Relationship between the Mechanical Properties and Topology of Cross-Linked Polymer Molecules: Parallel Strands Maximize the Strength of Model Polymers and Protein Domains

Kilho Eom,[†] Pai-Chi Li,[‡] Dmitrii E. Makarov,*[,‡,§,||] and Gregory J. Rodin[†,§]

Department of Aerospace Engineering and Engineering Mechanics, Department of Chemistry and Biochemistry, Institute for Computational Engineering and Sciences, and Institute for Theoretical Chemistry, The University of Texas, Austin, Texas 78712

Received: May 1, 2003; In Final Form: July 5, 2003

Proteins that perform mechanical functions in living organisms often exhibit exceptionally high strength and elasticity. Recent studies of the unfolding of single protein molecules under mechanical loading showed that their strength is mostly determined by their native topology rather than by thermodynamic stability. To identify the topologies of polymer molecules that maximize their resistance to unfolding, we have simulated the response of cross-linked polymer chains under tensile loading and have found that chain configurations that maximize the unfolding work and force involve parallel strands. Chains with such optimal topologies tend to unfold in an all-or-none fashion, in contrast to randomly cross-linked chains, most of which exhibit low mechanical resistance and tend to unfold sequentially. These findings are consistent with AFM studies and molecular mechanics simulations of the unfolding of β-sheet proteins. In particular, parallel strands give rise to the high strength of the immunoglobulin-like domains in the muscle protein titin.

I. Introduction

Proteins intended for load-bearing functions in living organisms (e.g., titin, tenascin, and spider silk proteins) are very resistant to unfolding under mechanical loading.[3,4,20,22,25,27,29] This high resistance to unfolding is believed to give rise to the unique strength and toughness of natural fibers.[27] In contrast, the resistance to unfolding is low in proteins that are not required to perform mechanical functions.[1,28,31] Several recent studies suggested that whereas the mechanical strength of proteins is often uncorrelated with their thermodynamic stability[2,19] their mechanical unfolding mechanism is largely determined by their native topology.[9,26] Here we use a solvable model of cross-linked Gaussian polymer chains to investigate the relationship between topology and the mechanical response of polymer molecules and identify the optimal topologies that maximize their mechanical load-bearing capacity.

II. Topological Optimization Problem

Consider a polymer consisting of $L + 1$ beads (numbered 0, 1, 2,..., L) connected into a chain by L links. Assume that the chain obeys Gaussian statistics such that the probability distribution for the distance between any two monomers i and j is Gaussian:

$$P(\mathbf{r}_i - \mathbf{r}_j) = \left(\frac{\gamma}{2\pi k_B T|i-j|}\right)^{3/2} \exp\left(-\frac{\gamma|\mathbf{r}_i - \mathbf{r}_j|^2}{2k_B T|i-j|}\right)$$

where $\mathbf{r}_i$ is the position of the ith bead, k_B is the Boltzmann constant, T is the temperature, and γ has units of a force constant and is related to the mean-square length of a link by $\gamma = 3k_BT/\langle|\mathbf{r}_{i+1} - \mathbf{r}_i|^2\rangle$.

Suppose now that there are cross links connecting some of the beads. A cross link between a pair of beads is formed by connecting them with a harmonic spring with a very large force constant and zero equilibrium length. A cross link between beads i and j is labeled $\{i, j\}$. The chain topology is then defined by the cross-link list: $C = \{\{i_1, j_1\}, \{i_2, j_2\},..., \{i_N, j_N\}\}$ (Figure 1).

Now suppose that we stretch the chain by moving its ends apart slowly enough that the time scale of stretching is much longer than that of the polymer's thermal motions. Further assume that a cross link ruptures once the force it transmits attains a critical value equal to f_c. When the distance between the ends of the chain e is short enough that the internal forces in the cross links are all below f_c, the force $F(e)$ generated by the chain in our model is a linear function of e.[17] As the extension e is increased and the cross links start rupturing, $F(e)$ becomes a piecewise linear function with jumps (Figure 1). Each jump corresponds to a rupture event, which may involve one or more cross links and results in an abrupt drop in the chain's stiffness and, consequently, a decrease in force F. Once all of the cross links are broken, $F(e)$ attains the slope γ/L of the "denatured" chain without cross links. In the following text, we quantify the chain resistance using two quantities: (1) The excess work $W = \int_0^e F(s)\,ds - (\gamma e^2/2L)$ performed when unravelling the cross-linked chain as compared to the work expended on extending the denatured chain. This excess work is equal to shaded area shown in Figure 1b. (2) The peak force $F_p = \max[F(e)]$ recorded in the course of polymer stretching. The choice of W and F_p as the measures of resistance is consistent with the notions of toughness and strength, respectively, which are well accepted in the characterization of solids.

