

Single-molecule pulling: phenomenology and interpretation

Ignacio Franco, Mark A. Ratner, and George C. Schatz

Department of Chemistry, Northwestern University, Evanston, Illinois 60208-3113

Single-molecule pulling techniques have emerged as versatile tools for probing the noncovalent forces holding together the secondary and tertiary structure of macromolecules. They also constitute a way to study at the single-molecule level processes that are familiar from our macroscopic thermodynamic experience. In this Chapter, we summarize the essential phenomenology that is typically observed during single-molecule pulling, provide a general statistical mechanical framework for the interpretation of the equilibrium force spectroscopy and illustrate how to simulate single-molecule pulling experiments using molecular dynamics.

Published in *Nano and Cell Mechanics: Fundamentals and Frontiers*, edited by H.D. Espinosa and G. Bao (Wiley, Microsystem and Nanotechnology Series, 2013), Chap. 14, pages 359–388.

Keywords: Molecular dynamics, macromolecules, mechanical properties, finite system thermodynamics, free energy reconstruction

Contents

I. Introduction	2
II. Force-extension behavior of single molecules	4
III. Single-molecule thermodynamics	8
A. Free energy profile of the molecule plus cantilever	9
B. Extracting the molecular potential of mean force $\phi(\xi)$	10
C. Estimating force-extension behavior from $\phi(\xi)$	13
IV. Modeling single-molecule pulling using molecular dynamics	14
A. Basic computational setup	14
B. Modeling strategies	15
C. Examples	17
1. Direct modeling: Oligorotaxanes	17
2. Indirect modeling: DNA dimer	19

V. Interpretation of pulling phenomenology	22
A. Basic structure of the molecular potential of mean force	22
B. Mechanical instability	23
C. Dynamical bistability	26
VI. Summary	28
Acknowledgments	30
References	30

I. INTRODUCTION

In recent years a number of techniques have been developed that allow us to access the properties of single molecules (see, e.g., Refs. [1–8]). These windows into the single-molecule world are teaching us with unprecedented detail how molecules behave, including how they move, their mechanical behavior, their conductivity, reactivity, optical properties, and other properties. While measurements made on bulk matter are averages over a whole ensemble of molecules, single-molecule measurements highlight the contributions of the constituent parts to the ensemble. A key feature of the properties of single molecules is that fluctuations in the observables are comparable with the average values. This contrasts with macroscopic behavior (where such fluctuations are usually negligible) and indeed fluctuation-induced emergent phenomena can arise that, simply put, cannot be observed in the macroscale.

In this Chapter we discuss basic aspects of single-molecule pulling experiments in which laser optical tweezers or atomic force microscopes (AFMs) are employed to exert mechanical forces on single molecules [2, 9–12], and the resulting stress/strain response is used as a probe of the molecular potential energy surface. The focus will be on how to simulate single-molecule pulling experiments using molecular dynamics, on what class of information can be extracted from the resulting force-extension isotherms and on how to interpret the basic phenomena typically observed upon pulling.

Figure 1 shows a schematic of a single-molecule pulling experiment using AFMs (an equivalent setup exists for experiments that employ laser optical tweezers). In it, one end of a molecule is attached to a surface, while the other end is attached to an AFM tip coupled to a cantilever. The distance between the surface and the cantilever L is controlled while the molecular end-to-end distance $\xi(t)$ is allowed to fluctuate. The pulling is performed by varying $L(t)$ in some prescribed

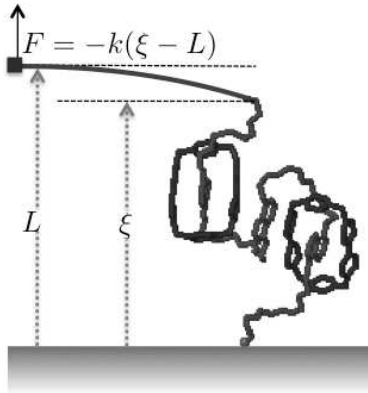


FIG. 1: Schematic of a single-molecule pulling experiment using AFMs. One end of the molecular system is attached to a surface and the other end to an AFM tip attached to a cantilever. During the pulling, the distance between the surface and the cantilever L is controlled while the deflection of the cantilever from its equilibrium position measures the instantaneous applied force $F(t) = -k(\xi(t) - L)$, where k is the cantilever spring constant and $\xi(t)$ the fluctuating molecular end-to-end extension.

fashion. The force F exerted during the process is determined by measuring the deflection of the cantilever from its equilibrium position $F = -k(\xi - L)$, where k is the cantilever stiffness. For a discussion of experimental aspects of single-molecule force spectroscopy see Ref. [13].

The experiments exert mechanical control over the molecular conformation while simultaneously making thermodynamic measurements of any mechanically-induced unfolding events. The stress maxima observed during pulling have been linked to the breaking [11, 14] of noncovalent interactions, thus providing insight into the secondary and tertiary structures of macromolecules. From the force-extension data it is possible to reconstruct the molecular free energy profile along the extension coordinate, quantifying in this way thermodynamic changes undergone by the molecule during folding.

The structure of this Chapter is as follows: Section II introduces the basic phenomenology observed during single molecule pulling through a discussion of representative experiments. Section III establishes the statistical mechanical theory behind single-molecule pulling experiments and describes a method for reconstructing the molecular free energy profile along the extension coordinate from the force-extension data. Section IV, in turn, introduces direct and indirect methods that can be used to model single-molecule pulling experiments using molecular dynamics. Section V focuses on the interpretation of single-molecule pulling experiments through an analysis of the basic structure of the molecular free energy profile. Section VI summarizes the main topics discussed in this Chapter.

II. FORCE-EXTENSION BEHAVIOR OF SINGLE MOLECULES

In this section, we consider a series of examples of single-molecule pulling experiments that illustrate the basic phenomenology that can be observed. Consider first the mechanically induced unfolding of a RNA hairpin [9]. The hairpin is attached to polystyrene beads via RNA/DNA hybrid handles and the pulling is performed in an optical tweezers arrangement. Figure 2 shows the experimentally determined force-extension behavior and the details of the hairpin structure. For fast pulling speeds hysteresis in the force-extension traces is evidenced by the fact that the traces observed upon pulling and subsequent contraction do not coincide. However, upon reduction of the pulling speed the traces obtained during extension and contraction do coincide, indicating that for all practical purposes the process occurs under reversible conditions. During pulling the force initially increases monotonically, then at ~ 14 pN it shows a drop, and then it increases again. In the process, the RNA hairpin undergoes a conformational transition from the hairpin structure to an extended coil. The drop in the force corresponds to the molecular unfolding event where the hydrogen bonding network holding the RNA hairpin together is broken. The local maximum in the force-extension isotherm is interpreted as the force required to break the secondary structure of the molecule.

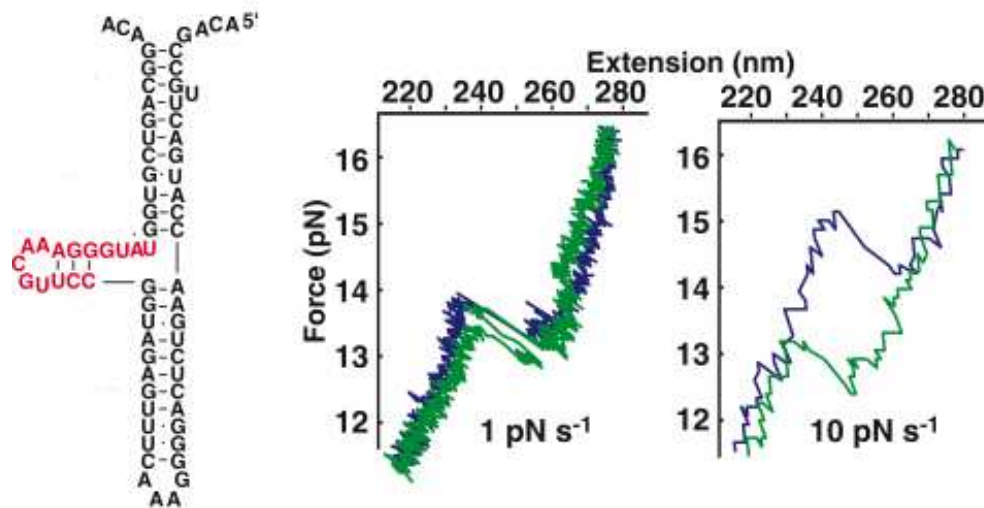


FIG. 2: Force-extension behavior of a RNA hairpin obtained in an optical tweezer setup. The left panel shows the sequence and secondary structure of the main structural motif of the pulled RNA hairpin. The right plots detail the force-extension curves during stretching (blue) and contraction (green) when pulling under equilibrium (left) and nonequilibrium (right) conditions. Note that the x -axis is the molecular end-to-end distance ξ and not the distance between the surface and the cantilever L . Figure adapted from [9].

The force-extension behavior of the RNA hairpin illustrates two basic phenomena that are commonly observed during single-molecule pulling. During pulling there is a region of *mechanical instability* where the average force decreases with the extension $\frac{\partial \langle F \rangle L}{\partial L} < 0$. This region occurs when the pulled molecule undergoes a conformational transition from a stable folded structure to an extended system. Note that around the unfolding region the molecule hops between the folded and the unfolded state. That is, around the region of mechanical instability the molecule plus cantilever system exhibits *dynamical bistability* where the molecule unfolds and refolds, leading to blinking in the force measurement between a stronger force when the molecule is folded and a weaker force regime when extended. In Sec. V we discuss the origin of these two phenomena.

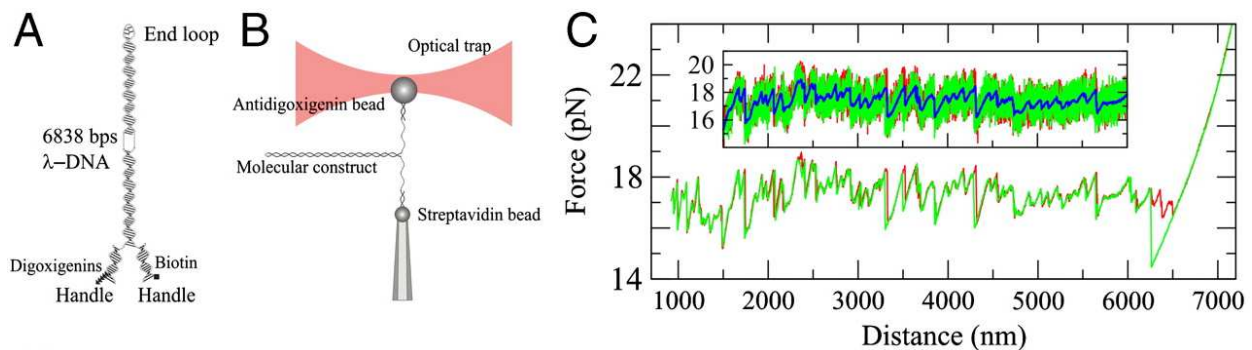


FIG. 3: In this experiment the 6838 base pair long DNA hairpin schematically shown in (A) is pulled in the optical tweezers arrangement in (B). The pulling is done by moving the optical trap relative to the pipette at low pulling speeds (10 nm/s). Panel (C) shows the force-extension profile during unfolding (in red) and refolding (in green). The inset shows the raw data obtained while the smooth curves are time-averaged. Except for some hysteresis in the last rip, the curves obtained during pulling and relaxing are almost identical. Figure adapted from [15].

