Mechanical unfolding and refolding of proteins: An off-lattice model study

Feng-Yin Li,[1] Jian-Min Yuan,[1,2] and Chung-Yuan Mou[1]
[1]*Department of Chemistry, National Taiwan University, Taipei, Taiwan 106, Republic of China*
[2]*Department of Physics, Drexel University, Philadelphia, Pennsylvania 19104*
(Received 27 December 1999; revised manuscript received 28 July 2000; published 19 January 2001)

Using an off-lattice model, we investigate the response of a protein molecule to external stretching and release. In particular we study the passive force in a protein as a function of the extension of the protein. These force-extension curves exhibit hysteresis loops, whose areas increase with the pulling rate and decrease with thermal noise. Most of these results seem to be appropriately described by a cusp catastrophe.

DOI: 10.1103/PhysRevE.63.021905 PACS number(s): 87.15.He, 87.15.Cc, 87.15.La, 47.50.+d

I. INTRODUCTION

Protein folding and unfolding is one of the most important problems in molecular biology. Because of the Genome Projects, an explosion of information on gene sequences has become available, and there is an urgency in determining the three-dimensional structures of proteins, coded by these genes. At the microscopic level the dynamics of protein folding and unfolding is governed by intramolecular and intermolecular noncovalent forces, except in the presence of disulfite bonds. Recent single-molecule experiments showed that it is possible to measure the mechanical response of single protein domains directly, when it is pulled by an external force. These experiments are carried out by using either atomic force microscopy (AFM) [1–5] or laser tweezers [6]. In the former type of experiments, a protein polymer tethered to a solid surface is pulled at another part of the polymer by the tip of an AFM lever [1–6]. In the latter type of experiments two ends of a protein are attached to latex beads, and one bead is then held in an optical trap, while the other bead is pulled by a micropipette [6]. In both cases, one can measure the passive force developed in a protein when the molecule is stretched, and the force-extension curves thus obtained often exhibit sawtooth patterns [1–5] and hysteretic behavior for a stretch-and-release cycle [1,2,6]. Furthermore, experiments using dynamic force spectroscopy show that the bond strength of a receptor-ligand complex [7] and the resistance of protein molecules to external forcing [6,5] depend on the loading rate.

Direct measurements of passive forces in single molecules find applications in materials and polymer sciences [8,9] as well. For instance, a recent experiment on nacre showed that the extraordinary strength of nacre may arise from the protein molecules sandwiched between the inorganic plates [8]. Single molecule measurements of related proteins show that a sawtooth pattern appears in their force-extension curves as well. This pattern reflects the unusual toughness of nacre, because a large dissipative energy is required to break materials made of protein domains, held together by noncovalent forces.

Single-molecule experiments are, in principle, most suitable for theoretical investigations and numerical simulations. In the present paper, using an off-lattice model, we study the response of a protein molecule when it is mechanically stretched and released. This model is built upon on the well-studied 46-bead Skolnick-Honeycutt-Thirumalai model [10–14] for β-barrel proteins. For reasons to become clear below, this model will be called the BPN model. Our objective is to shed light on the origins of the patterns of the force-extension curves, the hysteretic behavior, the dissipative energy involved, and the role of pulling rate in these curves. Numerical simulations of mechanical unfolding already appeared in the literature using lattice models [15,16] and all-atom molecular dynamics simulation [17,3,18,19]. However, a reduced off-lattice model allows us to investigate a dynamics which is more realistic than that of a lattice model. By focusing on the backbone dynamics, a minimal model like the BPN model has the advantage over all-atom simulations in the fact that we may be able to carry out longer-time integration and wider-range exploration in parameter space. Using such minimal models, Klimov and Thirumalai [20] were able recently to verify the hypothesis that stretching unfolding can be understood using only the topology of the native state (for more details, see Sec. V).

The organization of this paper is as follows: We briefly review some of the features of the BPN model in Sec. II, and present an extended model, modified for mechanical unfolding and refolding, in Sec. III. Results of molecular dynamics calculations of the extended model are presented in Sec. IV, in which we discuss the force-extension curves, hysteretic phenomenon, and temperature effects in mechanical unfolding and refolding. A concluding section follows.

II. *BPN* MODEL

The mechanical unfolding model that we shall use is a modification of the Skolnick-Honeycutt-Thirumalai β-barrel bead model [10,11], which was extensively studied for spontaneous protein folding [11–16,21–24]. We briefly review its essential features in this section. This model consists of a chain of 46 beads of three different colors, namely, hydrophobic (B), hydrophilic (P), and neutral (N). The specific bead sequence is $B_9N_3(PB)_4N_3B_9N_3(PB)_5P$, where B_9 denotes a segment made of nine beads of type B, $(PB)_4$ denotes four sets of PB pairs, and so forth. The strings B_9, $(PB)_4$, and $(PB)_5P$ are meant to mimic the strands of a β-sheet structure, which in this system is a barrel, because there are only four strands. Following Nymeyer *et al.* [14], hereafter we refer to this model as the BPN model.

The neutral beads act as very flexible separators to define

the four strands of the *BPN* model. The first strand includes the beads from the first bead to the ninth bead, the second strand ranges from the 13th bead to the 20th bead; the third strand from the 24th bead to the 32nd bead, and the last strand from the 36th bead to the 46th bead.

The potential functions and associated parameters are the those of Guo and Thiumalai [12], except that, instead of a rigid bond length, we assume an harmonic potential between any two connected beads, as used by Berry *et al.* [13]. The nonbonded interaction between any *B*-*B* pair (not next to each other in the sequence)is represented by a Lennard-Jones potential,

$$V_{BB}=4\varepsilon_h[(\sigma/r)^{12}-(\sigma/r)^6], \tag{2.1}$$

where r is the distance between the specific beads, σ the unit of length, and ε_h the strength of the hydrophobic interaction. The nonbonded potential between *P*-*P* and *P*-*B* pairs has a form given by

$$V_{P\alpha}=4\varepsilon_P[(\sigma/r)^{12}+(\sigma/r)^6] \quad (\alpha=P \text{ or } B), \tag{2.2}$$

where $\varepsilon_P=2/3\varepsilon_h$. The nonbonded interaction between an *N* bead with any other bead takes the form

$$V_{N\alpha}=4\varepsilon_h(\sigma/r)^{12} \quad (\alpha=P,\ B,\ \text{or } N). \tag{2.3}$$

