Acc. Chem. Res. **2005,** *38,* 504−513

ARTICLES

Free Energy Surfaces from Single-Molecule Force Spectroscopy

GERHARD HUMMER* AND ATTILA SZABO*

Laboratory of Chemical Physics, Building 5, National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health, Bethesda, Maryland 20892-0520

Received July 22, 2004

ABSTRACT

Single-molecule force spectroscopy has the potential to provide unprecedented insights into the mechanical properties of individual molecules. The unfolding of proteins and nucleic acids, the dissociation of molecular complexes, and other molecular transitions can be induced through mechanical forces exerted, for example, by laser optical tweezers or atomic force microscopes and monitored with subnanometer resolution. Can one obtain the equilibrium free energy of the molecular system along the pulling coordinate from such nonequilibrium force measurements? Jarzynski's remarkable identity does not immediately solve this problem because it relates the nonequilibrium work to free energy differences at different times, not positions. By surmounting this difficulty, we were able to express the free energy profile in terms of the integral of the force with respect to extension. Here we present the theory in a simple way and discuss various practical aspects in the context of pulling experiments. We illustrate our rigorous free energy reconstruction procedure by applying it to force-induced RNA unfolding experiments.

1. Introduction

In single-molecule force spectroscopy, individual molecules or molecular assemblies are pulled apart by atomic force microscopes or laser optical tweezers to learn about their structure, dynamics, interactions, and mechanical properties.[1−15] In such experiments, a spring is attached to an anchored molecule or molecular assembly, possibly via molecular linkers (Figure 1). Pulling induces mechanical stress in the molecular system and, eventually, forces a molecular transition such as the unfolding of a nucleic acid or protein or the dissociation of a molecular complex. Analogous computer simulations[9,16−19] have been used to provide atomistically detailed pictures of molecular rupture processes.

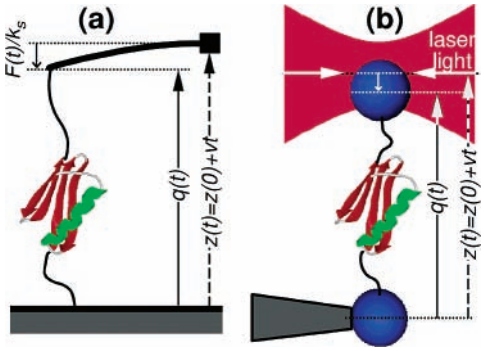

FIGURE 1. Single-molecule force spectroscopy using (a) atomic force microscopes and (b) laser optical tweezers. The anchored sample, indicated by the protein cartoons, is moved at a speed v relative to the pulling apparatus. The dashed vertical arrows indicate the controlled distance, $z(t) = z(0) + vt$, between the sample and pulling-spring anchors. The fluctuating molecular extension, $q(t)$, is indicated by solid vertical arrows and corresponds to the bead-to-bead distance in panel b. The deflections of (a) the cantilever or (b) the bead from their resting positions (small vertical arrows) indicate the instantaneous forces, $F(t) = k_s[z(t) - q(t)]$.

With spectacular resolutions of piconewton forces and angstrom molecular extensions, it may at first sight seem easy to extract useful equilibrium thermodynamic properties (binding constants, folding free energies, etc.) from single-molecule force spectroscopy. Unfortunately, in typical pulling experiments, a time-varying external force actively perturbs the molecular system. This leads to nonequilibrium effects and hysteresis. We know from the second law of thermodynamics that the work performed during a measurement will equal the free energy only if the experiment is performed reversibly (i.e., infinitely slowly). This may suggest that rigorous thermodynamic measurements at finite pulling speeds are impossible. However, Jarzynski's remarkable identity between thermodynamic free energy differences and the work along nonequilibrium trajectories[20,21] suggests that one should, at least in principle, be able to extract thermodynamic information from repeated nonequilibrium pulling experiments.

In practice, one is immediately faced with the difficulty that Jarzynski's identity relates the free energy difference between two thermodynamic states to the work performed in converting one state to another by continuously changing some control or switching parameter. In the context of pulling experiments, it gives the free energy of the entire system at two different times. Because the position of the pulling coordinate fluctuates, this is not the same as the free energy difference between two positions along the pulling coordinate.