Figure 3 shows the force-extension characteristics of a related system, a 6838 base-pair long DNA hairpin, also obtained in a optical tweezers arrangement [15]. Even for this complex system it is possible to find experimentally accessible pulling speeds (10 nm/s) where reversible behavior is recovered in the mechanically induced melting process. Melting in this case refers to dehybridization of the base pairs, and we see that unlike Fig. 2, there is an extended unfolding region where the average force is in the 15-19 pN range while the length L increases. This corresponds to sequential dehybridization (believed to occur a few base pairs at a time) rather than a concerted breaking of all the base pairs at the same time. The inset in the figure shows the raw data obtained in the experiment while the smooth curves shown are after further temporal averaging. The averaged data evidences the changes in the pulling force required to mechanically induce melting. It quantifies

the strength of the type of hydrogen bonds encountered as the elongation proceeds.

The raw data evidences yet another fundamental aspect of this class of experiments: the substantial thermal fluctuations in the force measurements that are typically observed during pulling indicating that the system is not in the thermodynamic limit. This observation has the important consequence of leading to a non-equivalence between different statistical ensembles. In particular, it leads to a non-equivalence between the isometric ensemble (where L is controlled and F fluctuates) and the isotensional ensemble (where F is controlled and L fluctuates). Throughout we will focus on the isometric ensemble since is the most common way to perform these experiments. For recent discussions of the nonequivalence between ensembles see [16, 17]. See [18] for an interpretation of the pulling phenomenology in the isotensional ensemble.

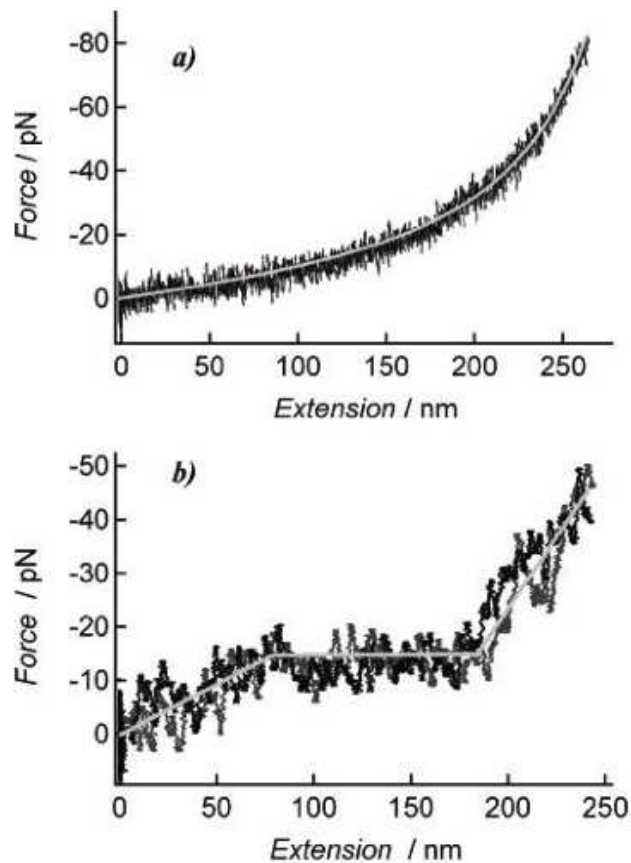


FIG. 4: Reversible pulling of individual polystyrene chains in toluene (top panel) and water (bottom panel) obtained using an AFM setup. Figure adapted from [19].

As a third example, consider the reversible force-extension behavior of a synthetic polymer pulled in poor and good solvent conditions using an AFM [19]. Figure 4 shows the force-extension isotherm of an individual polystyrene (PS) chain pulled in toluene (good solvent) and water (poor

solvent). In toluene (Fig. 4a), the deformation of PS exhibits a worm-like chain [20] behavior with a persistence length of (0.25 ± 0.05) nm. In water (Fig. 4b) the behavior is qualitatively different. In this case, the force initially rises linearly up to ~ 13 pN, then exhibits a plateau force at ~ 13 pN, followed by a third regime where the force rises again. This behavior is well predicted by a model [21] in which the initial linear restoring force is due to deformation of a collapsed spherical globule to an ellipsoid during which the polymer's surface energy increases. In this model, the plateau is associated with a conformational transition involving the coexistence of the globule and a stretched chain of polymer or polymer globules. For large stretching, where coexistence is no longer possible, the force extension behavior of a Gaussian chain [20] is recovered. Note that large-scale fluctuations in the force measurements are also evident in this system.

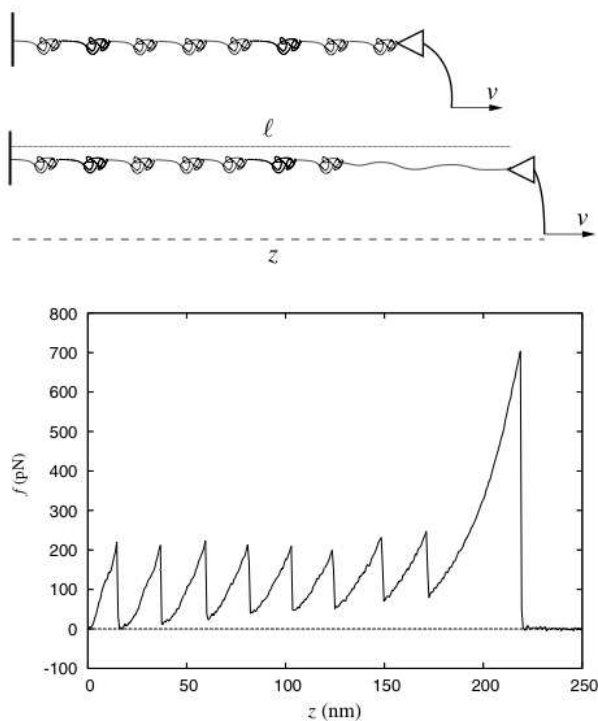


FIG. 5: Top panel: schematic of the protein attached to the AFM tip. Bottom panel: force-extension behavior of the polyprotein composed of eight repeat of the Ig27 domain. Figure taken from [22].

Consider last the forced unfolding of a recombinant polyprotein composed of eight repeats of the Ig27 domain of human titin obtained in an AFM setup [22]. Figure 5 shows the force-extension profile when pulling at a constant speed under nonequilibrium conditions. As the polyprotein is extended the force exerted on the molecule increases until one of the Ig27 domains suddenly unfolds. Unfolding of a domain reduces the overall force, but with increasing extension another local subunit unfolds. This occurs until all the domains are unfolded and the protein-tip interaction is broken.

This results in the sawtooth pattern shown in Fig. 5 where each peak in the force-extension profile marks the unfolding of one subunit in the polyprotein and the last peak signals the detachment of the protein from the tip.

III. SINGLE-MOLECULE THERMODYNAMICS

In what follows we specify what we mean about thermodynamics in the context of single-molecule pulling experiments [12] and the class of thermodynamic information that is accessible from the force-extension isotherms. The thermodynamic system consists of the molecule plus cantilever coupled to a thermal bath, typically the surrounding solvent. As such, the equilibrium state of the system is well described by the canonical ensemble with configurational partition function

$$Z(L) = \int d\mathbf{r} \exp[-\beta U_L(\mathbf{r})], \quad (1)$$

where \mathbf{r} denotes the coordinates of the N atoms of the molecule, $\beta = 1/k_B T$ is the inverse temperature, and

$$U_L(\mathbf{r}) = U_0(\mathbf{r}) + V_L[\xi(\mathbf{r})] \quad (2)$$

is the potential energy function of the molecule plus restraints. Here $U_0(\mathbf{r})$ is the molecular potential energy plus potential restraints not varied during the experiment, and $V_L[\xi(\mathbf{r})]$ is the potential due to the cantilever at extension L . To a good approximation, the quantity $V_L[\xi(\mathbf{r})]$ is given by a harmonic function

$$V_L[\xi(\mathbf{r})] = \frac{k}{2} [\xi(\mathbf{r}) - L]^2 \quad (3)$$

of stiffness k . The instantaneous force exerted by the cantilever on the molecule is determined by measuring the deflection of the cantilever from its equilibrium position and it is given by

$$F = -\nabla V_L = \frac{\partial V_L}{\partial L} = -k [\xi(\mathbf{r}) - L], \quad (4)$$

where the gradient ∇ is with respect to the $[\xi(\mathbf{r}) - L]$ coordinate. In writing Eqs. (3) and (4) the vector nature of L , $\xi(\mathbf{r})$ and F has been obviated since these quantities are typically collinear in the single-molecule pulling setup. For notational simplicity, in Eq. (1) and throughout this manuscript we ignore the constant multiplicative factor that makes partition functions dimensionless since they are irrelevant for determining relative free energy changes.

A. Free energy profile of the molecule plus cantilever

Because the state of the molecule plus cantilever under equilibrium conditions is well described by a canonical ensemble, the force exerted during the pulling yields the associated change in the Helmholtz free energy

$$\Delta A = A(L) - A(L_0) = -\frac{1}{\beta} \ln \frac{Z(L)}{Z(L_0)} \quad (5)$$

for the composite molecule plus cantilever system. To see this, note that in the canonical ensemble the average force at a given extension L is given by

$$\langle F \rangle_L = \frac{1}{Z(L)} \int d\mathbf{r} \left(\frac{\partial V_L}{\partial L} \right) \exp\{-\beta U_L(\mathbf{r})\} = -\frac{1}{\beta} \frac{1}{Z(L)} \frac{\partial Z(L)}{\partial L} = \frac{\partial A}{\partial L}. \quad (6)$$

Therefore, if the pulling is done under reversible conditions, such that the state of the system is well described by a canonical ensemble at each step of the pulling, the change in A when pulling from L_0 to L is determined by the reversible work exerted in the process

$$\Delta A = \int_{L_0}^L \frac{\partial A}{\partial L'} dL' = \int_{L_0}^L \langle F \rangle_{L'} dL' = W_{\text{rev}}. \quad (7)$$

Remarkably, even when the pulling is performed under nonequilibrium conditions it is still possible to determine the free energy changes during pulling by means of the Jarzynski non-equilibrium work fluctuation relation [23–25]

$$\exp[-\beta \Delta A] = \langle \exp[-\beta W] \rangle_{\text{non-eq}} \quad (8)$$

where the average in this expression is over nonequilibrium realizations of a given pulling protocol starting from a system initially prepared in the canonical ensemble. Eq. (8) relates the nonequilibrium work with the equilibrium free energy changes of the system and holds arbitrarily far away from equilibrium. Note that by using Jensen's inequality ($e^{\langle x \rangle} \leq \langle e^x \rangle$), Eq. (8) implies that the average nonequilibrium work is greater than or equal to the free energy change

$$\langle W \rangle_{\text{non-eq}} \geq \Delta A, \quad (9)$$

which is the usual statement of the second law of thermodynamics

The above analysis associates a free energy change to the pulling process although the system is not in the thermodynamic limit. This association relies on the assumption that the equilibrium state of the system is well described by a classical canonical ensemble. This feature has been experimentally tested: In [26] the authors experimentally demonstrated the validity of the Jarzynski

equality [Eq. (8)]. The Jarzynski equality supposes that the initial equilibrium state of the system is given by a classical canonical partition function and hence the test of the Jarzynski equality also tests whether this is a good description of the equilibrium state of the system.

B. Extracting the molecular potential of mean force $\phi(\xi)$

The properties that are measured during the pulling are those of the molecule plus cantilever. Thus, direct integration of the force-extension isotherms yields the free energy profile of the composite system [see Eq. (7)]. However, what one is really interested in are the properties of the molecule itself. Specifically, one is interested in the molecular Helmholtz free energy profile (also known as the potential of mean force PMF [27]) along the extension coordinate ξ . The PMF succinctly captures thermodynamic changes undergone by the molecule during folding and determines the basic phenomenology that is observed upon pulling [18, 28].