The bond angle among three successive beads is constrained by a harmonic potential written as

$$V(\theta)=k_\theta(\theta-\theta_0)^2/2. \tag{2.4}$$

The dihedral angle potential, involving four successive beads, is represented by

$$V(\phi)=A[1+\cos(\phi)]+B[1+\cos(3\phi)], \tag{2.5}$$

where ϕ is the dihedral angle. For all the beads except those in the loop regions, $A=B=1.2\varepsilon_h$, while for beads in the loop region, when two or more beads among the four defining the dihedral angle are of *N* type, we choose smaller values of *A* and *B*, e.g., $A=0$, and $B=0.2\varepsilon_h$. Guo and Thirumalai [21] established, through dynamics calculations, that two characteristic temperatures exist in this β-barrel model: the collapse temperature T_θ, at which extended chains collapse into compact conformations, and a folding temperature T_f, at which compact conformations transform into one of an ensemble of native structures. Shea *et al.* [23] studied the dynamic effects of attractive, non-native hydrophobic interactions, by varying the strength of such interactions and found that decreasing the strength of such interactions tends to move T_θ closer to T_f and thus reduces the ratio of $\sigma=(T_\theta-T_f)/T_\theta$. In the limit of zero strength of such interactions, the 46-bead model becomes a very efficient folder. This result is in agreement with those of Nymeyer *et al.* [14], who studied a Go-like version of the 46-bead β-barrel model. They observed that in this Go-like model T_θ becomes coincident with T_f and the chain folds fast, that is, becomes minimally frustrated. This in turn agrees with the proposal of Veitshans *et al.* [25] that σ determines the folding rate, and a smaller σ value corresponds to a larger folding rate. Berry *et al.* [13], on other hand, applied methods of analysis for atomic clusters to the β-barrel model, and discovered two low-energy intermediates (the sliding intermediates) in addition to the native structure, which represents the global minimum. In addition, they delineated the structures of the saddles that link these ''basin-bottom structures,'' the intermediates and the native structure discussed above.

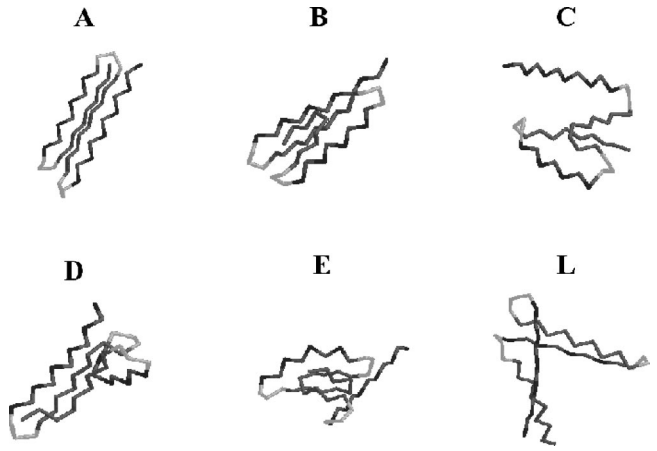

FIG. 1. The six typical conformation groups that appear during unfolding and refolding processes, including five groups of native and misfold conformations: *A*, *B*, *C*, *D*, and *E* (as described in Ref. [24]). Roughly speaking, the native state (*A*) consists of four strands well aligned: first and third strands are parallel to each other, but they are antiparallel to the second and fourth strands. Conformation group *B* differs from group *A* in that the first strand curls up. Group *C* differs from group *A* by curling up the third strand and group *E* differs from group *A* by curling up both the first and third strands. Group *D* differs from group *A* in that the first and third strands are antiparallel to each other, caused by a curling up of the second strand. Conformation group *L* is similar to the conformation group *A*, except that (2,3) and (1,4) pairs have not come together.

In addition, Guo and Thirumalai (GT) [21] found, through molecular dynamics simulations that roughly 60% (depending on the temperature used) of the 85 trajectories run reaches the native state rather fast, but the rest of the trajectories seem to be trapped in some misfolded intermediates. They explicitly showed two such examples (*B* and *C* intermediates in Ref. [24], replotted in Fig. 1) of the misfolded intermediates, which involve extra bends in the backbone chain. We carried out similar calculations [24], and found many intermediate states exist along the folding pathways [13,21,24]. Together with the results of GT, these can generally be classified into several conformational groups, as schematically represented in Fig. 1 [21,24]. One of the objectives of the present study is to see how the stability and dynamical roles of these intermediates change in the case of mechanical forcing.

III. EXTENDED *BPN* MODEL

To simulate the mechanical unfolding and refolding of a single protein molecule tethered to a solid surface and pulled

by an AFM tip, we extend the *BPN* model [11] by including a substrate surface and an AFM tip to the system. We model the substrate surface as a three-dimensional semi-infinite rigid lattice of atoms. Each bead of the protein chain, in principle, interacts with each atom of the surface lattice. If we consider an arbitrary bead of the *BPN* chain, its interaction with atoms of the surface can be approximated by a Lennard-Jones 6-12 potential, similar to Eq. (2.1). The integrated effects of all the atoms of the surface lattice on this bead is a 3-9 potential given by

$$V_{BB}=C_9(\sigma/z)^9-C_3(\sigma/z)^3, \tag{3.1}$$

where z is the vertical distance of this bead measured from the surface, and C_3 is equal to $(4\pi/9)\varepsilon_h\sigma^3$ and $C_9=5C_3$. It turns out that we have to switch off the attractive part of Potential (3.1), because otherwise the protein chain would quickly fall onto the surface, due to the long-range nature of the attractive part of the potential. The rationale for removing the attractive part comes from the fact that the surface attractive forces on the beads may be compensated for by the attractive interactions between the residues and the solvent molecules. Furthermore, to prevent the chain, drifting parallel to the surface, the tethered end of the bead is bound to a fixed point of the surface by a harmonic potential.

To simulate the AFM tip, we add an extra bead to the other end of the chain and assume that this tip bead is bound harmonically to the end of the chain. (However, to reduce high-frequency noise, the force constant is assumed to be only one tenth of that of a covalent bond). Furthermore, this tip bead interacts with all other beads in the chain. The interaction potential is assumed to be exactly the same as those between two neutral beads [Eq. (2.3)].

For the calculations reported below, we carry out molecular dynamics (MD) simulations using the velocity Verlet method to solve Hamilton's equations of motion for the beads [26,27]. As in Ref. [13], the masses of all beads are set to 40 a.u., and the time step of MD is 5 fsec. The MD runs are done isothermally, with the stochastic algorithm developed by Kast and co-workers [28,29]. The temperature is defined by the unit of the interaction energy, the well-depth parameter of the Lennard-Jones interaction between hydrophobic beads, ε_h, which is set to 121 K.

To simulate force-induced unfolding and refolding, we carry out isothermal MD runs by starting with a force-free native structure, and let the protein chain equilibrate at a desired temperature for a period of time. We then let the tip bead pull the chain vertically up from the substrate surface at a constant velocity, and compute the passive force of the chain as a result to this external forcing. In reality, this passive force is calculated as the net force acting on the tip bead by all beads of the chain. After the chain is fully extended, we reverse the direction of pulling, that is, downward at the same velocity, to allow the chain to have a chance to refold. After the same duration of pulling time as used for pulling up, we stop the motion and let the chain equilibrate again for a period of time. During this cycle of pulling up and down, we record the passive force of the chain as a function of the vertical distance of the tip to the surface to obtain a force-extension curve.