In this Account, we will describe how to obtain the free energy along the pulling coordinate and discuss how to

Gerhard Hummer received a doctoral degree in physics for work done jointly at the Max-Planck Institute for Biophysical Chemistry in Göttingen and the University of Vienna (1992). In 1996, he started his independent career in the Theoretical Division of Los Alamos National Laboratory after his postdoctoral work there. In 1999, Hummer joined the Laboratory of Chemical Physics in the National Institute of Diabetes and Digestive and Kidney Diseases at the National Institutes of Health where he is a Senior Investigator.

Attila Szabo also works in the Laboratory of Chemical Physics. He obtained his B.Sc. from McGill (1968) and his Ph.D. from Harvard (1972). After postdocs at the Institut de Biologie Physico-Chimique in Paris and the Medical Research Council in Cambridge, he joined the Chemistry Department at Indiana University. In 1980, he moved to the National Institutes of Health.

10.1021/ar040148d
Published on Web 01/08/2005

remove contributions due to molecular linkers and the pulling apparatus. A simplified version of our procedure has been used by Liphardt et al.[15] to analyze force-extension curves obtained by unfolding RNA and by Schulten and collaborators[19] to extract free energy surfaces for aquaporin conduction from steered molecular dynamics simulations. We begin by putting Jarzynski's identity into the context of traditional methods of free energy calculation.[22,23] A unified path−integral approach to kinetics with fluctuating rates, Kubo−Anderson line-shape theory,[24] and the Feynman−Kac theorem of quantum mechanics[25] will then provide us with the necessary background for a simple derivation of Jarzynski's identity and our results for the free energy surfaces. After discussing some practical aspects in the context of pulling experiments, we illustrate our rigorous procedure[26] and some approximations to it by considering the recent RNA-unfolding experiment of Liphardt et al.[15] We conclude by briefly describing our related work on how kinetic information can be extracted from pulling experiments.[27]

2. Theory

2.1. Free Energy Perturbation Theory. Perhaps the simplest physical example of the kind of problems that motivated Jarzynski's seminal work is the calculation of the free energy required to charge an ion in solution. Let $H_0(\mathbf{x})$ be the Hamiltonian describing a system consisting of a neutral atom dissolved in water with x being a point in phase space. Let $H_0(\mathbf{x}) + \lambda V(\mathbf{x})$ be the Hamiltonian for the same system with a partial charge λe on the atom. The (Helmholtz) free energy of such a partially charged system is

$$\beta G(\lambda) = -\ln Q(\lambda) \tag{1}$$

where $\beta^{-1} = k_B T$ (k_B being Boltzmann's constant and T being the absolute temperature), and $Q(\lambda)$ is the canonical partition function,

$$Q(\lambda) = \int \mathrm{e}^{-\beta(H_0+\lambda V)}\, \mathrm{d}\mathbf{x} \tag{2}$$

The integral is over all points $\mathbf{x}$ in phase space, and for simplicity, combinatorial prefactors are ignored. Differentiating eq 1 with respect to λ and using eq 2, we have

$$\frac{\partial G}{\partial \lambda} = \frac{\int V\, \mathrm{e}^{-\beta(H_0+\lambda V)}\, \mathrm{d}\mathbf{x}}{\int \mathrm{e}^{-\beta(H_0+\lambda V)}\, \mathrm{d}\mathbf{x}} \equiv \langle V \rangle_\lambda \tag{3}$$

where the subscript λ indicates that the average is for a system with Hamiltonian $H_0 + \lambda V$. Integrating this from $\lambda = 0$ to $\lambda = 1$, we get

$$G(1) - G(0) = \Delta G = \int_0^1 \langle V \rangle_\lambda \, \mathrm{d}\lambda \tag{4}$$