A useful method to extract the molecular PMF $\phi(\xi)$ from the force versus extension measurements is the Weighted Histogram Analysis Method (WHAM) [29–33]. The WHAM analysis combines all the force-extension data collected during the pulling to properly remove the bias due to the cantilever and extract the PMF. In fact, one can view the single-molecule pulling process as an experimental realization of the WHAM methodology using harmonic potential restraints. We now describe how to extract the PMF from equilibrium force-extension data using the WHAM. A schematic of the process is shown in Fig. 6.

The PMF is defined by [31]

$$\phi(\xi) = \phi(\xi^*) - \frac{1}{\beta} \ln \left[\frac{p_0(\xi)}{p_0(\xi^*)} \right], \quad (10)$$

where ξ^* and $\phi(\xi^*)$ are arbitrary constants and

$$p_0(\xi) = \frac{\int d\mathbf{r} \delta[\xi - \xi(\mathbf{r})] \exp[-\beta U_0(\mathbf{r})]}{\int d\mathbf{r} \exp[-\beta U_0(\mathbf{r})]} = \frac{Z_0(\xi)}{Z_0} = \langle \delta[\xi - \xi(\mathbf{r})] \rangle. \quad (11)$$

Here,

$$Z_0 = \int d\mathbf{r} \exp[-\beta U_0(\mathbf{r})] \quad (12)$$

is the configurational partition function of the molecule, and

$$Z_0(\xi) = \int d\mathbf{r} \delta[\xi - \xi(\mathbf{r})] \exp[-\beta U_0(\mathbf{r})]. \quad (13)$$

In the context of the pulling experiments $p_0(\xi)$ is the probability density that the molecular end-to-end distance function $\xi(\mathbf{r})$ adopts the value ξ in the unbiased (cantilever-free) ensemble. The

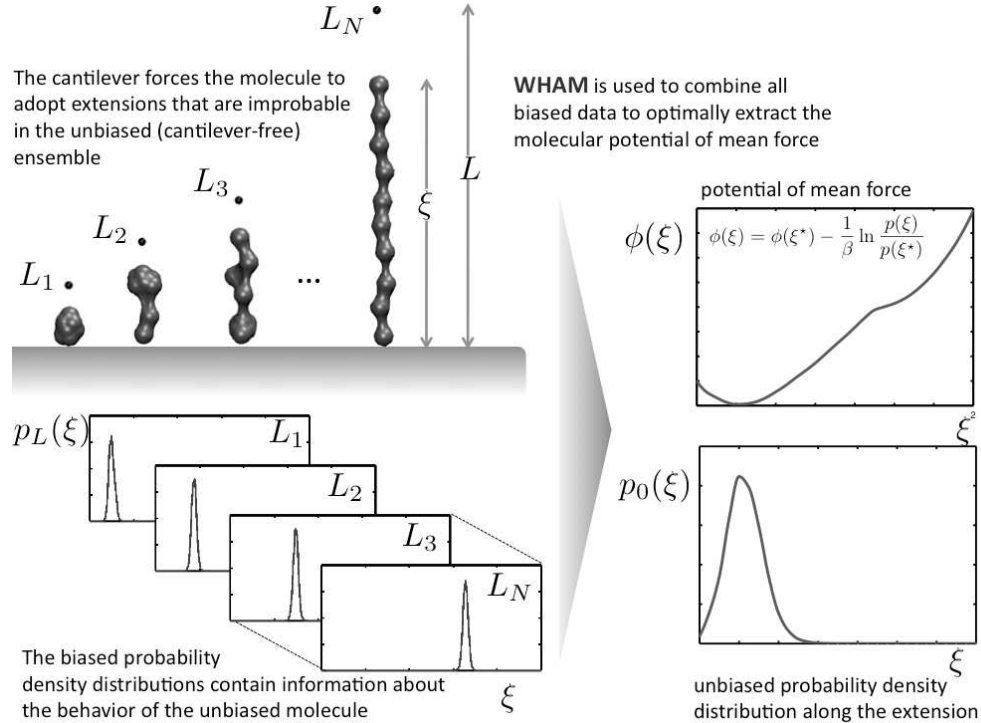


FIG. 6: Scheme of the free energy reconstruction from single-molecule pulling. The potential of mean force $\phi(\xi)$ along the extension coordinate ξ is determined by the probability density distribution $p_0(\xi)$ [Eq. (10)]. Sampling $p_0(\xi)$ directly is challenging because there are regions along ξ that have vanishingly small probability of being observed. In a single-molecule pulling experiments the molecule is forced to adopt those improbable extensions ξ by subjecting it to the bias due to the cantilever. For each L visited (L_1, L_2, \dots, L_N) during pulling the experiments measure the probability density distribution along ξ in the presence of the cantilever bias, $p_L(\xi)$. The $p_L(\xi)$ contain information about the unbiased probability density distribution $p_0(\xi)$ [Eq. (16)]. The WHAM combines all sampled $p_L(\xi)$ to properly remove the cantilever bias and extract $p_0(\xi)$ with minimal variance in the estimate.

quantity $p_0(\xi)$ determines the PMF up to a constant and can be estimated from the force measurements, as we now describe.

At this point it is convenient to discretize the pulling process and suppose that M extensions $L_1, \dots, L_i, \dots, L_M$ are visited during pulling. The potential of the system plus cantilever at extension L_i is $U_i = U_0 + V_i$, where $V_i = \frac{k}{2}[\xi(\mathbf{r}) - L_i]^2$ is the potential bias due to the cantilever. In the i -th biased measurement knowledge of the force F_i and length L_i gives the molecular end-to-end distance. The probability density of observing the value ξ at this given extension is given by

$$p_i(\xi) = \frac{\int d\mathbf{r} \delta[\xi - \xi(\mathbf{r})] \exp[-\beta U_i(\mathbf{r})]}{\int d\mathbf{r} \exp[-\beta U_i(\mathbf{r})]} = \frac{Z_i(\xi)}{Z_i} = \langle \delta[\xi - \xi(\mathbf{r})] \rangle_i, \quad (14)$$

where $Z_i = \int d\mathbf{r} \exp[-\beta U_i(\mathbf{r})]$ is the configurational partition function for the system plus can-

tilever at the i -th extension, and

$$Z_i(\xi) = \int d\mathbf{r} \delta[\xi - \xi(\mathbf{r})] \exp[-\beta U_i(\mathbf{r})]. \quad (15)$$

In principle, the unbiased probability density $p_0(\xi)$ can be reconstructed from each $p_i(\xi)$ since, by virtue of Eqs. (11) and (14), these two quantities are related by

$$p_0(\xi) = \exp[+\beta V_i(\xi)] \frac{Z_i}{Z_0} p_i(\xi), \quad (16)$$

where we have exploited the fact that $Z_0(\xi) = \exp[+\beta V_i(\xi)] Z_i(\xi)$. In practice this approach will not work because the range of ξ values where $p_0(\xi)$ and $p_i(\xi)$ differ significantly from zero need to overlap [32]. Nevertheless, the experiments do provide measurements at a wealth of values of L_i that can be combined to estimate $p_0(\xi)$:

$$p_0(\xi) = \sum_{i=1}^M w_i \exp[+\beta V_i(\xi)] \frac{Z_i}{Z_0} p_i(\xi), \quad (17)$$

where the w_i are some normalized ($\sum_i w_i = 1$) set of weights. In the limit of perfect sampling any set of weights should yield the same $p_0(\xi)$. In practice, for finite sampling it is convenient to employ a set that minimizes the variance in the $p_0(\xi)$ estimate from the series of independent estimates of biased distributions. Such a set of weights is precisely provided by the WHAM prescription [29, 30, 32]

$$w_i = \frac{\exp[-\beta V_i(\xi)] \frac{Z_0}{Z_i g_i}}{\sum_{i=1}^M \exp[-\beta V_i(\xi)] \frac{Z_0}{Z_i g_i}}, \quad (18)$$

where $g_i = 1 + 2\tau_i$ is the statistical inefficiency [34] and τ_i the integrated autocorrelation time [29, 34]. The difficulty in estimating the τ_i for each measurement generally leads to further supposing that the g_i 's are approximately constant and factor out of Eq. (18), so that

$$p_0(\xi) = \frac{\sum_{i=1}^M p_i(\xi)}{\sum_{i=1}^M \exp[-\beta V_i(\xi)] Z_0 / Z_i}. \quad (19)$$

Neglecting the g_i 's from Eq. (19) does not imply that the resulting estimate of $p_0(\xi)$ is incorrect; but simply that the weights selected do not precisely minimize the variance in the estimate. Experience with this method indicates that if the g_i 's do not differ by more than an order of magnitude their effect on $\phi(\xi)$ is small [30, 35].

The computation of the PMF then proceeds as follows. Suppose that N_i force measurements are done for each extension i . Knowledge of the force F_i^j ($j = 1, \dots, N_i$) and of L_i gives the molecular

end-to-end distance ξ_i^j in each of these measurements. The numerator in Eq. (19) is then estimated by constructing a histogram with all the available data,

$$\sum_{i=1}^M p_i(\xi) \approx \sum_{i=1}^M \sum_{j=1}^{N_i} \frac{C_i^j(\xi)}{N_i \Delta\xi}, \quad (20)$$

where $\Delta\xi$ is the bin size, and $C_i^j(\xi) = 1$ if $\xi_i^j \in [\xi - \Delta\xi/2, \xi + \Delta\xi/2)$ and zero otherwise. Estimating the denominator in Eq. (19) requires knowledge of the Z_i 's, the configurational partition functions of the system plus cantilever at all extensions $\{L_i\}$ considered. There are two ways to obtain these quantities. If reversible force-extension data is available, the most direct one is to employ the Helmholtz free energies for the system plus cantilever obtained through the thermodynamic integration in Eq. (7), as they determine the Z_i 's up to a constant multiplicative factor. Alternatively, it is also possible to determine the ratio of the configurational partition functions between the i -th biased system and its unbiased counterpart using $p_0(\xi)$:

$$\frac{Z_i}{Z_0} = \int d\xi \exp[-\beta V_i(\xi)] p_0(\xi). \quad (21)$$

Equations (19) and (21) can be solved iteratively. Starting from a guess for the Z_i/Z_0 , $p_0(\xi)$ is estimated using Eq. (19) and normalized. The resulting $p_0(\xi)$ is then used to obtain a new set of Z_i/Z_0 through Eq. (21), and the process is repeated until self-consistency. This latter approach is the usual procedure to solve the WHAM equations. For computational implementations of the WHAM methodology see Refs. [35–38].