One should note here that there are more than one ways to simulate forced unfolding. For instance, Paci and Karplus [17] performed a calculation of the forced unfolding of a module of titin protein by a biased molecular dynamics simulation. A ratchetlike biased force was applied to the end-to-end coordinate. A steered molecular dynamics simulation was also applied to force-induced unfolding of titin's domains and of other globular proteins [18–20]. Our way of calculating the force is similar, but not identical, to the latter approach.

IV. THE FORCE-EXTENSION CURVES: HYSTERESIS LOOPS AND PHASE DIAGRAM TABLES

Using the extended *BPN* model described above, we have studied the mechanical unfolding and refolding of proteins. This was done by obtaining the force-extension curves for two different configurations: the tail-pulled case and the head-pulled case. In the tail-pulled case, the first bead is attached to the substrate surface and the 46th bead is pulled. In the head-pulled case, the roles of the first and 46th beads are reversed, that is, the 46th bead is attached to the surface and the first bead is pulled. We discuss the results of these two cases, separately, in the subsections below.

A. Tail-pulled case

To prepare for the initial ''native'' conformation, we start with a force-free native conformation, and let the chain of 46 beads equilibrate at a constant temperature for 2.5 nsec under the constraint that the first bead is tethered to the surface and the 46th bead is attached to the tip. The system quickly settles down to a nativelike conformation [see the structure at 10 Å of Fig. 5(a)], which is then used as an initial state for the forced unfolding and refolding calculations of this and following sections.

With the first bead attached to the surface, we pull the 46th bead upward at a constant velocity away from the surface. We observe that the unfolding always takes place with the fourth strand opening up first. This is reasonable, because the fourth strand is least tightly bound to the hydrophobic core, comprised of the first and third purely hydrophobic strands bound neatly together. As stated above, at a large end-to-end separation, we reverse the direction of pulling and pull down the chain for the same duration of time and observe how the chain refolds when the chain comes back close to the original end-to-end separation. The system is then allowed to equilibrate for a period of 2.5 nsec. During this loop cycle, we periodically sample the conformation and record the net force exerted on the tip bead as a function of the extension, represented by the z component of the tip coordinate.

We have studied the effects of the pulling rate on the force-extension curve of the extended *BPN* model. Some results of our simulations are plotted in Fig. 2 and summarized in Table I. In Fig. 2, we present the force-extension curves for several pulling rates (0.36, 0.54, 0.72, and 1.44

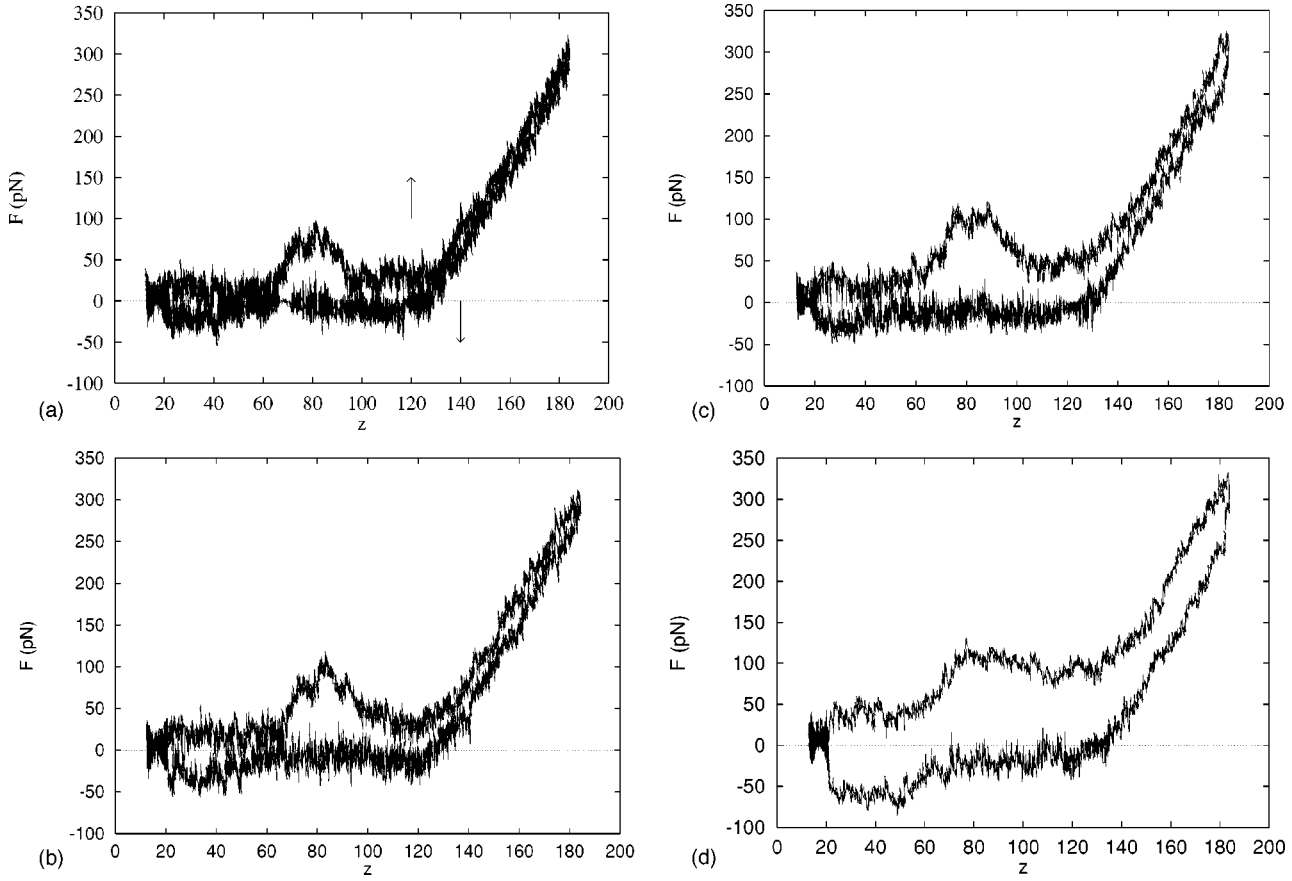

FIG. 2. The force-extension curve as a function of the pulling rate (P) at a temperature of $T=20$ K. (a) $P=0.36$ m/s, (b) $P=0.54$ m/s, (c) $P=0.72$ m/s, and (d) $P=1.44$ m/s. For this series the first bead is tethered to the solid surface (tail-pulled case). The unit of the horizontal axis is Å.

m/s) at a fixed temperature of 20 K. These pulling rates are 5 to 6 orders of magnitude larger than the experimental values, but are about 2–3 orders of magnitude smaller than those used in steered MD simulations [18,19]. All the curves in Fig. 2 are noisy, and show hysteresis. These curves almost, but not quite, form closed curves. Assuming they do, our results show that the enclosed area of the hysteresis loop increases with the pulling rate. The enclosed area measures the work or the energy dissipated during a pulling cycle. Thus the amount of energy dissipated increases with the pulling rate. The pulling-up trace shows several peaks, which are missing, in the pulling-down trace. Corresponding dynamics shows that these peaks are due to the resistance of the chain to break up the nucleus forming by the first and third strands. We see roughly three peaks, which may correspond to the three sliding intermediates (sliding between the first and third strands) seen in our previous study [24]. The maximal heights of these peaks also increase with pulling rate; that is, a protein shows greater resistance to unfolding as the pulling rate increases. A larger dissipated energy and a larger resistance to unfolding for greater pulling rate are interesting properties of noncovalent bonds in biomolecules, such as proteins, which are related to the strength or toughness of biomolecules [7].