Our optimization problem is as follows: For given L, N, γ, and f_c, we seek the optimal sets C maximizing (a) the excess

* Corresponding author. E-mail: makarov@mail.cm.utexas.edu.
[†] Department of Aerospace Engineering and Engineering Mechanics.
[‡] Department of Chemistry and Biochemistry.
[§] Institute for Computational Engineering Sciences.
[||] Institute for Theoretical Chemistry.

10.1021/jp035178x CCC: $25.00
Published on Web 08/05/2003

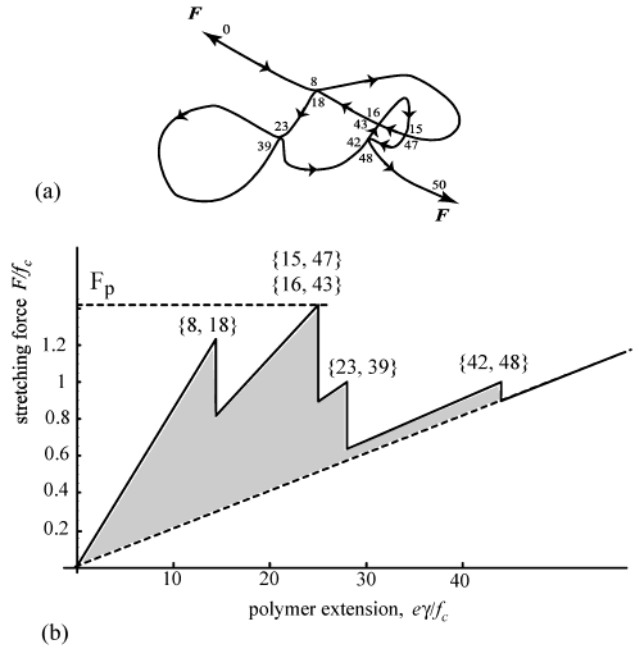

Figure 1. Unfolding of a cross-linked polymer. (a) The configuration of a polymer with $L = 50$ and with cross links {{8, 18}, {15, 47}, {16, 43}, {42, 48}, {23, 39}}. (b) The force-extension curve of this polymer. The force is measured in units of f_c, and the extension, in units of f_c/γ, where γ is the effective force constant of a single chain link. Each peak corresponds to the rupture of one or more cross links, as indicated on top of each peak. The shaded area is equal to the excess work W required to extend the cross-linked chain as compared to that for the "denatured" chain.

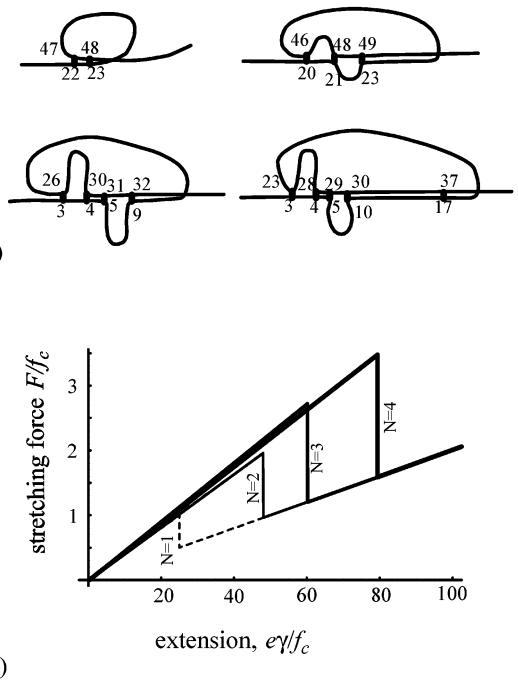

Figure 2. Optimal polymer configurations. (a) The cross-link configurations for polymers with $L = 50$ links and $N = 2-5$ cross links, maximizing the excess work W. These configurations also maximize the peak unfolding force F_p. (b) The force-extension curves for the optimal polymer configurations with $N = 1-4$.

work W and (b) the peak force F_p, provided that any two monomers can be connected by no more than one cross link.