C. Estimating force-extension behavior from $\phi(\xi)$

The PMF is central to the interpretation of the phenomena observed upon pulling and in the estimation of the force-extension behavior when pulling with cantilevers of arbitrary stiffness [28]. Its utility relies on the fact that the configurational partition function of the molecule plus cantilever at extension L [Eq. (1)] can be expressed as

$$Z(L) = \int d\xi \exp\{-\beta[\phi(\xi) + V_L(\xi)]\}, \quad (22)$$

where we have neglected the constant and L -independent multiplicative factor that arises in the transformation since it is irrelevant for the present purposes. The above relation implies that the extension process can be viewed as thermal motion along a one-dimensional effective potential determined by the PMF $\phi(\xi)$ and the bias due to the cantilever $V_L(\xi)$:

$$U_L(\xi) = \phi(\xi) + V_L(\xi). \quad (23)$$

Further, by means of the PMF it is possible to estimate the force vs. extension characteristics for the composite system for any value of the force constant k . This is because the average force exerted on the system at extension L can be expressed as:

$$\langle F \rangle_L = -\frac{1}{\beta} \frac{\partial \ln[Z(L)/Z_0]}{\partial L} = \frac{\int d\xi \frac{\partial V_L(\xi)}{\partial L} \exp\{-\beta[\phi(\xi) + V_L(\xi)]\}}{\int d\xi \exp\{-\beta[\phi(\xi) + V_L(\xi)]\}}, \quad (24)$$

where, for convenience, we have introduced Z_0 in the logarithm just to make the argument dimensionless. Since $\phi(\xi)$ is a property of the isolated molecule it is independent of k . Hence, Eq. (24) can be employed to estimate the F - L isotherms for arbitrary k . Another quantity of interest is the probability density in the force measurements as a function of the extension L , given by

$$p_L(F) = \langle \delta[F - F(\xi)] \rangle_L = \frac{\int d\xi \delta[F - F(\xi)] \exp\{-\beta[\phi(\xi) + V_L(\xi)]\}}{\int d\xi \exp\{-\beta[\phi(\xi) + V_L(\xi)]\}}, \quad (25)$$

as this quantity highlights the fluctuations in the force measurements during pulling.

IV. MODELING SINGLE-MOLECULE PULLING USING MOLECULAR DYNAMICS

A. Basic computational setup

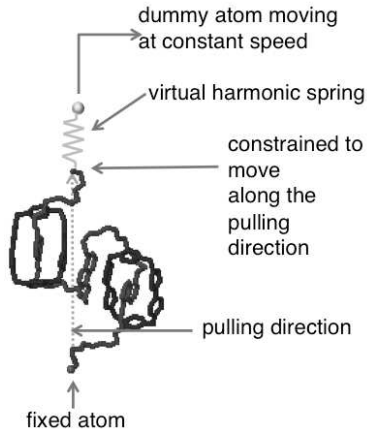


FIG. 7: Setup of a MD pulling simulation. In it, one end of the molecule is fixed while the opposite end is attached to a dummy atom via a virtual harmonic spring. The position of the dummy atom is the simulation analogue of the cantilever position and is controlled throughout. The pulling is performed by moving the dummy atom away from the molecule at a constant speed. The fluctuating deflection of the harmonic spring measures the force exerted during the pulling and is the simulation analog of the cantilever deflection in the AFM pulling experiment.

Computationally, the forced unfolding of single molecules can be studied using “pulling” or “steered” molecular dynamics (MD) simulations [39–41]. The general setup of this type of simu-

lations is schematically described in Fig. 7. The stretching computation begins by attaching one end of the molecule to a rigid isotropic harmonic potential that mimics the molecular attachment to the surface. Simultaneously, the opposite molecular end is connected to a dummy atom via a virtual harmonic spring. The position of the dummy atom is the simulation analogue of the cantilever position L , and is controlled throughout. In turn, the varying deflection of the virtual harmonic spring measures the force exerted during the pulling. The stretching is caused by moving the dummy atom away from the molecule at a constant speed. The pulling direction is defined by the vector connecting the two terminal atoms of the complex. Since cantilever potentials are typically stiff in the direction perpendicular to the pulling, in the simulations the terminal atom that is being pulled is typically forced to move along the pulling direction by introducing appropriate additional restraining potentials.

In the simulations, the molecule plus any surrounding solvent are assumed to be weakly coupled to a heat bath and the effect of the heat bath is modeled through a thermostat. Thermostats that are useful in simulating the force spectroscopy include the Nosé-Hoover chain [32, 42–44] and the Langevin thermostat [45]. In turn, velocity rescaling thermostats, like the Berendsen thermostat [46], can be problematic because they can lead to a spurious violation of the second law during the pulling because the ensemble generated by them does not satisfy the equipartition theorem [47]. The thermal dynamics of the molecule, cantilever and any surrounding solvent are then followed in a NVT or NPT ensemble using MD. Freely available MD codes such as NAMD [48] and GROMACS [49] have pulling routines implemented in them. We have, in addition, developed [28] a pulling routine for TINKER [50] and created a graphical interface for it called MOLpull [36]. MOLpull is accessible through the NanoHub (<http://nanohub.org/>) and can be used to perform the pulling and reconstruct the PMF along the extension coordinate of arbitrary molecules.

B. Modeling strategies

One of the challenges in simulating pulling experiments using MD is to bridge the large disparity between the pulling speeds v that are employed experimentally and those that can be accessed computationally. While typical experiments often use $v \sim 1 - 10^{-3} \mu\text{m/s}$, current computational capabilities require pulling speeds that are several (~ 5 to 8) orders of magnitude faster.

One possible strategy that can be used to overcome this difficulty is to strive for pulling speeds that are slow enough that reversible behavior is recovered. Under such conditions the results

become independent of v , and the simulations comparable to experimental findings. The way to determine if the pulling is done under quasistatic conditions is by making sure that the force-extension curves during extension and subsequent contraction essentially coincide, such that the net work exerted during a pulling/retraction cycle is negligible. The advantage of this approach is that there is a systematic way to test if quasistatic behavior is recovered by estimating the amount of work dissipated in a pulling/retraction cycle. Its limitation is that the pulling speeds that are required to obtain reversible behavior may be computationally unfeasible for pulling large macromolecules. In fact, the examples that exist in the literature where reversible behavior has been computationally recovered are for relatively small molecular systems [28, 51, 52] using pulling speeds between $10^{-4} - 10^{-7}$ nm/ps. When pulling large macromolecules, present computational limitations require using pulling speeds of $\sim 10^{-3}$ nm/ps [53] which is far too fast for recovering reversible behavior even in simple systems.

A second possible strategy is to pull under irreversible conditions. A single nonequilibrium trajectory can recover some qualitative structural features encountered during the pulling [41, 53], but it is not sufficient to reconstruct the molecular potential of mean force. In order to reconstruct the PMF it is necessary to pull many times and use a nonequilibrium work fluctuation theorem [23–25, 54] to reconstruct the free energy profile [22, 55–61]. The advantage of this strategy is that it can be used when equilibrium data is difficult to obtain, for example during the forced unfolding of a complex protein like the one in Fig. 5. Note, however, that this strategy requires adequately estimating the average of an exponential quantity (Eq. (8)) that often has poor convergence properties [60, 62]. For a computational implementation of some nonequilibrium methods to reconstruct the PMF see [38].

As a third *indirect* alternative it is possible to reconstruct the molecular potential of mean force without pulling continuously but rather by sampling the fluctuating force at selected points along the extension coordinate. For each L the system is allowed to thermally equilibrate and the force is measured for a given amount of simulation time. The data obtained is used to reconstruct the PMF using WHAM as described in Sec. III, and then the PMF is employed to estimate the force-extension behavior of the molecule plus cantilever. While the free energy reconstruction using continuous equilibrium force-extension data uses a data set with a large (essentially continuous) set of L values with very few data points for the force at each L , this strategy works in a different limit where one has few extensions L but extensive sampling of the force at each extension. The advantage of this procedure is that it can be used to reconstruct the PMF of molecules that are computationally challenging to pull under equilibrium or near equilibrium conditions. The limitation is that there is

no entirely satisfactory way to judge when the system has achieved a state of thermal equilibrium in the MD trajectory. This contrasts with the direct strategy involving explicit pulling in which it is easy to judge when quasistatic behavior has been recovered by measuring the amount of dissipative work in extension/contraction cycles. As in the direct equilibrium pulling, the sampling for each L has to be sufficient to capture all relevant molecular events that are consistent with the extension.

Another important aspect that has to be taken into consideration when modeling single-molecule pulling is the cantilever stiffness employed. Experimentally accessible AFM cantilevers stiffnesses are in the $10^{-1} - 10^4$ pN/Å range [13]. While experiments tend to favor soft cantilever potentials because they provide high resolution in the force measurements, in the simulations stiff cantilevers are preferable. This is because when employing stiffer cantilever potentials: (i) a shorter range of L values is required to explore all the molecular extension coordinates thus reducing the simulation time; (ii) the possible values of ξ that are accessible for the molecule for a given L are further restricted making it faster to sample all accessible conformational space. As a consequence of (ii), the amount of dissipative work in a pulling/retraction cycle decreases with increasing cantilever stiffness for fixed pulling speed [63]. In our experience with molecular pulling, a rule of thumb that we frequently use is that a cantilever stiffness of ~ 1 pN/Å can be regarded as soft and one with stiffness of $\sim 10^2$ pN/Å as stiff. These numbers, however, will clearly depend on the particular system that is being pulled.

C. Examples

1. Direct modeling: Oligorotaxanes

As an example of an explicit simulation of a single-molecule pulling experiment consider the forced unfolding of a [3]rotaxane [28, 64, 65]. Figure 8 shows the force-extension curves during the extension (blue) and subsequent contraction (red) for two different pulling speeds and a schematic of the structure of the molecule that is being pulled. The oligorotaxane[71] consists of two cyclobis(paraquat-*p*-phenylene) tetracationic cyclophanes (the blue boxes) threaded onto a linear chain (in red) composed of three naphthalene units linked by oligoethers with oligoether caps at each end. As the system is stretched the oligorotaxane undergoes a conformational transition from a folded globular state to an extended coil. In the process, the force initially increases approximately linearly with L , then drops, and subsequently increases again. For the faster pulling speed hysteresis in the force-extension cycle is evident while for $v = 10^{-4}$ nm/ps the F - L curves obtained

during extension and contraction essentially coincide and, for all practical purposes, the pulling occurs under equilibrium conditions. Note that thermal fluctuations in the force measurements are roughly comparable to the average values indicating that the system is far from the thermodynamic limit.

The force-extension behavior predicted by the MD simulations for the oligorotaxane is qualitatively similar to that experimentally observed for the RNA hairpin (recall Fig. 2). Specifically, there is a region of mechanical instability where the force drops with increasing extension ($\frac{\partial(F)L}{\partial L} < 0$). Further, if one fixes L around the region of mechanical instability ($L = 70.0 \text{ \AA}$ in this case) and allows the system to evolve, the molecule undergoes conformational transitions between a folded globular state and an extended coil, see Fig. 9. The right panel in Fig. 9 shows the probability density of the distribution of ξ values obtained from a 20 ns trajectory. The system exhibits a clear dynamical bistability along the ξ coordinate. As in the RNA case, the bistability along the end-to-end distance leads to blinking in the force measurements from a high force to a low force regime during the pulling.

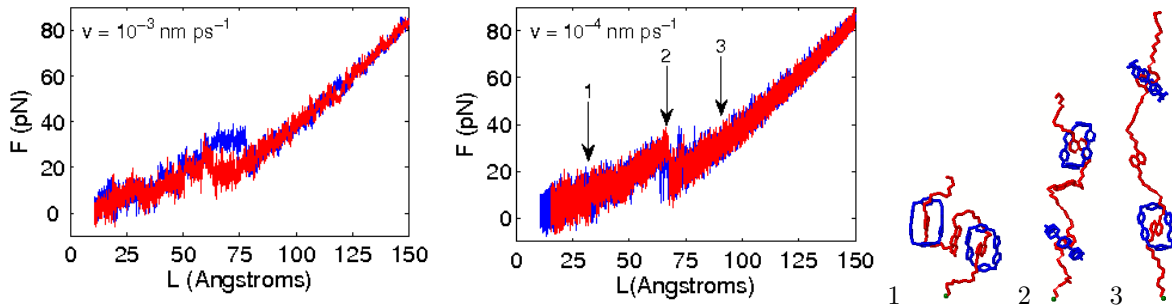


FIG. 8: Simulated force-extension behavior of a [3]rotaxane immersed in a high dielectric medium at 300K obtained using MD and the MM3 [66] force field. The plots show the force-extension curves during stretching (blue) and retraction (red) using a pulling speed v of 10^{-3} and 10^{-4} nm/ps. The cantilever employed has a soft spring constant of $k = 1.1 \text{ pN/\AA}$. Snapshots of structures (labels 1, 2 and 3) encountered during pulling are shown in the far right.