We have also studied temperature effects on the force-extension curve. One obvious effect of increasing temperature is that the force-extension curve becomes increasingly noisy, i.e., the magnitude of fluctuations increases. Because of these thermal fluctuations, the enclosed area of the averaged curve of the hysteresis loop shrinks with temperature, and above a certain temperature hysteresis is no longer discernible. A sequence of plots demonstrating this phenomenon is given in Fig. 3, where we plot the force-extension curve at a fixed pulling rate (0.36 m/s) as a function of temperature. Another type of temperature effect is illustrated by the results listed in Table I, where we use symbols of Fig. 1 for intermediates to represent the conformation at the end of a loop cycle. This represents, in some sense, a phase diagram for our system. We start a loop with essentially the same initial nativelike conformation. Table I shows how the final conformation varies with the control parameters, the pulling rate and temperature. In Table I, symbol A denotes an ensemble of conformations which are nativelike, and symbol L represents an ensemble of conformations in which strands 2 and 3 stick together to form one pair and strands 1 and 4 form another pair. Altogether they form a geometry of the shape of an upside-down L (or Γ). An L-shaped conformation differs from an A type only in the fact that the (2,3) pair and (1,4) pair have not come together. If the (2,3) pair just comes down close to the surface to form an aggregate with the (1,4) pair, an L-shaped conformation has a good chance to become nativelike. Besides the letter symbols representing

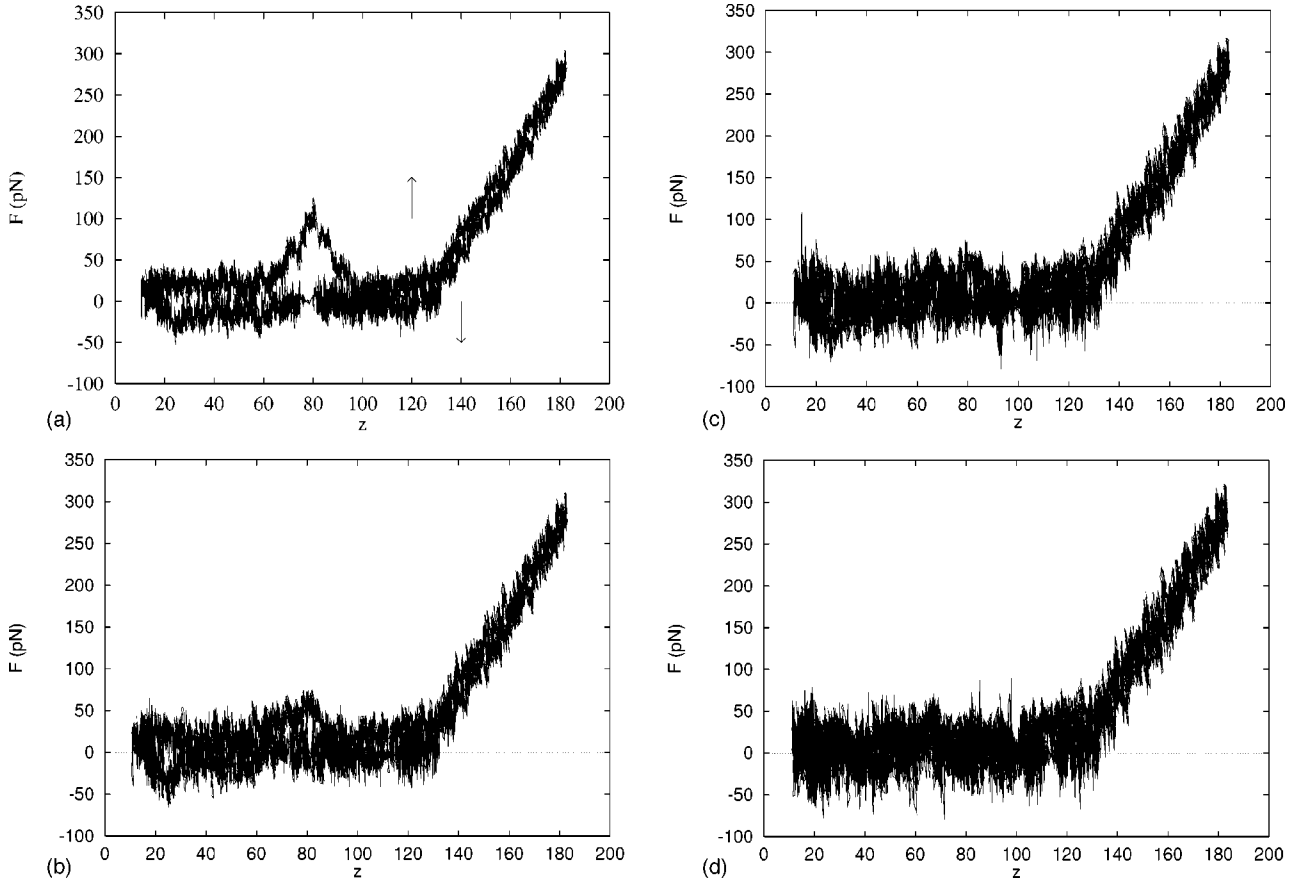

FIG. 3. The force-extension curve as a function of temperature for the tail-pulled case at a pulling rate of $P=0.36$ m/s. (a) $T=20$ K, (b) $T=40$ K, (c) $T=60$ K, and (d) $T=80$ K. The unit of the horizontal axis is Å.

the final configurations, we also include in the table the strand number of the strand which opens up first. Table I shows, as discussed above, that it is always the fourth chain which opens up first in the tail-pulled case. Table I also shows that at a lower pulling rate the final conformation of the *BPN* chain tends to be nativelike and, at a higher pulling rate, it tends to become *L* shaped. At higher temperature, thermal fluctuations are such that the conformation of the *BPN* chain jumps between a *C* conformation and an *L* or *A* conformation.

