Sleep, and M. Simons, Nature (London) **387**, 308 (1997).
[3] C. Bustamante, Y. R. Chemla, N. R. Forde, and D. Izhaky, Annu. Rev. Biochem. **73**, 705 (2004).
[4] M. S. Li, Biophys. J. **93**, 2644 (2007).
[5] M. S. Li, A. M. Gabovich, and A. I. Voitenko, J. Chem. Phys. **129**, 105102 (2008).
[6] H. Dietz and M. Rief, Phys. Rev. Lett. **100**, 098101 (2008).
[7] M. J. Buehler and Y.-C. Yung, Nature Mater. **8**, 175 (2009).
[8] J. I. Sułkowska and M. Cieplak, J. Phys.: Condens. Matter **19**, 283201 (2007).
[9] K. W. Plaxco, K. T. Simon, and D. Baker, J. Mol. Biol. **277**, 985 (1998).
[10] R. Best, S. Fowler, J. Herrera, A. Steward, E. Paci, and J. Clarke, J. Mol. Biol. **330**, 867 (2003).
[11] H. Li and J. M. Fernandez, J. Mol. Biol. **334**, 75 (2003).
[12] H. Li, A. F. Obserhauser, S. B. Fowler, J. Clarke, and J. M. Fernandez, Proc. Natl. Acad. Sci. U.S.A. **97**, 6527 (2000).
[13] M. Rief, M. Gautel, M. Schemmel, and H. G. Gaub, Biophys. J. **75**, 3008 (1998).
[14] Y. Oberdorfer, H. Fuchs, and A. Janshoff, Langmuir **16**, 9955 (2000).
[15] A. Oberhauser, C. Badilla-Fernandez, M. Carrion-Vazquez, and J. Fernandez, J. Mol. Biol. **319**, 433 (2002).
[16] M. Schlierf and M. Rief, J. Mol. Biol. **354**, 497 (2005).
[17] I. Schwaiger, M. Schlierf, A. Noegel, and M. Rief, EMBO Rep. **6**, 46 (2005).
[18] M. Schlierf and M. Rief, Biophys. J. **90**, L33 (2006).
[19] M. Carrion-Vazquez, A. Oberhauser, T. Fisher, P. Marszalek, H. Li, and J. Fernandez, Prog. Biophys. Mol. Biol. **74**, 63 (2000).
[20] K. Watanabe, C. Muhle, M. S. Z. Kellermayer, S. Labeit, and H. L. Granzier, J. Struct. Biol. **137**, 248 (2002).
[21] K. Watanabe, P. Nair, D. Labeit, M. S. Z. Kellermayer, M. Greaser, and H. L. Granzier, J. Biol. Chem. **277**, 11549 (2002).
[22] D. J. Brockwell, G. Beddard, E. Paci, D. West, P. Olmsted, D. Smith, and S. Radford, Biophys. J. **89**, 506 (2005).
[23] C. Chyan, F. Lin, H. Peng, J. Yuan, C. Chang, S. Lin, and G. Yang, Biophys. J. **87**, 3995 (2004).
[24] R. Best, B. Li, A. Steward, V. Daggett, and J. Clarke, Biophys. J. **81**, 2344 (2001).
[25] S. Ainavarapu, L. Li, and J. Fernandez, Biophys. J. **89**, 3337 (2005).
[26] H. Dietz and M. Rief, Proc. Natl. Acad. Sci. U.S.A. **101**, 16192 (2004).
[27] G. Yang, C. Cecconi, W. A. Baase, I. R. Vetter, W. A. Breyer, J. A. Haack, B. W. Matthews, F. W. Dahlquist, and C. Bustamante, Proc. Natl. Acad. Sci. U.S.A. **97**, 139 (2000).
[28] C. Cecconi, E. A. Shank, C. Bustamante, and S. Marquese, Science **309**, 2057 (2005).
[29] M. Rief, J. Pascual, M. Saraste, and H. Gaub, J. Mol. Biol. **286**, 553 (1999).
[30] R. Law, P. Carl, S. Harper, D. W. Speicher, and D. E. Discher, J. Biol. Chem. **279**, 16410 (2004).
[31] N. Bhasin, P. Carl, S. Harper, G. Feng, H. Lu, D. W. Speicher, and D. E. Discher, J. Biol. Chem. **279**, 45865 (2004).
[32] G. Lee, K. Abdi, Y. Jiang, P. Michaely, V. Bennett, and P. E. Marszalek, Nature (London) **440**, 246 (2006).
[33] L. Li, S. Wetzel, A. Pluckthun, and J. Fernandez, Biophys. J. **90**, L30 (2006).
[34] L. Dougan, G. Feng, H. Lu, and J. M. Fernandez, Proc. Natl. Acad. Sci. U.S.A. **105**, 3185 (2008).
[35] H. Lu and K. Schulten, Biophys. J. **79**, 51 (2000).
[36] A. Ikai, Biophys. Chem. **116**, 187 (2005).
[37] G. C. Yoon, H. J. Park, K. Eom, and S. Na, J. Comput. Chem. **30**, 873 (2009).
[38] J. R. Banavar, A. Maritan, C. Micheletti, and A. Trovato, Proteins: Struct., Funct., Genet. **47**, 315 (2002).
[39] S. Keten and M. J. Buehler, Phys. Rev. Lett. **100**, 198301 (2008).
[40] S. P. Timoshenko, *Strength of Materials: Part I: Elementary Theory and Problems* (Van Nostrand, Princeton, 1955).
[41] K. Beck and B. Brodsky, J. Struct. Biol. **122**, 17 (1998).

[42] L. Yang, K. O. van der Werf, C. F. C. Fitie, M. L. Bennink, P. J. Dijkstra, and J. Feijen, Biophys. J. **94**, 2204 (2008).
[43] B. Brodsky and J. A. M. Ramshaw, Matrix Biol. **15**, 545 (1997).
[44] E. Evans and K. Ritchie, Biophys. J. **72**, 1541 (1997).
[45] M. Carrion-Vazquez, H. Li, H. Lu, P. E. Marszalek, A. F. Oberhauser, and J. M. Fernandez, Nat. Struct. Biol. **10**, 738 (2003).
[46] H. Dietz, F. Berkemeier, M. Bertz, and M. Rief, Proc. Natl. Acad. Sci. U.S.A. **103**, 12724 (2006).
[47] D. J. Brockwell, E. Paci, R. C. Zinober, G. S. Beddard, P. D. Olmsted, D. A. Smith, R. N. Perham, and S. E. Radford, Nat. Struct. Biol. **10**, 731 (2003).
[48] M. Kouza, C. K. Hu, and M. S. Li, J. Chem. Phys. **128**, 045103 (2008).