The molecular PMF of the [3]rotaxane reconstructed from the force-extension data is shown in Fig. 10. The thermodynamic native state of the oligorotaxane, i.e. the minimum in $\phi(\xi)$, corresponds to a globular folded structure with an average end-to-end distance of 7 \AA . The PMF consists of a convex region (a region of positive curvature) for $\xi < 44 \text{ \AA}$ when the molecule is folded, another convex region for $\xi > 52 \text{ \AA}$ when the molecule is extended and a region of concavity (where $\frac{\partial^2 \phi}{\partial \xi^2} < 0$) for $\xi = 44 - 52 \text{ \AA}$ where the conformational transition occurs. The region of concavity along the PMF arises because the molecule is inherently bistable along some natural

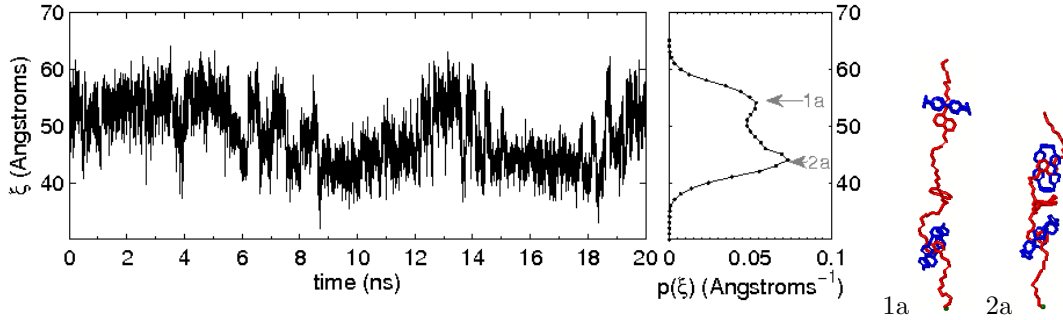


FIG. 9: Dynamical bistability when L is fixed around the region of mechanical instability. The figure shows the time dependence and probability density distribution of the end-to-end molecular extension ξ when the [3]rotaxane plus cantilever is constrained to reside in the unstable region of Fig. 8, with $L = 70.0$ Å. Typical structures encountered in this regime (labels 1a-2a) are shown in the right.

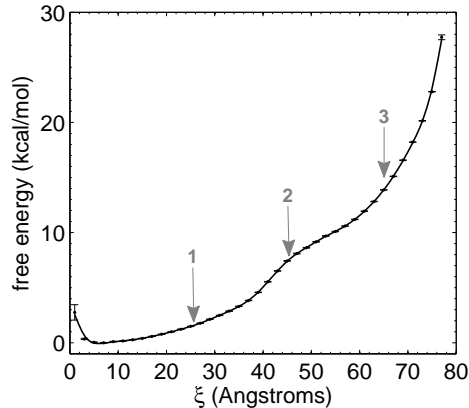


FIG. 10: Molecular potential of mean force for the [3]rotaxane along the end-to-end distance ξ . The solid lines result from a spline interpolation of the available data points. Typical structures (labels 1-3) are shown in Fig. 8.

unfolding pathway. As discussed in Sec. V, this characteristic structure of the PMF is responsible for many of the interesting features observed during pulling.

2. Indirect modeling: DNA dimer

As an example of indirect modeling of the force spectroscopy, we now consider the forced extension of the DNA dimer shown in Fig. 11 along the O5'-O3' coordinate in explicit water [67]. The dimer is composed of two guanine-cytosine base pairs capped by a *trans*-azobenzene linker. Computationally, it is described by the CHARMM27 force field and the total system is composed of the 178 DNA hairpin atoms, 4 sodium ions, and 1731 TIP3P water molecules in a $27.1\text{Å} \times 34.7\text{Å} \times 56.3\text{Å}$ box

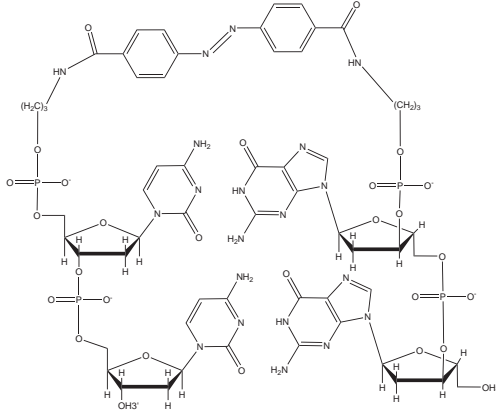


FIG. 11: DNA dimer consisting of two guanine-cytosine base pairs connected by a *trans*-azobenzene linker.

in the NVT ensemble at 300K (see Ref. [67] for details). Even for this relatively simple system it is challenging to computationally obtain equilibrium force-extension data by pulling directly because the pulling has to be slower than any characteristic timescale of the system. An indirect approach is thus more convenient.

The sampling along the extension coordinate required to reconstruct the potential of mean force proceeded as follows: The distance, L , between the surface and the cantilever was fixed at several different values along the extension coordinate. The system was first allowed to equilibrate for 4 ns. Subsequently, the dynamics was followed for 8 ns and the end-to-end distance ξ recorded every 1 ps. The cantilever was taken to have a stiffness of $k_0 = 110 \text{ pN}/\text{\AA}$. In total, 155 extensions were simulated. Values of L ranged from 6.75 \AA to 44.75 \AA while ξ ranged from 2.9 \AA to 44.9 \AA . Total analyzed simulation time was $1.24 \mu\text{s}$.

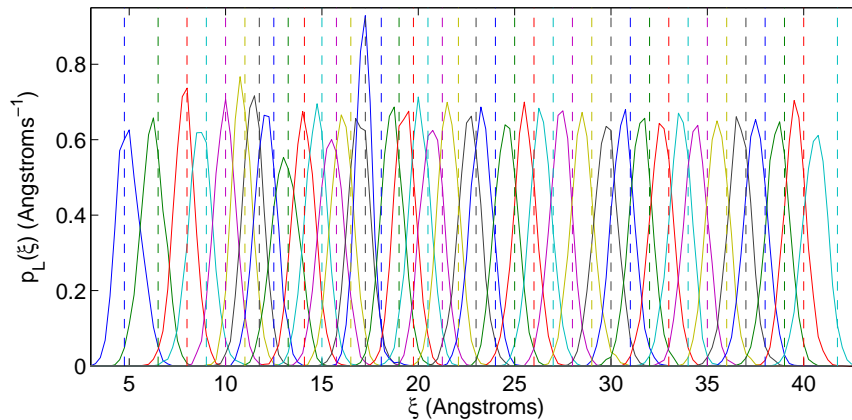


FIG. 12: Probability density distribution of ξ when the cantilever bias is fixed at position L for the DNA dimer. Each curve corresponds to a different L . The dotted lines mark the L chosen in each case.

Figure 12 shows some of the sampled $p_L(\xi)$ [defined in Eq. (14)] that are used as input for the PMF reconstruction. The resulting Helmholtz free energy profile is shown in Fig. 13. It consists of a convex region for $\xi < 17 \text{ \AA}$ representing the folded state, a mostly concave region for $17 \text{ \AA} < \xi < 40 \text{ \AA}$ where the molecule unfolds, followed by another region of convexity that corresponds to the unfolded state. The unfolding region exhibits some convex intervals signaling marginally stable intermediates encountered during the unfolding.

Using the PMF one can then estimate the force-extension characteristics of the system when pulling with cantilevers of arbitrary stiffness as described in Sec. III C. Figure 14 shows the predicted average and probability density in the force measurements when the DNA dimer is pulled using a cantilever stiffness of 1.1 pN/\AA . Note the degree of detail of the pulling process revealed by the probability density.

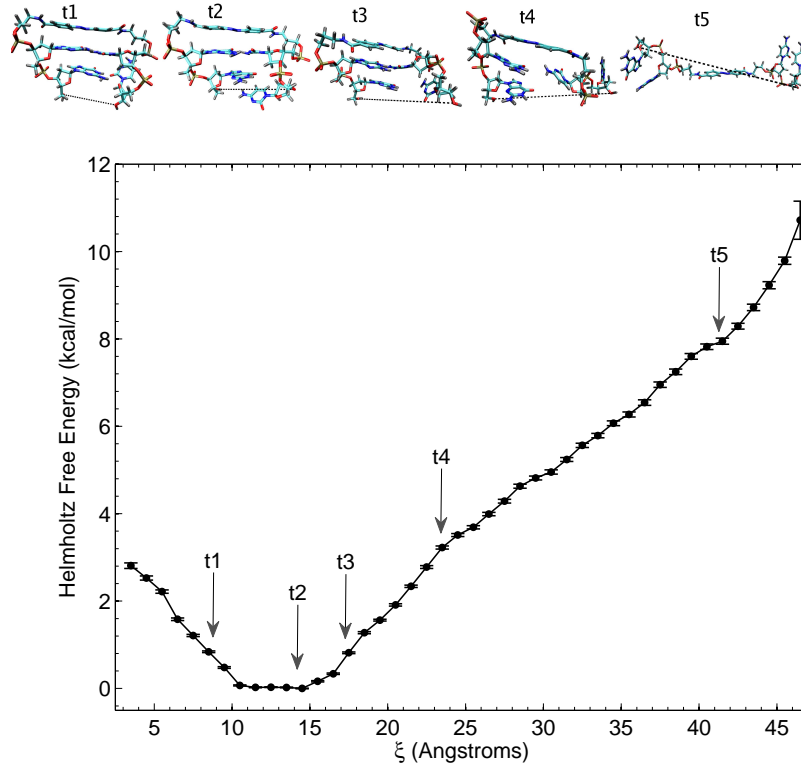


FIG. 13: Potential of mean force along the end-to-end distance coordinate for the azobenzene capped DNA hairpin. Snapshots of structures adopted during the unfolding are shown in the top panel. The error bars correspond to twice the standard deviation obtained from a bootstrapping analysis.

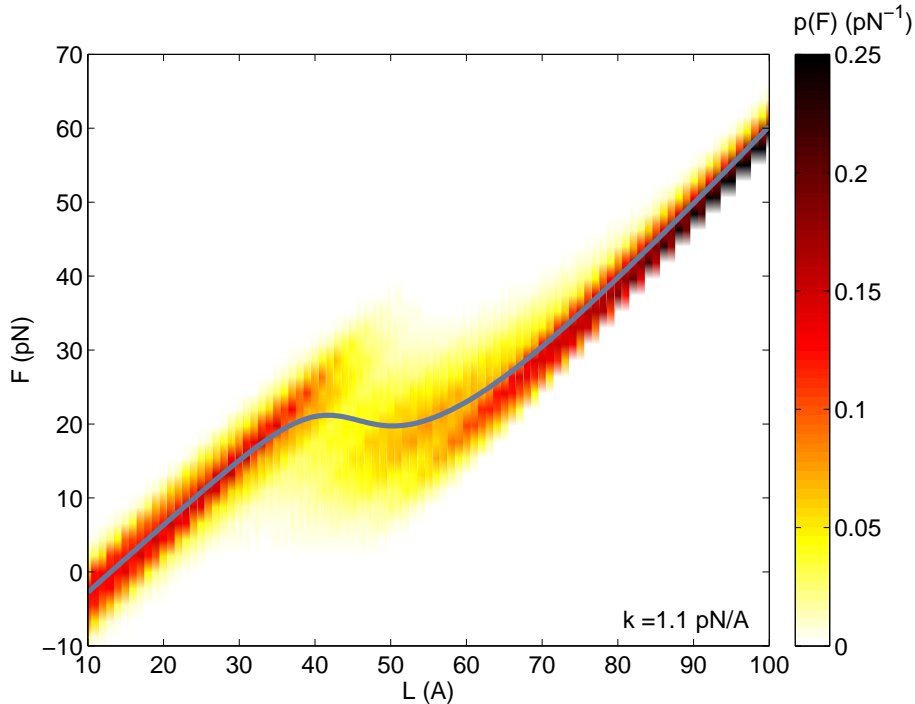


FIG. 14: Force-extension profile of the DNA dimer when pulled with a cantilever of stiffness $1.1 \text{ pN}/\text{\AA}$ estimated from the PMF in Fig. 13. The solid line represents the average force-extension profile, the color plot the probability density distribution of the force measurements $p_L(F)$ [Eq. (25)].