B. Head-pulled case

In this subsection we discuss results of the head-pulled case, where the 46th bead is attached to the surface, while the first bead is being pulled. As in Table I, Table II presents a partial ''phase diagram'' for the present case. We see in Table II that at a low pulling rate (for example, 0.36 m/s at 20 K), the fourth strand opens up first; at a large pulling rate (for example, 0.72 m/s or higher at 20 K), however, the first strand opens up first. At an intermediate pulling rate, where the transition takes place, we may observe cases where both the first and fourth strands open up simultaneously. If we define this rate as the transition pulling rate, we see that it increases with the temperature, if the temperature is below 80 K. This change of the pathways to unfolding can be easily understood in terms of the competition between the pulling rate and the rate of force transmission through the chain. At a low pulling rate, the force of pulling is transmitted through the more tightly bound core (made up of the first, second, and third strands) to the fourth strand, which is the most loosely bound strand, and the fourth strand opens up first. At a large pulling rate, on the other hand, there is not enough time for the force to transmit to the fourth chain, and the first strand is pulled apart first. Furthermore, the effect of temperature is such that thermal agitation makes the originally loosely bound fourth strand even more loosely bound, and thus easier to pull apart. However, when temperature becomes too high (80 K or higher), the first strand becomes almost as loosely bound as the fourth strand, thus both strands open up when the pulling rate is equal to or above 0.54 m/s.

At a fixed temperature of 20 K, in Fig. 4 we plot a series of force-extension curves as a function of pulling rate for the head-pulled case. As in Fig. 2, this series again show that the dissipated energy increases with the pulling rate and that the force-extension curve opens wider and becomes more negative, and thus moves farther away from forming a closed curve. The curve eventually closes onto itself, when the system is allowed to equilibrate for a period of time. To understand the origin of the large negative force, corresponding to a net upward force on the tip at a greater pulling rate, we examine the dynamics of the pulling cycle in more detail.

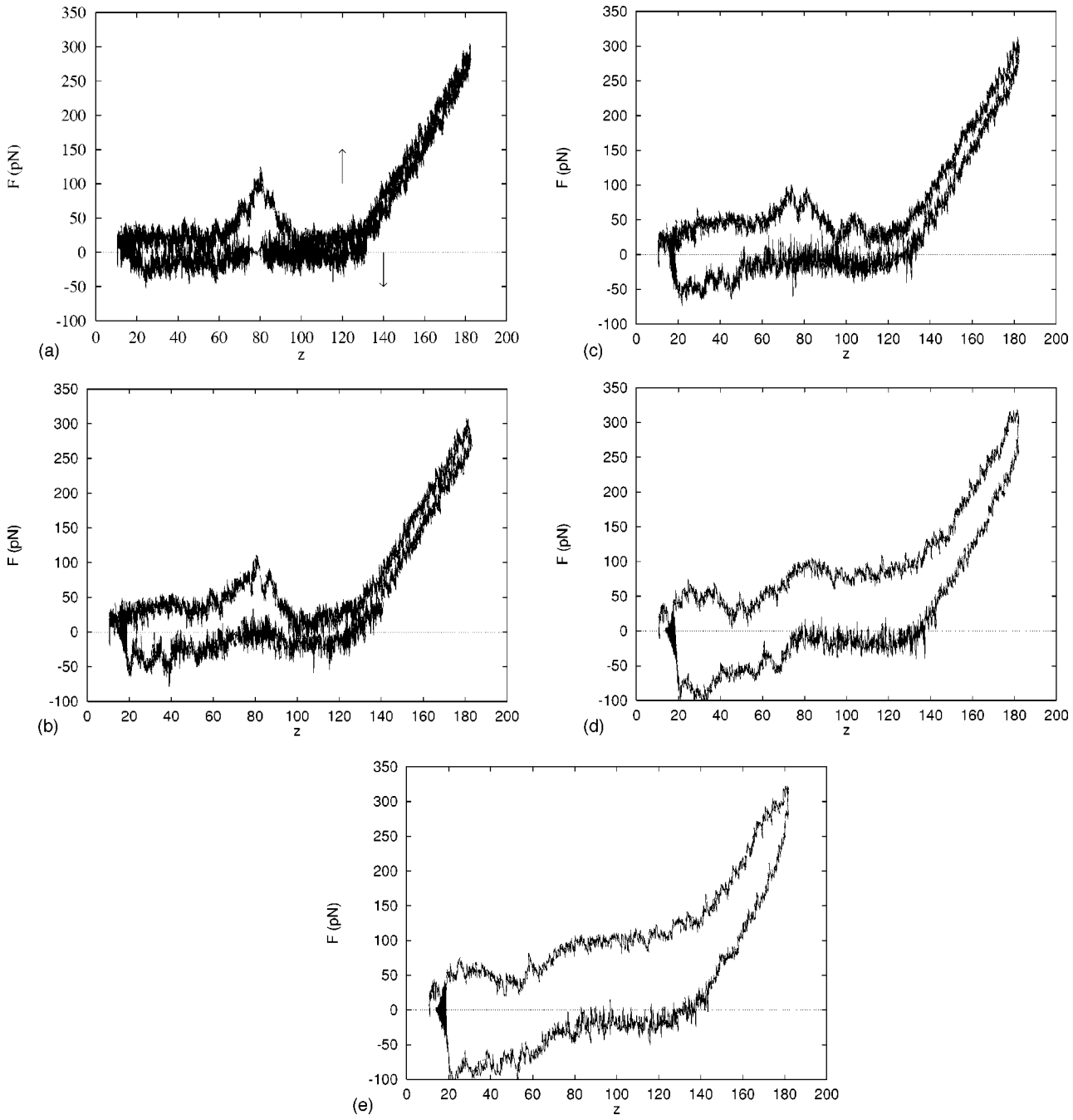

FIG. 4. The force-extension curves as a function of pulling rate at $T=20$ K. (a) $P=0.36$ m/s, (b) $P=0.54$ m/s, (c) $P=0.72$ m/s, (d) $P=1.44$ m/s, and (e) $P=1.80$ m/s. For this series the 46th bead is tethered to the solid surface (head-pulled case). The unit of the horizontal axis is Å.