This expression dates back at least to the work of Born[22] in 1920. It shows how one can obtain the free energy change from equilibrium averages of the perturbation when the perturbation is turned on infinitely slowly (and thus reversibly). Can one obtain the free energy difference when the perturbation is turned on instantaneously, that is, infinitely fast, by calculating some equilibrium average for the initial unperturbed state with Hamiltonian H_0? The answer is not only "yes" but also the result follows immediately from the definition of the free energy difference:[23]

$$\mathrm{e}^{-\beta\Delta G} = \frac{Q(1)}{Q(0)} = \frac{\int \mathrm{e}^{-\beta(H_0+V)}\, \mathrm{d}\mathbf{x}}{\int \mathrm{e}^{-\beta H_0}\, \mathrm{d}\mathbf{x}} = \langle \mathrm{e}^{-\beta V} \rangle_0 \tag{5}$$

Thus, the free energy difference can be obtained exactly by calculating the equilibrium average of the exponential of the perturbation using trajectories generated from the initial state. Although this result is formally exact, it is not very useful in practice when the perturbation is large, and relevant regions of state "1" in phase space have little population in state "0". To find configurations that contribute significantly to the exponential average (i.e., are "probable"), one has to run trajectories on the initial state for a very long time. In such cases, second-order perturbation theory (for technical reasons also called the "cumulant" expansion) can give better results. To obtain the perturbation expansion of the free energy, one expands the exponential in eq 5 in powers of V and then takes the logarithm of this expansion. In this way, it is easy to show that

$$\Delta G = \langle V \rangle_0 - \frac{\beta}{2}(\langle V^2 \rangle_0 - \langle V \rangle_0{}^2) \pm \ldots \tag{6}$$

With this background we are in a position to appreciate what Jarzynski did. He showed that if the perturbation is turned on during a finite time interval t [i.e., $\lambda(0) = 0$, $\lambda(t) = 1$], then the free energy difference between states "0" and "1" is exactly given by

$$\mathrm{e}^{-\beta\Delta G} = \left\langle \exp\left[-\beta \int_0^t \frac{\mathrm{d}\lambda(\tau)}{\mathrm{d}\tau} V[\mathbf{x}(\tau)]\, \mathrm{d}\tau \right] \right\rangle = \langle \mathrm{e}^{-\beta W_t} \rangle \tag{7}$$

The averages are taken over trajectories that start out from the equilibrium distribution corresponding to H_0 and are propagated according to the time-dependent Hamiltonian $H(t) = H_0 + \lambda(t)V$. W_t is the total work required to carry out this transformation. This relationship can be viewed as a simple result of the statistical mechanics of trajectory space.[28]

At first sight, Jarzynski's identity appears to be truly remarkable. It establishes an exact relation between equilibrium free energy changes and the irreversible work done to make these changes. The second law of thermodynamics tells us only that $\langle W_t \rangle \geq \Delta G$ (which incidentally follows immediately from eq 7 using Jensen's inequality, $\langle \mathrm{e}^{-x} \rangle \geq \mathrm{e}^{-\langle x \rangle}$), so it is unexpected that an equality actually exists. Moreover, in nonequilibrium statistical mechanics for systems far from equilibrium exact results are few and far between. However, in retrospect, the existence of such an identity is, as often happens, not all that surprising. If one can obtain the free energy difference by making the transformation infinitely slowly (eq 4) and infinitely fast (eq 5), why can't it be done in intermediate cases? In

addition, one must remember that Jarzynski's identity tells us nothing about the nature of nonequilibrium states. It only allows someone who is too impatient to carry out the transformation sufficiently slowly to obtain the equilibrium free energy differences by repeatedly making the transformation in a finite time. Can we prove Jarzynski's identity in a simple way as we did above for the limiting cases? The answer is "yes", but a little additional background is needed.

2.2. Fluctuating Rates, NMR Line Shapes, and Feynman Path Integrals. One understands new things by relating them to what one already knows or to what is familiar. Specialists in kinetics, NMR spectroscopy, or quantum mechanics may require explanations that sound different but are fundamentally the same.