V. INTERPRETATION OF PULLING PHENOMENOLOGY

A striking feature of single-molecule pulling is that the elastic behaviors of vastly different molecules often have qualitatively similar features. In Secs. II and IV we discussed two of these features: (i) mechanically unstable regions where $\frac{\partial(F)L}{\partial L} < 0$ and (ii) regions of dynamical bistability where blinking in the force measurements for fixed L is observed. In this section, we discuss the origin of both of these effects. Emphasis will be placed on the basic requirements on the molecule and the pulling device for the emergence of the phenomena.

A. Basic structure of the molecular potential of mean force

Recall the basic structure of the PMF for the oligorotaxanes (Fig. 10): It consists of a convex region ($\frac{\partial^2 \phi(\xi)}{\partial \xi^2} > 0$) where the molecule is folded, another region of convexity when the molecule is extended and a region of concavity ($\frac{\partial^2 \phi(\xi)}{\partial \xi^2} < 0$) where the molecule unfolds. This basic structure in the PMF is not unique to the rotaxanes but it is actually a common feature of the extension behavior of single molecules of sufficient structural complexity. Regions of convexity mark me-

mechanically stable conformations while regions of concavity mark molecular unfolding events. We had already encountered this very same basic structure when studying the elastic properties of the DNA dimer shown in Fig. 13. Perhaps a more spectacular example of this structure is provided by the polyprotein composed of eight repeats of the Ig27 domain of human titin that we introduced in Fig. 5. The PMF reconstructed from experimental force-extension data is shown in Fig. 15. In this case also, mechanically stable conformations are described by regions of convexity in the PMF, while regions of concavity mark unfolding events. The reason why in this case there are 8 regions of concavity along the extension pathway is that there are 8 unfolding events, each marking the unfolding of one of the Ig27 domains in the polyprotein. Below we discuss the implications of these regions of concavity along the PMF.

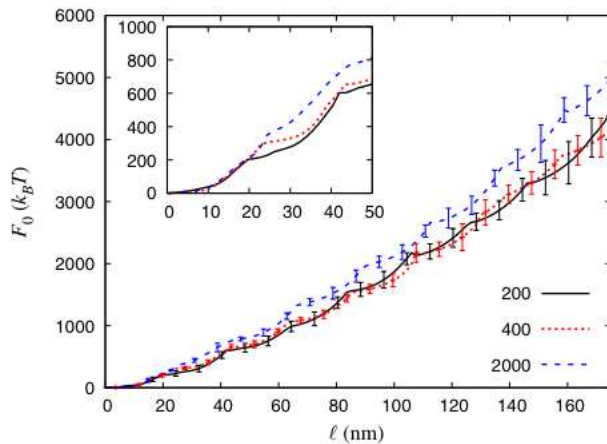


FIG. 15: Molecular potential of mean force for a polyprotein composed of eight repeats of the Ig27 domain of human titin. The free energy profile was reconstructed using a nonequilibrium method and pulling at different pulling speeds of 200 (black curve), 400 (red dotted curve) and 2000 (blue dashed curve) nm/s. The number of nonequilibrium unfolding trajectories used in the estimate was 66, 35 and 29, respectively. The force-extension profile of this molecule is shown in Fig. 5. Figure taken from [22]

B. Mechanical instability

A first consequence of the regions of concavity along the molecular PMF is that they lead to mechanically unstable regions in the F - L isotherms. To see this consider the configurational partition function of the molecule plus cantilever at extension L [Eq. (22)] for large k . In this regime, most contributions to the integral will come from the region where $\xi = L$. Consequently,

$\exp[-\beta\phi(\xi)]$ can be expanded around this point to give:

$$\exp[-\beta\phi(\xi)] = \exp[-\beta\phi(L)] \left[1 - \beta A_1(L)(\xi - L) - \frac{\beta}{2} A_2(L)(\xi - L)^2 + \dots \right], \quad (26)$$

where $A_1(L) = \frac{\partial\phi(L)}{\partial L}$ and $A_2(L) = \frac{\partial^2\phi(L)}{\partial L^2} - \beta \left(\frac{\partial\phi(L)}{\partial L} \right)^2$. Here, the notation is such that $\frac{\partial\phi(L)}{\partial L} = \frac{\partial\phi(\xi)}{\partial\xi} \Big|_{\xi=L}$. Introducing Eq. (26) into Eq. (22), integrating explicitly the different terms, and performing an expansion around $1/k = 0$ one obtains:

$$Z(L) = \sqrt{\frac{2\pi}{k\beta}} \exp[-\beta\phi(L)] \left[1 - \frac{1}{2k} A_2(L) + \mathcal{O}(1/k^2) \right]. \quad (27)$$

Using Eq. (24), the derivative of the force can be obtained from this approximation to the partition function. To zeroth order in $1/k$ it is given by:

$$\frac{\partial\langle F \rangle_L}{\partial L} = \frac{\partial^2\phi(L)}{\partial L^2} + \mathcal{O}(1/k). \quad (28)$$

That is, an unstable region in the force vs. extension where $\frac{\partial\langle F \rangle_L}{\partial L} < 0$ requires a region of concavity in the PMF.

Note, however, that since the properties that are measured during pulling are those of the molecule plus cantilever, the observed mechanical stability properties will depend on the cantilever stiffness employed. In fact, when employing very soft cantilevers no mechanically unstable region can be observed independent of the molecule that is being pulled, provided that no bond-breaking occurs during the pulling. To see this, consider the soft-spring approximation of the configurational partition function of the system plus cantilever [Eq. (22)]:

$$\frac{Z(L)}{Z_0} \approx \frac{\exp(-\beta k L^2 / 2)}{Z_0} \int d\xi \exp[-\beta\phi(\xi)] \exp(\beta k L \xi) = \exp(-\beta k L^2 / 2) \langle \exp(\beta k L \xi) \rangle, \quad (29)$$

where the notation $\langle f \rangle$ stands for the unbiased (cantilever-free) average of f . In writing Eq. (29) we have supposed that in the region of relevant ξ (in which the integrand is non-negligible) $k\xi^2/V_L(\xi) \ll 1$ and, hence, that the cantilever potential is well approximated by $V_L(\xi) \approx \frac{kL^2}{2} - kL\xi$. This approximation is valid provided that that no bond breaking is induced during pulling and permits the introduction of a cumulant expansion [68] in the configurational partition function. Specifically, the average $\langle \exp(\beta k L \xi) \rangle$ can be expressed as:

$$\langle \exp(\beta k L \xi) \rangle = \sum_{n=0}^{\infty} \frac{(\beta k L)^n}{n!} \langle \xi^n \rangle = \exp \left\{ \sum_{n=1}^{\infty} \frac{(\beta k L)^n}{n!} \kappa_n(\xi) \right\} \quad (30)$$

where $\kappa_n(\xi)$ is the n -th order cumulant. In view of Eqs. (24), (29) and (30) the slope of the F - L curves can be expressed as:

$$\frac{\partial\langle F \rangle_L}{\partial L} = k - \beta k^2 \sum_{n=0}^{\infty} \frac{(\beta k L)^n}{n!} \kappa_{n+2}(\xi). \quad (31)$$

It then follows that to lowest order in k ,

$$\frac{\partial \langle F \rangle_L}{\partial L} \approx k > 0. \quad (32)$$

That is, no unstable region in the F - L curve can arise in the soft-spring limit, irrespective of the specific form of $\phi(\xi)$.

Figure 16 illustrates the dependence of the region of mechanical instability on the cantilever stiffness using the [3]rotaxane case discussed in Sec. IV C 1 as an example. The figure shows the k -dependence of the local maximum F^+ and minimum F^- in the average force measurements that enclose the region of mechanical instability. The critical forces F^+ and F^- show a smooth and strong dependence on the cantilever spring constant. For large k the system persistently shows a region of instability in the isotherms. However, as the cantilever spring is made softer, F^+ and F^- approach each other and for small k the mechanical instability in the F - L isotherms is no longer present.

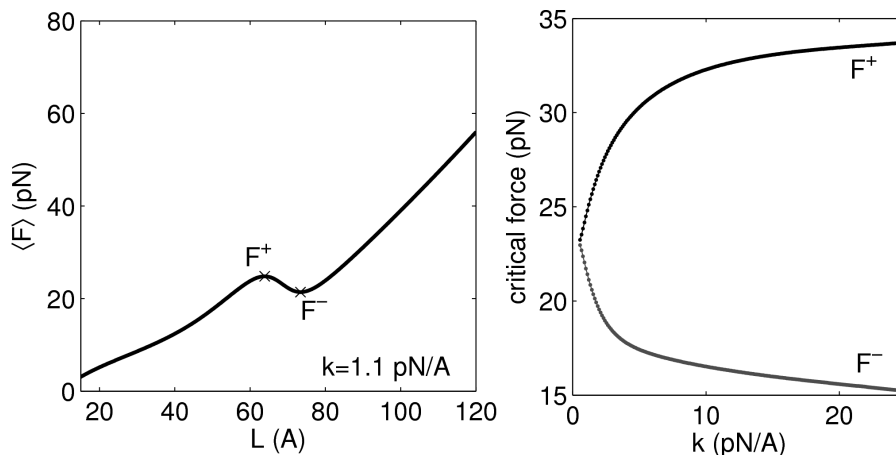


FIG. 16: Dependence of the mechanical instability on the cantilever stiffness. Left panel: force-extension isotherm for a [3]rotaxane when pulled with a cantilever of stiffness $k = 1.1$ pN/Å. Here, F^+ and F^- correspond to the values of the force when the F - L curves exhibit a maximum and minimum, and enclose the region of mechanical instability. Right panel: dependence of F^+ and F^- on k . Note the decay of the region of mechanical instability for soft cantilevers.