III. Optimal Configurations

We found analytically that for $N = 1$ and $L = 2l$ the optimum is realized by any cross link of the type $\{i, i + l\}$. For $N = 2$, through an exhaustive enumeration of all cross-link configurations we found that the optimal sets are of the form $\{\{i, i + l\}, \{i + 1, i + 1 + l\}\}$. To find the optimal chain configurations for $N \geq 3$ and $L = 50$, we had to resort to random search strategies described in the Methods section. Some of the optimal configurations found for $N = 3-5$ are shown in Figure 2. All of these configurations share two features: (i) they optimize *both* W and F_p and (ii) they have the topology of *parallel strands* (i.e., they can be represented in the form $\{\{i_1, j_1\}, \{i_2, j_2\},...\{i_N, j_N\}\}$ where $i_1 < i_2 < ... i_N < j_1 < j_2 < ... < j_N$). For the best configurations found for $N = 2-4$, the cross links rupture at once as soon as a critical force is reached, leading to a single peak in the force-extension curves shown in Figure 2b.

To gain more insight into the structure of the optimal configurations that were found, consider a continuous optimization problem where chain length L and indices $0 < i \leq L$ enumerating beads are regarded as continuous quantities. Suppose we create a "super cross link" (SCL) by placing all N cross links between two points so that the cross links are connected in parallel and each cross link bears the force F/N. Then the optimum with respect to W is achieved by choosing the two points as $\{x, x + L/2\}$ with $0 \leq x \leq L/2$, as dictated by the optimal solution for $N = 1$. As a result, we obtain $W = (LN^2 f_c^2/8\gamma)$, $F_p = Nf_c$.

The optimal C's shown in Figure 2 are close to the above SCL configuration but satisfy the additional constraint that no more than one cross link can be placed between two points. As N increases, this constraint becomes increasingly more significant, and as our numerical experiments indicate, for moderate N/L one should not expect quadratic scaling in N for the optimal W and linear scaling in N for the optimal F_p. Similar to the SCL configuration in the continuum case, parallel strand configurations tend to spread the load (nearly) equally among the N cross links such that the rupture of one cross link results in an increase of the load on the remaining cross links, leading to an avalanche-like separation of the strands.

The fact that this avalanche rupture scenario optimizes not only the unfolding force F_p but also the work W is rather surprising: One could expect that having to "restretch" the polymer performing extra work against the polymer's entropic elasticity each time a cross link is ruptured and a "slack" is created in the chain would provide a useful strategy to maximize the work W. Our results indicate that this strategy is inferior to the one that maximizes the rupture force by ensuring that all cross links share the load and break simultaneously. The optimal conformations are "unique" in that an overwhelming majority of random cross-link configurations were found to unfold in a sequential fashion (exemplified in Figure 1) and to exhibit a much lower unfolding force and unfolding work. In particular, for $N = 4$, the mean excess work $\langle W \rangle$ found by averaging W calculated for random cross-link configurations is about 5 times smaller than the excess work W required to unfold the optimum conformation shown in Figure 2a.

IV. Discussion

Our simple model cannot be expected to be an adequate description of protein unfolding. It entirely ignores any details of the energetics of protein folding, which are known to affect the mechanical resistance of globular proteins[10] significantly. It further ignores the stiffness of the protein backbone. We also note that our analysis does not apply to the case of isotropic deformations of biological or artificial polymer networks. With