We have found that the extended *BPN* model follows different refolding paths at different pulling rates. In Fig. 5, we show sequences of snapshots of conformations at several z values for two different cases of pulling rates: low at 0.36 m/s and high at 1.44 m/s. In Figs. 5(a) and 5(b), respectively, we present unfolding and refolding processes at a low pulling rate of 0.36 m/s, corresponding to the force-extension curve shown in Fig. 4(a). We see in Fig. 5(b) that the first and second strands remain attached while they are pulled down by the tip, and the final conformation of the refolding process resembles a nativelike structure. However, at a high pulling rate, for example, 1.44 m/s, as shown in Fig. 5(c), a resistance (or an upward force on the tip) starts to develop around 80 Å above the surface and reaches a maximum around 30 Å above the surface [see the pulling-down half of Fig. 4(d)]. The dynamic sequence in Fig. 5(c) shows that the first and second strands are torn apart by the tip beginning at around $z=80$ Å, and achieved around $Z\geqslant 40$ Å, and the first and third strands are pulled away from each other (but not separated) between 30 and 20 Å; the final conformation

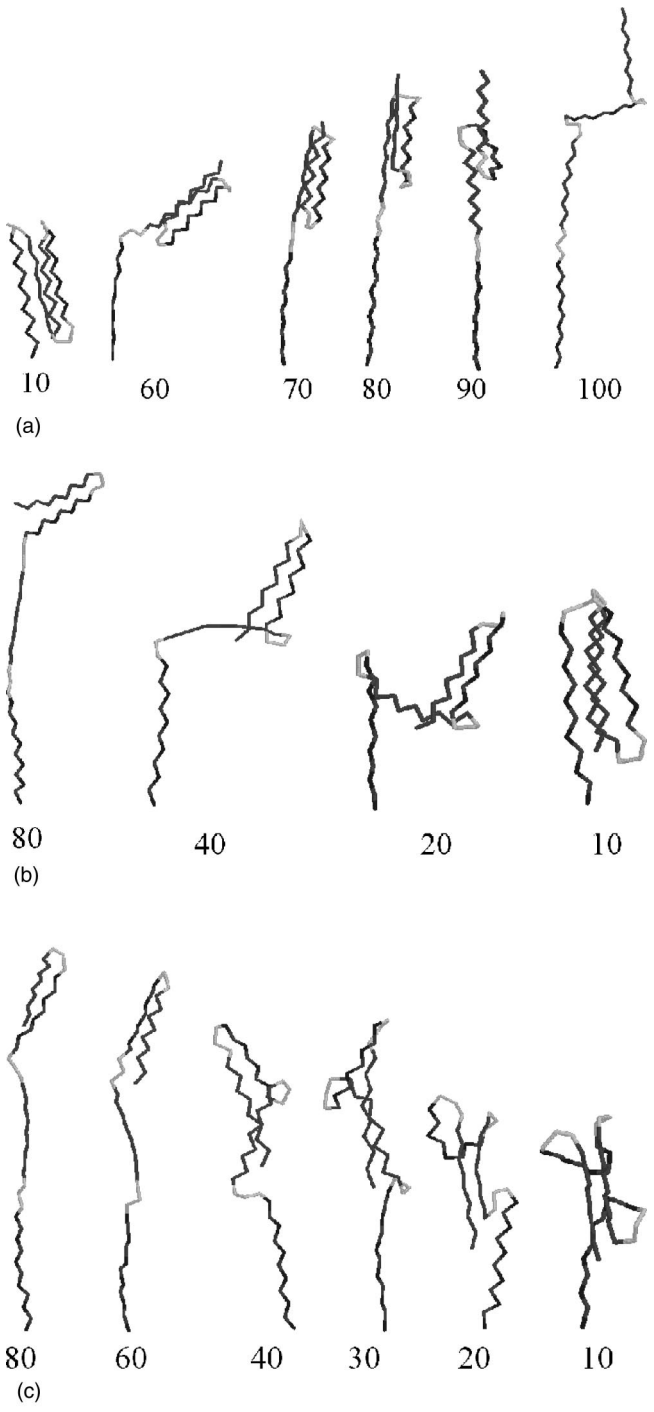

FIG. 5. (a) A sequence of structures in the pulling-up process, corresponding to the upper curve of Fig. 4(a). The structures from left to right are those at around 10, 60, 70, 80, 90, and 100 Å above the solid surface. (b) A sequence of structures in the pulling-down process, corresponding to the lower curve of Fig. 4(a). The structures from left to right are those at around 80, 40, 20, and 10 Å above the solid surface. (c) A sequence of structures in the pulling-down process, corresponding to the lower curve of Fig. 4(d). The structures from left to right are those at around 80, 60, 40, 30, 20, and 10 Å above the solid surface.

of this process resembles a *D*-type conformation. Therefore, the large negative force in the pulling-down process at a high pulling rate is generated by tearing apart the first and second and first and third strands.

TABLE I. The process and results of mechanical unfolding and refolding, with the first bead attached to the substrate surface and pulling the 46th bead at various pulling rates and temperatures. The number in the table indicates which strand opens up first. For example, ''4'' means that the fourth strand opens up first. The letters represent the basins of conformation groups which the system visits during the refolding process, for example, $(L \leftrightarrow C)$ indicates that the system moves between the basins of *L* and *C*.

Pulling rates and temperature	20 K	40 K	60 K
0.36 m/s	$4 \rightarrow A$	$4 \rightarrow A$	$4 \rightarrow A$
0.54 m/s	$4 \rightarrow A$	$4 \rightarrow L$	$4 \rightarrow C \rightarrow (A \leftrightarrow C)$
0.72 m/s	$4 \rightarrow L$	$4 \rightarrow L$	$4 \rightarrow (L \leftrightarrow C)$
1.44 m/s	$4 \rightarrow L$	$4 \rightarrow L \rightarrow C$	$4 \rightarrow (L \leftrightarrow C)$

As mentioned above in Table II the letter symbols *A*, *B*, *D*, *E*, and *L* (defined in Fig. 1 for the intermediates or misfolds for the force-free protein folding process) denote the final conformation group that the *BPN* chain ends up with after a pulling cycle. This table shows that the final conformation for the second type of attachment often assumes the misfold structure *D*, characterized by a bent second strand. Without an external force, the *D*-type structure is very unstable and can only be reached at temperature higher than 90 K [24]. Thus the external force stabilizes the *D* structure with reference to the nativelike structure. This may not be too surprising, because in a *D* structure both ends of the first and fourth strands can come together and point roughly in the same direction, favored by the pulling-down portion of the cycle. This is also true for structure *A*. However, the two ends point in opposite directions in structures *B* and *C*, which may explain why *B* and *C* structures rarely appear in Table II (or I) at lower temperature. At temperature of 80 K and higher, basins of attraction start to merge, we see more intermediates and competition of pathways.

V. DISCUSSIONS

The intrinsic differences in the interaction strengths among the four strands of the *PBN* model play a major role in the various phenomena observed in our mechanical unfolding and refolding simulations. As seen in Sec. III A, in the tail-pulled case the unfolding process always starts by opening up the last strand, but this is not always so in the head-pulled case. In the latter case, there is a competition between either the first or last strand to open up first, depending on the magnitude of the pulling rate. Thus mechanical unfolding and refolding is sensitive to the location of the AFM tip and to the way that a protein molecule is pulled. Furthermore, external forcing may alter the unfolding pathway.