Another interesting aspect of single-molecule pulling is that there is a close relationship between the mechanical stability properties and the fluctuations in the force measurements. To see this, note that the average slope of the F - L curves can be expressed in terms of the fluctuations in the force:

$$\frac{\partial \langle F \rangle_L}{\partial L} = \left\langle \frac{\partial^2 V_L(\xi)}{\partial L^2} \right\rangle_L - \beta [\langle F^2 \rangle_L - \langle F \rangle_L^2] = k - \beta [\langle F^2 \rangle_L - \langle F \rangle_L^2], \quad (33)$$

where we have employed Eq. (24). The sign of $\frac{\partial \langle F \rangle_L}{\partial L}$ determines the mechanical stability during the extension and hence Eq. (33) relates the thermal fluctuations in the force measurements with the stability properties of the F - L curves. Specifically, for the stable regions for which $\frac{\partial \langle F \rangle_L}{\partial L} > 0$ the force fluctuations satisfy

$$\langle F^2 \rangle_L - \langle F \rangle_L^2 < \frac{k}{\beta}. \quad (34a)$$

In turn, in the unstable regions the force fluctuations are larger than in the stable regions and satisfy the inequality

$$\langle F^2 \rangle_L - \langle F \rangle_L^2 > \frac{k}{\beta}. \quad (34b)$$

The fact that the fluctuations are larger around the region of mechanical bistability is noticeable in some of the force-extension curves that we have discussed here (see Figs. 8 and 14). Figure 17 illustrates these general observations in the specific case of the pulling of [3]rotaxane. As shown, for $k = k_0$ and $k = 2k_0$ (where $k_0 = 1.1$ pN/Å) there is a region where the force fluctuations become larger than k/β and, consequently, unstable behavior in F - L develops. For $k = 0.4k_0$ or less the fluctuations in the force are never large enough to satisfy Eq. (34b) and no critical points in the average F - L curve develop, as can be confirmed in Fig. 16.

C. Dynamical bistability

The observation of force measurements that blink between a high-force and a low force regime for selected extensions requires the composite molecule plus cantilever system to be bistable along the end-to-end distance for some L . That is, the effective potential $U_L(\xi)$ [Eq. (23)] must have a double minimum. Since the molecular PMF is not usually bistable along ξ the bistability must be introduced by the cantilever potential. To see how this bistability arises, consider the effective potential for the [3]rotaxane for selected L shown in Figure 18. For L in the mechanically stable regions of the force-extension isotherms ($L = 40$ Å and $L = 90$ Å) the effective potential exhibits a single minimum along the ξ coordinate. However, when $L = 70$ Å a bistability in the potential develops. At this extension the cantilever potential turns the region of concavity in $\phi(\xi)$ into a barrier between two minimum. The secondary minimum is the cause for the blinking in the force measurements observed at this L (cf. Fig. 9).

What are the minimum requirements for the emergence of bistability along ξ ? A necessary

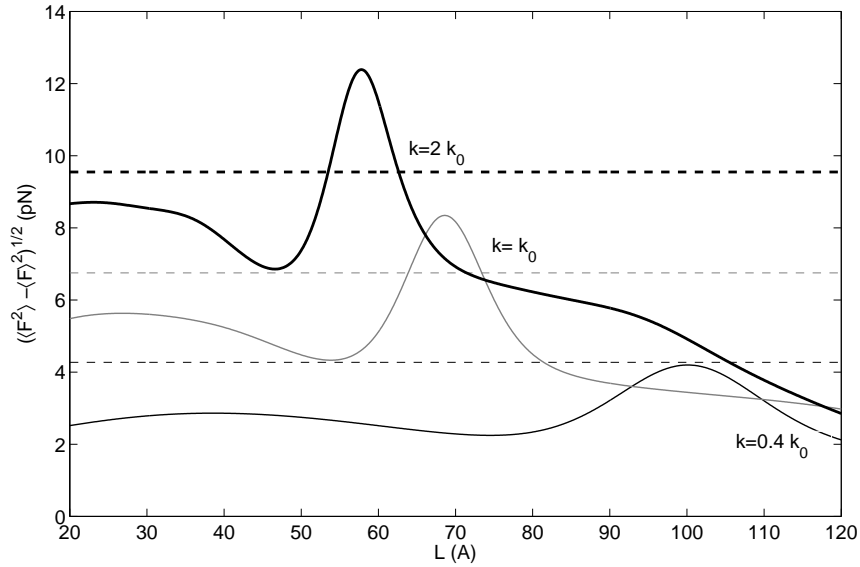


FIG. 17: Standard deviation in the force measurements $\sigma_F = \sqrt{\langle F^2 \rangle_L - \langle F \rangle_L^2}$ as a function of L for three different cantilever spring constants k during the pulling of the [3]rotaxane. In each case, the dotted line indicates the value of $\sqrt{k/\beta}$ which sets the limit between the stable and unstable branches in the extension, see Eq. (34). Here, $k_0 = 1.1$ pN/Å. Figure taken from [28].

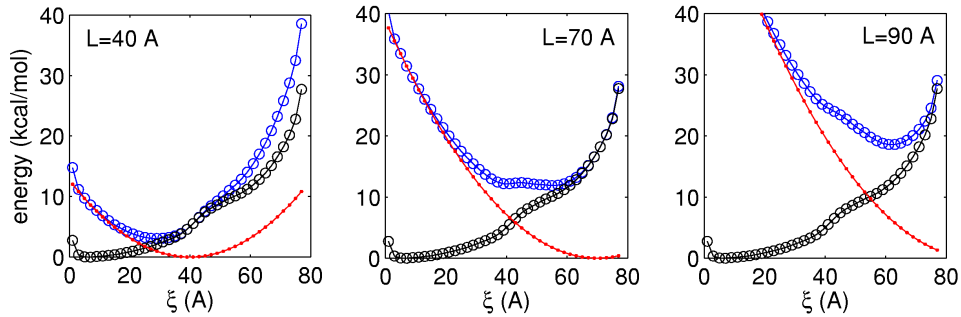


FIG. 18: Effective potential $U_L(\xi) = \phi(\xi) + V_L(\xi)$ for the molecule plus cantilever for different values of the extension L . In the panels, the blue circles correspond to $U_L(\xi)$, the black circles to the PMF $\phi(\xi)$, and the solid line to the cantilever potential $V_L(\xi)$ with $k = 1.1$ pN/Å. Note the bistability in the effective potential for $L = 70$ Å.

condition is that the effective potential $U_L(\xi)$ is concave for some region along ξ , that is:

$$\frac{\partial^2 U_L(\xi)}{\partial \xi^2} = \frac{\partial^2 \phi(\xi)}{\partial \xi^2} + k < 0. \quad (35)$$

For Eq. (35) to be satisfied it is required that both: (i) the PMF of the isolated molecule has a

region of concavity where $\frac{\partial^2 \phi(\xi)}{\partial \xi^2} < 0$, and (ii) the cantilever employed is sufficiently soft such that

$$k < -\min \left(\frac{\partial^2 \phi(\xi)}{\partial \xi^2} \right), \quad (36)$$

for some ξ . Equation (36) imposes an upper bound on $k > 0$ for bistability to be observable. If k is very stiff the inequality would be violated for all ξ and bistability would not be manifest. Note, however, that there is no lower bound for k that prevents bistability along ξ .

Since the region of concavity in the PMF is a common feature of molecules that have stable folded conformations, this bistability for selected extensions is a common feature of the pulling. We have already encountered this phenomenon during the pulling of the oligorotaxane and the RNA hairpin (Fig. 2). The DNA dimer (Fig. 14) and the polyprotein in Fig. 5 are also expected to show this behavior since their PMF has regions of concavity along ξ . The bistability has also been predicted to be measurable when pulling π -stacked molecules [51, 69]. For a discussion of the bistability in the isotensional ensemble, see Ref. [18]. Figure 19 shows a striking recent experimental demonstration of the bistability obtained during the pulling of a leucine zipper [70].

VI. SUMMARY

The power of single-molecule pulling techniques is that they allow for mechanical control over the molecular conformation while simultaneously making thermodynamic measurements of any mechanically induced unfolding events. The force-extension data often exhibits mechanically unstable regions where the force decreases with increasing extension and bistable regions where, for fixed extension, blinks between a high-force and a low-force regime are observed.

From the force-extension isotherms it is possible to reconstruct the PMF along the extension coordinate using the WHAM. In fact, the pulling process can be seen as an experimental realization of the WHAM methodology. The PMF summarizes the changes in the free energy during folding and determines the elastic properties of the molecule. Quite remarkably, the PMF's of widely different molecules often shares the same basic structure. It consists of regions of convexity that represent mechanically stable molecular conformations interspersed by regions of concavity that signal molecular unfolding events.

The starting point in the interpretation of single-molecule pulling experiments is Eq. (22) which shows that the pulling process can be viewed as thermal motion along a one-dimensional potential $U_L(\xi) = \phi(\xi) + V_L(\xi)$ that is determined by the PMF $\phi(\xi)$ and the cantilever potential $V_L(\xi)$. Several important conclusions follow from this simple observation. The first and most obvious one

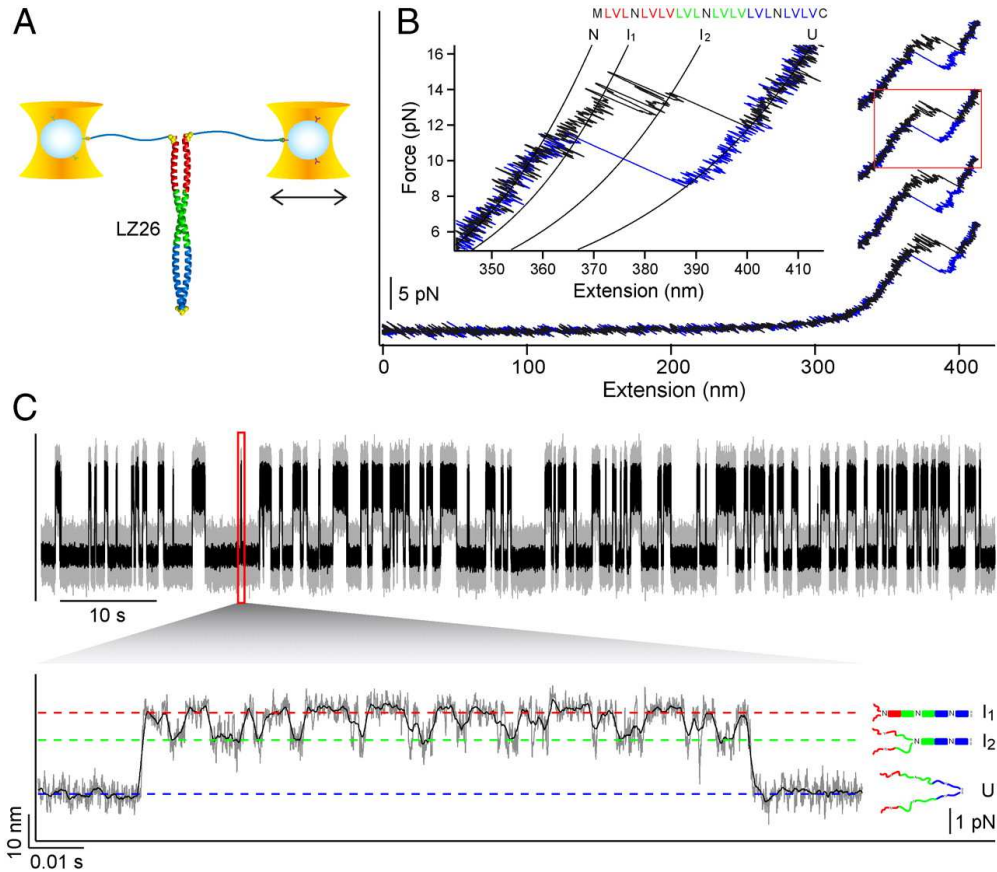


FIG. 19: Bistability during the pulling of a leucine zipper measured in an optical lattice arrangement. (A) Schematic of the experimental setup; (B) Sample force-extension traces; (C) Force vs time record for fixed L in the region of dynamical bistability. Figure taken from [70].

is that the properties that are measured during pulling are those of the molecule plus cantilever and hence that the basic phenomenon observed during pulling depend on the cantilever stiffness used. It also follows that in order for the mechanical instability and dynamical bistability to be observable, the molecular PMF has to have a region of concavity. However, while the dynamical bistability survives for soft cantilever and decays for rigid cantilevers [see Eq. (36)], the opposite is true for the mechanical instability [recall Eq. (32)]. The mechanical stability properties during pulling were also shown to be intimately related to the fluctuations in the force measurements [see Eq. (34)].