Let us begin by providing the necessary background for a kineticist. Consider a first-order irreversible process described by a time-dependent rate coefficient $k(t)$. The concentration of interest then satisfies

$$\frac{\mathrm{d}C}{\mathrm{d}t} = -k(t)C \tag{8}$$

The solution of this first-order rate equation is

$$\frac{C}{C_0} = \mathrm{e}^{-\int_0^t k(\tau)\mathrm{d}\tau} \tag{9}$$

where C_0 is the initial concentration. Now suppose that the reason we used a time-dependent rate coefficient in the first place is that the rate coefficient fluctuates. As a simple example, consider Förster resonance energy transfer between two dyes attached to the ends of a polymer. Because the rate of transfer depends on the end-to-end distance and this distance fluctuates as the result of conformational dynamics, the transfer rate fluctuates. Suppose that at $t = 0$ we excite the donor in an equilibrium ensemble of polymers with a short light pulse. If we ignore the intrinsic lifetime, the survival probability of excited donors is then given by

$$S(t) = \frac{C}{C_0} = \langle \mathrm{e}^{-\int_0^t k(\tau)\mathrm{d}\tau} \rangle \tag{10}$$

where the angular brackets denote averaging over equilibrium trajectories of the end-to-end distance generated in the absence of energy transfer. This relation was, for instance, used to compare the quenching kinetics expected from molecular dynamics simulations[29] to that observed experimentally.[30]

How can we calculate the average in eq 10 for some model of the dynamics? Let us start with the simplest model: the polymer has only two states that interconvert according to the kinetic scheme $1 \rightleftharpoons 2$. If k_1 and k_2 are the transfer rates in these conformations, then

$$S(t) = \langle \mathrm{e}^{-\int_0^t k(\tau)\mathrm{d}\tau} \rangle = C_1(t) + C_2(t) \tag{11}$$

where $C_1(t)$ and $C_2(t)$ satisfy the rate equation

$$\frac{\mathrm{d}}{\mathrm{d}t}\begin{pmatrix} C_1 \\ C_2 \end{pmatrix} = \begin{pmatrix} -a & b \\ a & -b \end{pmatrix}\begin{pmatrix} C_1 \\ C_2 \end{pmatrix} - \begin{pmatrix} k_1 & 0 \\ 0 & k_2 \end{pmatrix}\begin{pmatrix} C_1 \\ C_2 \end{pmatrix} \tag{12}$$

subject to the initial conditions that at $t = 0$, C_1 and C_2 are the equilibrium concentrations for the kinetic scheme $1 \underset{b}{\overset{a}{\rightleftharpoons}} 2$ [i.e., $C_1^{\mathrm{eq}} = b/(a + b)$, $C_2^{\mathrm{eq}} = a/(a + b)$]. The first term on the right-hand side of eq 12 describes the dynamics of the two-state system, while the second takes into account the irreversible decay. It is easy to generalize this to many discrete states and even to states labeled by the continuous multidimensional vector **x**. The generalization of eq 11 is

$$\langle \mathrm{e}^{-\int_0^t k(\tau)\mathrm{d}\tau} \rangle = \int C(\mathbf{x},t)\,\mathrm{d}\mathbf{x} \tag{13}$$

where now $C(\mathbf{x},t)$ satisfies the generalization of eq 12

$$\frac{\partial C(\mathbf{x},t)}{\partial t} = \mathscr{L}C(\mathbf{x},t) - k(\mathbf{x})C(\mathbf{x},t) \tag{14}$$

$\mathscr{L}$ is the operator that describes the dynamics of the system at temperature T in the absence of the "sink" term $k(\mathbf{x})$ and has the property

$$\mathscr{L}\mathrm{e}^{-\beta H} = 0 \tag{15}$$

This is the analogue of

$$\begin{pmatrix} -a & b \\ a & -b \end{pmatrix}\begin{pmatrix} C_1^{\mathrm{eq}} \\ C_2^{\mathrm{eq}} \end{pmatrix} = 0 \tag{16}$$

and means that the Boltzmann distribution is stationary. Equation 14 must be solved subject to the equilibrium initial condition

$$C(\mathbf{x},0) = \frac{\mathrm{e}^{-\beta H(\mathbf{x})}}{\int \mathrm{e}^{-\beta H(\mathbf{x}')}\,\mathrm{d}\mathbf{x}'} \tag{17}$$

In applications, it is almost never possible to solve eq 14 analytically.