Mechanical forcing induces some changes in the potential energy landscape of proteins. These changes are reflected in the temperature effects on the dynamical behavior of the *BPN* model. We describe some of these temperature effects below. At temperatures below 40 K and at a pulling rate below 0.72 m/s, the up-pulling part of the force-extension curves reveals the presence of several peaks in both head-

TABLE II. The same as Table I, but with the 46th bead attached to the substrate surface and pulling the first bead. The symbol ''1+4'' means that the first and fourth strands open up at the same time. Again, letters represent the basins of conformation groups which the system visits during the refolding process, for example, $(E \leftrightarrow D)$ indicates that the system moves between basins of E and D.

Pulling rates and temperature	20 K	40 K	60 K	80 K
0.36 m/s	$4 \rightarrow A$	$4 \rightarrow D$	$4 \rightarrow D$	$4 \rightarrow D \rightarrow B \rightarrow E \rightarrow C \rightarrow A$
0.54 m/s	$1+4 \rightarrow D$	$4 \rightarrow B$	$4 \rightarrow D$	$1+4 \rightarrow D \rightarrow B \rightarrow E$
0.72 m/s	$1 \rightarrow D$	$4 \rightarrow D$	$4 \rightarrow D$	$1+4 \rightarrow D \rightarrow B \rightarrow E$
1.44 m/s	$1 \rightarrow D$	$1 \rightarrow L$	$1 \rightarrow L \rightarrow E$	$1+4 \rightarrow L \rightarrow D \rightarrow (E \leftrightarrow D)$

and tail-pulled cases. These peaks represent the response of the folding nucleus, formed by an aggregation of the first and third strands, to the mechanical force, and originate from the same sliding motion between the first and third strands observed in the force-free folding process [24]. These peaks disappear at temperatures above 40 K, because the thermal fluctuations destroy the integrity of the folding nucleus. A previous study of the force-free case [24] revealed that the integrity of the folding nucleus is in danger only when temperature is above 80 K, much higher than 40 K found in the present case. Table II indicates that the protein undergoes rapid conformational changes at 80 K, but this phenomenon appears only above 95 K in the force-free case [24]. It seems natural to attribute the above differences between the forced and unforced cases to force-induced changes in the potential energy surface of the protein molecule; in particular, it seems that the external force has lowered the energy barriers separating stable conformations.

The pulling rate is another factor influencing the way by which a protein molecule explores its energy landscape. At low pulling rates, the resistance of the folding nucleus to external force, i.e., the barriers against the sliding motion between the first and third strands, dominates the pulling-up process; however, there is no such resistance observed in the pulling-down process. As the pulling rate increases, the rate of transmission of the pulling force through the chain becomes slower in comparison and the sliding motion is gradually replaced by the peeling motion; in this case resistance to the sliding motion appears in the pulling-down process as the negative passive force observed in the force-extension curves.

To understand the hysteretic behavior obtained in our simulations, At this point we introduce a simple phenomenological model of bistability in terms of a cusp catastrophe [30–32], shown schematically in Fig. 6(a). If both the upper and lower branches of the catastrophe are stable and the middle branch is unstable, a cycle of pulling will result in a hysteresis loop, shown in Fig. 6(b), qualitatively similar to those seen in Fig. 2. The effect of temperature, on the other hand, is that we have only a fuzzy hysteresis loop, because thermal noise can induce early transitions from the upper branch to the lower branch. The net effect is that the hysteresis loop shrinks as temperature increases, as seen in Fig. 3. Furthermore, since the hysteresis loop increases with the pulling rate P, the axis perpendicular to the F and z axes of Fig. 6(a) can be identified as the pulling rate. Therefore, a cusp catastrophe seems to be useful in qualitatively describing some of the results obtained in our simulation. For a system with several folded domains, it is possible that we have a series of such cusp regions [32], and the pulling trace of the force-extension curve will then become sawtoothlike [Fig. 6(c)], as seen in experiments [1–6].

The bistability that we discuss above is between a folded (upper branch) and an unfolded structure (lower branch). In reality, we have multistability, because there exist, for instance, for the present system, several folded structure groups, as shown in Fig. 1. How these states are interconnected can be complicated. In these complicated cases, higher-order catastrophes [30,31] are needed for description. However, the picture we suggest here, although incomplete, does fit some of the simulated and experimental results ob-

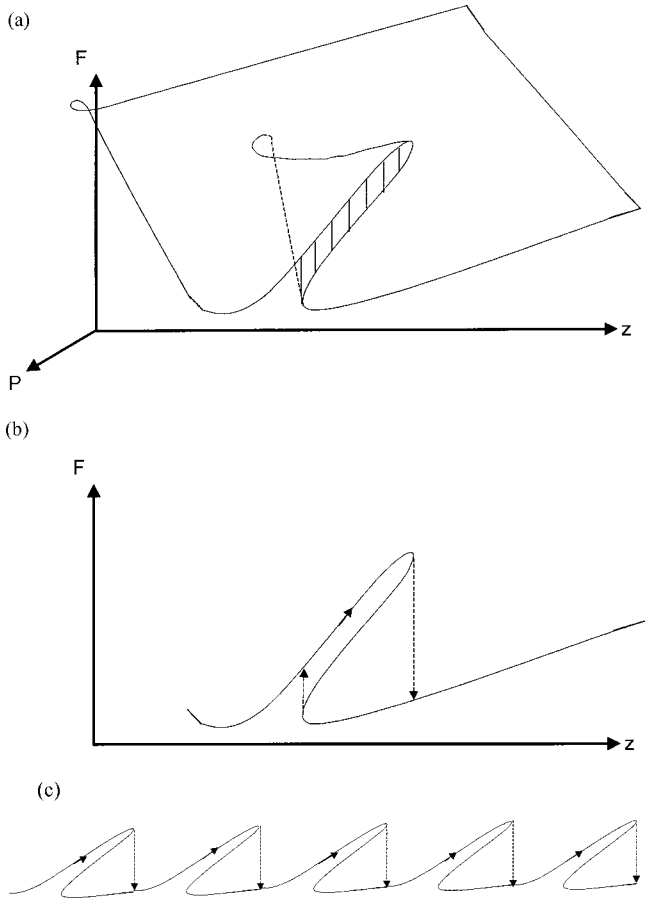

FIG. 6. (a) A schematic plot of a cusp catastrophe, where force F is considered as a function of the extension z and the pulling rate P (b) The hysteresis loop implied by the cusp catastrophe at zero temperature. (c) A case of multiple cusp regions.

tained so far. It would be interesting to improve upon it and to find its microscopic foundation.

Analysis of the results of our simulations shows that external forcing, the pulling rate, and the pulling location on the protein molecule can induce changes in the energy landscape. These changes in turn result in shifts of the relative stabilities among intermediates. In a force-induced unfolding and refolding experiment, we explored the dynamical behavior of a protein molecule along certain reaction coordinates, which may or may not be a pathway belonging to the force-free system. However, if we can express the dynamical properties as functions of the force, we can recover the dynamical behavior of the force-free system in the $F \rightarrow 0$ limit. In principle, we can make use of this fact to investigate the uncharted regions of the energy landscape related to the force-free folding process. Furthermore, we can, in principle, use this information to manipulate the energy landscape to enhance certain binding reactions [33].