One basic challenge in modeling the force spectroscopy using MD is to bridge the several orders of magnitude gap that exists between the pulling speeds used experimentally and those that are computationally feasible. We detailed two possible strategies that can be employed to overcome this difficulty: either strive for pulling speeds that are slow enough such that reversible behavior

is recovered or use an indirect approach in which the PMF is first reconstructed and the force-extension behavior is then estimated from the PMF. Other strategies based on nonequilibrium pulling are also possible. From the simulation of the pulling one can gain insights into the conformations encountered during molecular unfolding and into structure-function relations responsible for the elastic behavior of single molecules.

Currently, the accurate reconstruction of the PMF of molecular systems using equilibrium sampling is a serious computational challenge except for relatively modest systems. The recent discovery of nonequilibrium work fluctuation relations has led to a surge of activity in seeking nonequilibrium methods to reconstruct the PMF that can be used in situations where proper equilibrium sampling is unfeasible. The full consequences of this venue of research is still to be determined.

Acknowledgments

This work was supported by the Non-equilibrium Energy Research Center (NERC) which is an Energy Frontier Research Center funded by the U.S. Department of Energy, Office of Science, Office of Basic Energy Sciences under Award Number de-sc0000989. The authors thank Dr. Martin McCullagh for useful discussions.

-
- [1] A. Ashkin, Proc. Natl. Acad. Sci. U. S. A. **94**, 4853 (1997).
 - [2] M. Rief, M. Gautel, F. Oesterhelt, J. M. Fernandez, and H. E. Gaub, Science **276**, 1109 (1997).
 - [3] S. Weiss, Science **283**, 1676 (1999).
 - [4] W. E. Moerner and M. Orrit, Science **283**, 1670 (1999).
 - [5] T. R. Strick, V. Croquette, and D. Bensimon, Nature **404**, 901 (2000).
 - [6] A. Nitzan and M. A. Ratner, Science **300**, 1384 (2003).
 - [7] F. Ritort, J. Phys.: Condens. Matter **18**, R531 (2006).
 - [8] J. P. Camden, J. A. Dieringer, Y. Wang, D. J. Masiello, L. D. Marks, G. C. Schatz, and R. P. Van Duyne, J. Am. Chem. Soc. **130**, 12616 (2008).
 - [9] J. Liphardt, B. Onoa, S. B. Smith, I. Tinoco, and C. Bustamante, Science **292**, 733 (2001).
 - [10] A. F. Oberhauser, P. K. Hansma, M. Carrion-Vazquez, and J. M. Fernandez, Proc. Natl. Acad. Sci. U. S. A. **98**, 468 (2001).
 - [11] E. Evans, Annu. Rev. Biophys. Biomol. Struct. **30**, 105 (2001).
 - [12] C. Bustamante, J. Liphardt, and F. Ritort, Phys. Today **7**, 43 (2005).
 - [13] K. C. Neuman and A. Nagy, Nat. Meth. **5**, 491 (2008).

- [14] L. B. Freund, Proc. Natl. Acad. Sci. U. S. A. **106**, 8818 (2009).
- [15] J. M. Huguet, C. V. Bizarro, N. Forns, S. B. Smith, C. Bustamante, and F. Ritort, Proc. Natl. Acad. Sci. U. S. A. **107**, 15431 (2010).
- [16] H. J. Kreuzer and S. H. Payne, Phys. Rev. E **63**, 021906 (2001).
- [17] M. Süzen, M. Sega, and C. Holm, Phys. Rev. E **79**, 051118 (2009).
- [18] S. Kirmizialtin, L. Huang, and D. E. Makarov, J. Chem. Phys. **122**, 234915 (2005).
- [19] N. Gunari, A. C. Balazs, and G. C. Walker, J. Am. Chem. Soc. **129**, 10046 (2007).
- [20] M. Rubinstein and R. H. Colby, *Polymer Physics* (Oxford University Press, New York, 2003).
- [21] A. Halperin and E. B. Zhulina, Europhys. Lett. **15**, 417 (1991).
- [22] A. Imparato, F. Sbrana, and M. Vassalli, Europhys. Lett. **82**, 58006 (2008).
- [23] C. Jarzynski, Phys. Rev. Lett. **78**, 2690 (1997).
- [24] C. Jarzynski, Eur. Phys. J. B **64**, 331 (2008).
- [25] C. Jarzynski, Annu. Rev. Condens. Matter Phys. **2**, 329 (2011).
- [26] J. Liphardt, S. Dumont, S. B. Smith, I. Tinoco, and C. Bustamante, Science **296**, 1832 (2002).
- [27] J. G. Kirkwood, J. Chem. Phys. **3**, 300 (1935).
- [28] I. Franco, G. C. Schatz, and M. A. Ratner, J. Chem. Phys. **131**, 124902 (2009).
- [29] A. M. Ferrenberg and R. H. Swendsen, Phys. Rev. Lett. **63**, 1195 (1989).
- [30] S. Kumar, D. Bouzida, R. H. Swendsen, P. A. Kollman, and J. M. Rosenberg, J. Comp. Chem. **13**, 1011 (1992).
- [31] B. Roux, Comput. Phys. Commun. **91**, 275 (1995).
- [32] D. Frenkel and B. Smit, *Understanding Molecular Simulation* (Academic Press, 2002), 2nd ed.
- [33] M. E. J. Newman and G. T. Barkena, *Monte Carlo Methods in Statistical Physics* (Oxford University Press, New York, 2007).
- [34] J. D. Chodera, W. C. Swope, J. W. Pitner, C. Seok, and K. A. Dill, J. Chem. Theory Comput. **3**, 26 (2007).
- [35] J. S. Hub, B. L. de Groot, and D. van der Spoel, J. Chem. Theory Comput. **6**, 3713 (2010).
- [36] L. Felberg, I. Franco, M. McCullagh, M. A. Ratner, G. C. Schatz, and M. Carignano, *Molpull: a tool for molecular free energy reconstruction along a pulling coordinate*, DOI:10254/nanohub-r9583.2 (2010).
- [37] A. Grossfield, *The weighted histogram analysis method*, <http://membrane.urmc.rochester.edu/content/wham>.
- [38] D. D. L. Minh, *Ferbe free energy reconstruction from biased experiments*, <https://simtk.org/home/ferbe>.
- [39] S. Izrailev, S. Stepaniants, B. Isralewitz, D. Kosztin, H. Lu, F. Molnar, W. Wriggers, and K. Schulten, in *Computational Molecular Dynamics: Challenges, Methods, Ideas*, edited by P. Deuffhard, J. Hermans, B. Leimkuhler, A. E. Marks, S. Reich, and R. D. Skeel (Springer-Verlag, 1998), vol. 4 of *Lecture Notes in Computational Science and Engineering*, pp. 39–65.
- [40] B. Isralewitz, M. Gao, and K. Schulten, Curr. Opin. Struct. Biol. **11**, 224 (2001).
- [41] B. Heymann and H. Grubmüller, Chem. Phys. Lett. **305**, 202 (1999).
- [42] S. Nosé, Mol. Phys. **52**, 255 (1984).

- [43] W. G. Hoover, *Phys. Rev. A* **31**, 1695 (1985).
- [44] G. J. Martyna, M. L. Klein, and M. Tuckerman, *J. Chem. Phys.* **97**, 2635 (1992).
- [45] J. M. Thijssen, *Computational Physics* (Cambridge University Press, New York, 2007), 2nd ed.
- [46] H. J. C. Berendsen, J. P. M. Postma, W. F. Vangunsteren, A. Dinola, and J. R. Haak, *J. Chem. Phys.* **81**, 3684 (1984).
- [47] S. C. Harvey, R. K. Z. Tan, and T. E. Cheatham, *J. Comp. Chem.* **19**, 726 (1998).
- [48] J. C. Phillips, R. Braun, W. Wang, J. Gumbart, E. Tajkhorshid, E. Villa, C. Chipot, R. D. Skeel, L. Kalé, and K. Schulten, *J. Comp. Chem.* **26**, 1781 (2005).
- [49] B. Hess, C. Kutzner, D. van der Spoel, and E. Lindahl, *J. Chem. Theory Comput.* **4**, 435 (2008).
- [50] J. Ponder, *Tinker: Software tools for molecular design 4.2*, <http://dasher.wustl.edu/tinker/> (2004).
- [51] I. Franco, C. B. George, G. C. Solomon, G. C. Schatz, and M. A. Ratner, *J. Am. Chem. Soc.* **133**, 2242 (2011).
- [52] L. V. Schäfer, E. M. Müller, H. E. Gaub, and H. Grubmüller, *Angew. Chem.* **119**, 2282 (2007).
- [53] E. H. Lee, J. Hsin, E. von Castelmur, O. Mayans, and K. Schulten, *Biophys. J.* **98**, 1085 (2010).
- [54] G. E. Crooks, *Phys. Rev. E* **60**, 2721 (1999).
- [55] G. Hummer and A. Szabo, *P. Natl. Acad. Sci. U.S.A.* **98**, 3658 (2001).
- [56] S. Park and K. Schulten, *J. Chem. Phys.* **120**, 5946 (2004).
- [57] G. Hummer and A. Szabo, *Acc. Chem. Res.* **38**, 504 (2005).
- [58] N. C. Harris, Y. S., and C.-H. Kiang, *Phys. Rev. Lett.* **99**, 068101 (2007).
- [59] D. D. L. Minh and A. B. Adib, *Phys. Rev. Lett.* **100**, 180602 (2008).
- [60] A. Pohorille, C. Jarzynski, and C. Chipot, *J. Phys. Chem. B* **114**, 10235 (2010).
- [61] G. Hummer and A. Szabo, *Proc. Natl. Acad. Sci. U. S. A.* **107**, 21441 (2010).
- [62] C. Jarzynski, *Phys. Rev. E* **73**, 046105 (2006).
- [63] S. Marsili and P. Procacci, *J. Phys. Chem. B* **114**, 2509 (2010).
- [64] I. Franco, M. A. Ratner, and G. C. Schatz, *J. Phys. Chem. B* **115**, 2477 (2011).
- [65] S. Basu, A. Coskun, D. C. Friedman, M. A. Olson, D. Benítez, E. Tkatchouk, G. Barin, J. Yang, A. C. Fahrenbach, W. A. Goddard, et al., *Chem. Eur. J.* **17**, 2107 (2011).
- [66] N. L. Allinger, Y. H. Yuh, and J. H. Lii, *J. Am. Chem. Soc.* **111**, 8551 (1989).
- [67] M. McCullagh, I. Franco, M. A. Ratner, and G. C. Schatz, *J. Am. Chem. Soc.* **133**, 3452 (2011).
- [68] R. Kubo, *J. Phys. Soc. Jpn.* **17**, 1100 (1962).
- [69] J. S. Kim, Y. J. Jung, J. W. Park, A. D. Shaller, W. Wan, and A. D. Q. Li, *Adv. Mater.* **21**, 786 (2009).
- [70] J. C. M. Gebhardt, T. Bornschlögl, and M. Rief, *Proc. Natl. Acad. Sci. U. S. A.* **107**, 2013 (2010).
- [71] To be precise, we actually deal with pseudorotaxanes that lack the steric blocking units at the chain termini that force the cyclophanes to remain attached to the oligomeric chain. For simplicity, we will use the term rotaxane.