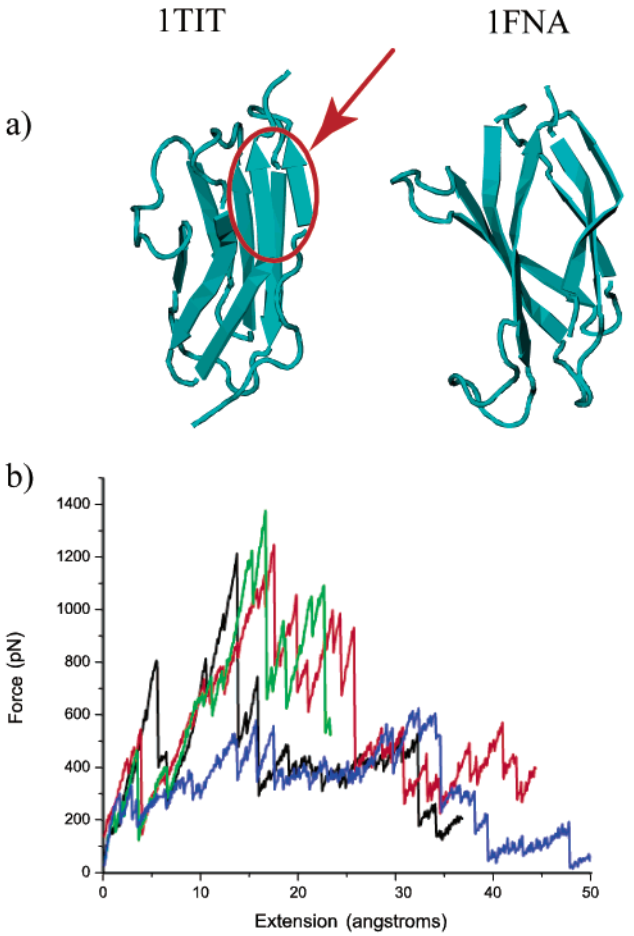

Figure 3. Molecular mechanics stretching of β-sheet protein domains. (a) The structure of the I27 (pdb code 1TIT) and FNIII (1FNA) domains. This plot was created by using VMD software.[7] The "clamp" formed by parallel strands in I27 is indicated by an arrow. (b) Force-extension curves of Ig27 (black, green, and red lines) and of FNIII (blue line) were calculated as described in the Methods section.

these caveats in mind, one may wonder whether forming parallel strands could provide a possible mechanism by which the mechanical resistance of proteins could be optimized. Indeed, the known mechanical properties of single protein molecules appear to be consistent with the above findings. The immunoglobulin-like I27 domain in the muscle protein titin is one of the best studied examples of a protein domain showing remarkable mechanical strength.[3,8,18,22,24,25] Recent simulations[11,12,14] suggest that the unfolding of I27 involves the simultaneous rupture of hydrogen bonds forming parallel strands (Figure 3). The present study supports the conjecture[2] that those strands are responsible for the high strength of I27.

Additional support for the optimality of parallel strands comes from our molecular mechanics studies of the unfolding of I27 and several other β-sheet proteins. In those studies, the distance between the ends of a molecule was increased in small increments, and constrained energy minimization was performed for each distance.[26] Such simulations can be thought of as "zero-temperature" stretching experiments. Because in the experimentally accessible regime thermal fluctuations allow barrier-crossing events to speed up the domain unfolding, the forces that are measured experimentally are much lower than those observed in our simulations.

Shown in Figure 3 are several simulated force-extension curves for the I27 domain from titin (pdb code 1TIT) and for the 10th type-III cell-adhesion module of human fibronectin (pdb code 1FNA). The two domains have comparable lengths and exhibit very similar β-sandwich folds. The three different force-extension curves for I27 shown in Figure 3 demonstrate the sensitivity of $F(e)$ to the initial molecular configuration, which was selected in each case by quenching, via steepest descent minimization, of three random configurations of the I27 molecule encountered in the course of an equilibrium, room-temperature molecular dynamics simulation. All three curves exhibit similar values of the peak force. The largest peak (at about 1200−1400 pN) seen in all of the I27 pulling curves corresponds to the separation of the parallel strands. In contrast, the fibronectin domain exhibits much lower strength and unfolds in a sequential manner. The key structural difference between the two domains is the presence of a "clamp" formed by the parallel strands in the I27 domain,[11,12] as shown in Figure 3. These findings are consistent with the previous simulations of the force-induced unfolding of fibronectin and immunoglobulin domains at room temperature.[12,21] We obtained similar results for other β-sheet protein domains that did not involve parallel strands, and none of them was as strong as the I27 domain.

We finally note that the present study does not take into account kinetic effects and assumes that the mechanical unfolding of polymers is a deterministic process. Depending on the temperature and the rate of loading, the rupture of chemical bonds in polymers may be driven by thermal fluctuations, leading to a distribution of rupture forces and to the inherently stochastic nature of the measured force-extension curves. Whether the optimality of the obtained polymer configurations will be changed by these effects remains to be seen in future studies.