During the time that this article was being revised, a paper by Klimov and Thirumalai (KT) [20] appeared in the literature, which, among other systems, also investigated mechanical unfolding of the *BPN* model. There is some overlap between their work and the present one, but there also exist some main differences, which we describe below. A protein molecule is pulled at both ends by a harmonic force in KT, while in our simulations we explicitly consider the interaction potential between each bead and atoms of a substrate surface. Furthermore, we represent the AFM tip as an extra bead and the force on this tip is then the force exerted on the AFM cantilever by the protein molecule. Our pulling rates are about a factor of 2–50 times larger. This effect, plus the relative temperature of the two systems, may explain why they see a clear peak, corresponding to the unfolding of the fourth strand, while our force-extension curve seems to show only a very broad peak ($z \sim 25$ Å in the tail-pulled case), which is hardly discernible. Finally, we have clearly seen a hysteresis loop generated by the pulling cycle, but such a phenomenon was not discussed in the work of KT. Their work is interesting, because they provided strong evidence for the hypothesis that the unfolding pathway is mainly determined by the topology of the native structure of protein.

VI. CONCLUSION

By varying three control factors, i.e., the pulling rate, the temperature, and the way in which protein is pulled, we explored the response of noncovalent forces during mechanical unfolding and refolding of a model protein. Our results indicate that both external forcing and the pulling location can induce modifications of the energy landscape, and thus cause shifts of the relative stability among intermediates. During the up-pulling process, there are barriers to the unfolding of the backbone due to the disintegration of the nucleation core. No barriers, however, were observed in the pulling-down process (at a low pulling rate). The resulting hysteresis reflects the fact that the unfolding and refolding follow different pathways.

ACKNOWLEDGMENTS

We thank Dr. Guoliang Yang for stimulating discussions on his AFM experiments. This work was supported by the National Research Council of Taiwan under the Contract No. NSC 88-2811-M-002-0030.

[1] M. R. Rief, M. Gautel, F. Oesterhelt, J. M. Fernandez, and H. E. Gaub, Science **276**, 1109 (1997).
[2] G. Yang, C. Cecconi, W. A. Baase, I. R. Vetter, J. A. Haack, B. W. Matthews, F. W. Dahlquist, and C. B. Bustamante, Proc. Natl. Acad. Sci. U.S.A. **97**, 139 (2000).
[3] P. E. Marszalek, H. Lu, H. Li, M. Carrion-Vazquez, A. F. Oberhauser, K. Schulten, and J. M. Fernandez, Nature (London) **420**, 100 (1999).
[4] A. F. Oberhauser, P. E. Marszalek, M. Carrion-Vazquez, and J. M. Fernandez, Nat. Struct. Biol. **6**, 1025 (1999).
[5] M. Carrion-Vazquez, A. F. Oberhauser, S. B. Fowler, P. E. Marszalek, S. E. Broedel, J. Clarke, and J. M. Fernandez, Proc. Natl. Acad. Sci. U.S.A. **96**, 3694 (1999).
[6] M. S. Z. Kellermayer, S. B. Smith, H. L. Granzier, and C. Bustamante, Science **276**, 1112 (1997).
[7] D. A. Simson, M. Strigl, M. Hohenadl, and R. Merkel, Phys. Rev. Lett. **83**, 652 (1999).
[8] B. L. Smith, T. E. Schafer, M. Viani, J. B. Thompson, N. A. Frederick, J. Kindt, A. Belcher, G. D. Stucky, D. E. Morse, and P. E. Hansma, Nature (London) **399**, 761 (1999).
[9] H. Li, B. Liu, X. Zhang, C. Gao, J. Shen, and G. Zhou, Langmuir **15**, 2120 (1999).
[10] A. Kolinsky, J. Skolnick, and R. Yaris, Biopolymers **26**, 937 (1987).
[11] J. D. Honeycutt and D. Thirumalai, Proc. Natl. Acad. Sci. U.S.A. **87**, 3526 (1990).
[12] Z. Guo and D. Thirumalai, J. Chem. Phys. **97**, 525 (1992).
[13] R. S. Berry, N. Elmaci, J. P. Rose, and B. Vekhter, Proc. Natl. Acad. Sci. U.S.A. **94**, 9520 (1997).
[14] H. Nymeyer, A. E. Garcia, and J. Onuchic, Proc. Natl. Acad. Sci. U.S.A. **95**, 5921 (1998).
[15] N. D. Socci, J. Onuchic, and P. Wolynes, Proc. Natl. Acad. Sci. U.S.A. **96**, 2031 (1999).
[16] D. K. Klimov and D. Thirumalai, Proc. Natl. Acad. Sci. U.S.A. **96**, 6166 (1999).
[17] E. Paci and M. Karplus, J. Mol. Biol. **288**, 441 (1999).
[18] H. Lu, B. Isralewitz, A. Krammer, V. Vogel, and K. Schulten, Biophys. J. **75**, 662 (1998).
[19] H. Lu and K. Schulten, Biophys. J. **79**, 51 (2000); Proteins: Struct., Funct., Genet. **35**, 453 (1999).
[20] D. K. Klimov and D. Thirumalai, Proc. Natl. Acad. Sci. U.S.A. **97**, 7254 (2000).
[21] Z. Guo and D. Thirumalai, Biopolymers **36**, 83 (1995).
[22] D. Thirumalai and D. K. Klimov, Curr. Opin. Struct. Biol. **9**, 197 (1999).
[23] J. E. Shea, Y. D. Nochomovit, Z. Y. Guo, and C. L. Brooks III, J. Chem. Phys. **109**, 2895 (1998).

[24] F. Y. Li, J. M. Yuan, C. Y. Mou, R. S. Berry, and B. Vehkter (unpublished).
[25] T. Veitshans, D. K. Klimov, and D. Thirumalai, Fold. Des. **2**, 1 (1997).
[26] O. M. Becker, Proteins: Struct., Funct., Genet. **27**, 213 (1997).
[27] M. P. Tildesley and D. J. Allen, *Computer Simulation of Liquids* (Oxford University Press, Oxford, 1989).
[28] S. M. Kast, L. Nicklas, H. J. Bar, and J. Brickmann, J. Chem. Phys. **100**, 566 (1994).
[29] S. M. Kast and J. Brickmann, J. Chem. Phys. **104**, 3732 (1996).
[30] T. Poston and I. Stewart, *Catastrophe Theory and its Applications* (Pitman, Boston, 1978).
[31] R. Gilmore, *Catastrophe Theory for Scientists and Engineers* (Wiley, New York, 1981).
[32] J. M. Yuan and G. C. Lie, Adv. Chem. Phys. **85**, 1 (1994).
[33] E. Freire, Proc. Natl. Acad. Sci. U.S.A. **96**, 10 118 (1999).