V. Methods

Simulating the Unfolding of Cross-Linked Polymer Chains. The method by which we calculate the entropic elasticity of cross-linked Gaussian chains is described in ref 17. This approach utilizes the mathematical equivalence of the mechanical response of a network of Gaussian polymers and that of a system of mechanical springs with suitably chosen force constants.[15−17] Specifically, a segment of the chain between monomers i and j such that it contains no cross links is replaced by a mechanical spring with a force constant of $k_{ij} = \gamma/|i - j| = 3k_BT/(|i - j|s^2)$, where the effective link length s is related to the mean-square end-to-end distance for a free chain consisting of $|i - j|$ links, $\langle r^2\rangle = s^2|j-i|$. Each cross link is treated as a spring with a large force constant of $k_{cl} \gg 3k_BT/s^2$. The effective force constant of the equivalent system of linear springs and the force in each cross link were calculated by the finite element method described in ref 17. To obtain the force-extension curve for a given initial set of cross links C, we (i) calculated the overall force constant $k = dF/de$ of the chain for the current set of cross links, (ii) calculated the critical extension e and the corresponding force F, for which the force in the most loaded cross link achieves the critical value f_c, (iii) removed the most loaded cross link, and (iv) repeated steps i−iii with the new cross-link configuration. These steps were repeated until all of the cross links were eliminated.

Topological Optimization. To find optimum cross-link configurations maximizing either W or F_p, we used either exhaustive searches or "random hill-climbing" strategies that tested local maxima. In the latter case, one creates a random cross-link configuration C_0 with the prescribed number of cross-links N and subsequently attempts to improve upon it via a series of local moves. To this end, a random cross link $\{i, j\}$ is selected from C_0, and the optimal configuration among $\{i, j\}$ and $\{i \pm$

1, $j \pm 1$} is determined. Then the search proceeds to another cross link. Once C_0 can no longer be improved, one generates a new initial configuration C_1 and so forth. A typical number of configurations tested in the course of a search was on the order of 10^7.

Molecular Mechanics Studies. Simulations of the stretching of protein domains were performed with Tinker molecular dynamics software (http://dasher.wustl.edu/tinker/) using the GB/SA continuum model for solvation[23] and the Charm27 force field.[13] A harmonic restraint was imposed on the distance R between the outermost α-carbon atoms of the protein by introducing the penalty term $(^1/_2)\, k(R-R_0)^2$ in the energy. This restraint plays a role similar to that of a cantilever in the AFM protein-pulling experiments.[5,6,24,25,30] The value R_0 was repeatedly incremented by a small amount $\Delta R = 0.1$ Å, $R_0(i) = R_0(0) + i\Delta R$, where $i = 1, 2,...$ and $R_0(0)$ is the distance between the outermost α-carbon atoms in the minimum-energy structure obtained from the pdb structure via unconstrained energy minimization. For each value $R_0(i)$, constrained minimization was performed starting from the minimum-energy structure found in the $(i - 1)$-th step.[26] The actual value of the distance R that was found for each minimum-energy structure was different from the value R_0 that was set by the constraint. Then the force acting on the protein molecule was equal to $k(R_0 - R)$. This force was plotted in Figure 3 as a function of protein extension $e = R - R_0(0)$.

Acknowledgment. This work was supported in part by the Robert A. Welch Foundation. G.J.R. acknowledges the hospitality of the Laboratoire de Mecanique des Solides at Ecole Polytechnique and the French CNRS for an appointment as Associate Researcher. We thank Helen G. Hansma, Emin Oroudjev, and Kevin W. Plaxco for helpful discussions.

References and Notes

(1) Best, R. B.; Li, B.; Steward, A.; Daggett, V.; Clarke, J. *Biophys. J.* **2001**, *81*, 2344.

(2) Brockwell, D. J.; Beddard, G. S.; Clarkson, J.; Zinober, R. C.; Blake, A.; Trinick, J.; Olmsted, P. D.; Smith, D. A.; Radford, S. E. *Biophys. J.* **2002**, *83*, 458.

(3) Erickson, H. P. *Science* **1997**, *276*, 1090.

(4) Fisher, T. E.; Oberhauser, A. F.; Vezquez, M. C.; Marsalek, P. E.; Fernandez, J. *TIBS* **1999**, *24*, 379.

(5) Hansma, H. G.; Pietrasanta, L. *Curr. Opin. Chem. Biol.* **1998**, *2*, 579.

(6) Hansma, H. G.; Pietrasanta, L. I.; Auerbach, I. D.; Sorenson, C.; Golan, R.; Holden, P. A. *J. Biomater. Sci., Polym. Ed.* **2000**, *11*, 675.

(7) Humphrey, W.; Dalke, A.; Schulten, K. *J. Mol. Graphics* **1996**, *17*, 33.

(8) Kellermayer, M. S. Z.; Smith, S. B.; Granzier, H. L.; Bustamante, C. *Science* **1997**, *276*, 1112.

(9) Klimov, D. K.; Thirumalai, D. *Proc. Natl. Acad. Sci. U.S.A.* **2000**, *97*, 7254.

(10) Li, H.; Carrion-Vazquez, M.; Oberhauser, A. F.; Marsalek, P. E.; Fernandez, J. M. *Nat. Struct. Biol.* **2000**, *7*, 1117.

(11) Lu, H.; Isralewitz, B.; Krammer, A.; Vogel, V.; Schulten, K. *Biophys. J.* **1998**, *75*, 662.

(12) Lu, H.; Schulten, K. *Chem. Phys.* **1999**, *247*, 141.

(13) MacKerell, A. D., Jr.; Bashford, D.; Bellott, R. L.; Dunbrack, R. L., Jr.; Evanseck, J. D.; Field, M. J.; Fischer, S.; Gao, J.; Guo, H.; Ha, S.; Joseph-McCarthy, D.; Kuchnir, L.; Kuczera, K.; Lau, F. T. K.; Mattos, C.; Michnick, S.; Ngo, T.; Nguyen, D. T.; Prodhom, B.; Reiher, W. E., III; Roux, B.; Schlenkrich, M.; Smith, J. C.; Stote, R.; Straub, J.; Watanabe, M.; Wiórkiewicz-Kuczera, J.; Yin, D.; Karplus, M. *J. Phys. Chem. B* **1998**, *102*, 3586.

(14) Makarov, D. E.; Hansma, P. K.; Metiu, H. *J. Chem. Phys.* **2001**, *114*, 9663.

(15) Makarov, D. E.; Keller, C.; Plaxco, K. W.; Metiu, H. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 3535.

(16) Makarov, D. E.; Metiu, H. *J. Chem. Phys.* **2002**, *116*, 5205.

(17) Makarov, D. E.; Rodin, G. J. *Phys. Rev. E* **2002**, *66*, 011908.

(18) Marsalek, P. E.; Lu, H.; Li, H.; Carrion-Vazquez, M.; Oberhauser, A. F.; Schulten, K.; Fernandez, J. *Nature* **1999**, *402*, 100.

(19) Oberhauser, A. F.; Badilla-Fernandez, C.; Carrion-Vazquez, M.; Fernandez, J. M. *J. Mol. Biol.* **2002**, *319*, 433.

(20) Oberhauser, A. F.; Marszalek, P. E.; Erickson, H.; Fernandez, J. M. *Nature* **1998**, *393*, 181.

(21) Paci, E.; Karplus, M. *J. Mol. Biol.* **1999**, *288*, 441−459.

(22) Pennisi, M. E. *Science* **1999**, *283*, 168.

(23) Qiu, D.; Shenkin, P. S.; Hollinger, F. P.; Still, W. C. *J. Phys. Chem. A* **1997**, *101*, 3005.

(24) Rief, M.; Fernandez, J. M.; Gaub, H. E. *Phys. Rev. Lett.* **1998**, *81*, 4764.

(25) Rief, M.; Gautel, M.; Oesterhelt, F.; Fernandez, J. M.; Gaub, H. E. *Science* **1997**, *276*, 1109−1112.

(26) Rohs, R.; Etchebest, C.; Lavery, R. *Biophys. J.* **1999**, *76*, 2760−2768.

(27) Smith, B. L.; Schaffer, T. E.; Viani, M.; Thompson, J. B.; Frederick, N. A.; Kindt, J.; Belcher, A.; Stucky, G. D.; Morse, D. E.; Hansma, P. K. *Nature* **1999**, *399*, 761.

(28) Thompson, J. B.; Hansma, H. G.; Hansma, P. K.; Plaxco, K. W. *J. Mol. Biol.* **2002**, *322*, 645−652.

(29) Tskhovrebova, L.; Trinic, J. A.; Sleep, J. A.; Simmons, R. M. *Nature* **1997**, *387*, 308.

(30) Viani, M. B.; Schaffer, T. E.; Paloczi, G. T.; Pietrasanta, I.; Smith, B. L.; Thompson, J. B.; Richter, M.; Rief, M.; Gaub, H. E.; Plaxco, K. W.; Cleland, A. N.; Hansma, H. G.; Hansma, P. K. *Rev. Sci. Instrum.* **1999**, *70*, 4300.

(31) Yang, G.; Cecconi, C.; Baase, W. A.; Vetter, I. R.; Breyer, W. A.; Haack, J. A.; Matthews, B. W.; Dahlquist, F. W.; Bustamante, C. *Proc. Natl. Acad. Sci. U.S.A.* **2000**, *97*, 139.