Let us now turn to the problem of determining the NMR line shape in the presence of chemical exchange. When the resonance frequency depends on the conformation, the line shape $I(\omega)$ is given by

$$I(\omega) = \frac{1}{2\pi}\int \mathrm{e}^{i\omega t}\langle \mathrm{e}^{-i\int_0^t \omega_0(\tau)\mathrm{d}\tau} \rangle\,\mathrm{d}t \tag{18}$$

If $\omega_0(t) = \omega_0$ is independent of time, $I(\omega) = \delta(\omega - \omega_0)$. Again, the simplest case is when the molecule has only two conformations with resonance frequencies ω_1 and ω_2. If the interconversion between the two conformations is described by the kinetic scheme $1 \underset{b}{\overset{a}{\rightleftharpoons}} 2$, then the line shape function can be obtained from eq 11 by solving eq 12 with $k_1 \rightarrow i\omega_1$ and $k_2 \rightarrow i\omega_2$. This is Kubo–Anderson line shape theory.[24] For more complicated dynamics, one must solve eq 14 with $k(\mathbf{x}) \rightarrow i\omega(\mathbf{x})$.

Finally, we mention the relationship to the Feynman path integral formulation of quantum mechanics (e.g., see ref 25). The time-dependent Schrödinger equation is

$$\frac{\partial \psi}{\partial t} = \frac{i\hbar}{2m}\nabla^2\psi - \frac{iV}{\hbar}\psi \tag{19}$$

The propagator or Greens function $G(\mathbf{x},t|\mathbf{x}_0,0)$ is the solution of this equation with initial condition $\psi(\mathbf{x},0) = \delta(\mathbf{x} - \mathbf{x}_0)$. In analogy with eqs 13 and 14, we can write

$$G(\mathbf{x},t|\mathbf{x}_0,0) = \langle e^{-i\int_0^t V[\mathbf{x}(\tau)]d\tau/\hbar}\rangle \tag{20}$$

where the average is over paths of the free quantum particle starting from $\mathbf{x}_0$ at $t = 0$ and ending in $\mathbf{x}$ at time t. In practice,[25] one normally works in imaginary time where the free-particle paths correspond to trajectories of a Brownian particle. The Feynman−Kac theorem states that the path integral representation of the propagator in eq 20 satisfies the Schrödinger equation, eq 19.

2.3. Proof of Jarzynski's Identity. We now show that Jarzynski's identity[20,21] relating free energy changes and nonequilibrium work follows almost immediately from the above results.[26] Rather than using a time-dependent coupling parameter, we consider a time-dependent Hamiltonian $H(\mathbf{x},t)$. This is the natural description of the entire system in a pulling experiment where the spring moves at, say, a constant velocity. We are interested in the free energy difference between equilibrium systems corresponding to Hamiltonians parametrized by times $t = 0$ and t

$$e^{-\beta\Delta G(t)} = \frac{Q(t)}{Q(0)} = \frac{\int e^{-\beta H(\mathbf{x},t)}\,d\mathbf{x}}{\int e^{-\beta H(\mathbf{x}',0)}\,d\mathbf{x}'} = \int C(\mathbf{x},t)\,d\mathbf{x} \tag{21}$$

where we have defined

$$C(\mathbf{x},t) = \frac{e^{-\beta H(\mathbf{x},t)}}{\int e^{-\beta H(\mathbf{x}',0)}\,d\mathbf{x}} \tag{22}$$

Note that $C(\mathbf{x},0)$ is simply given by eq 17 with $H(\mathbf{x}) \equiv H(\mathbf{x},0)$. We assume that the dynamics of the system is described by a time-dependent operator $\mathscr{L}_t$ that satisfies

$$\mathscr{L}_t e^{-\beta H(\mathbf{x},t)} = \mathscr{L}_t C(\mathbf{x},t) = 0 \tag{23}$$

as is, for instance, the case for the Liouville operator of classical mechanics or the Fokker−Planck operator for diffusion in phase space (i.e., Langevin dynamics). From eq 22 and the stationarity of the Boltzmann distribution, eq 23, it follows immediately by differentiating eq 22 with respect to time that

$$\frac{\partial C(\mathbf{x},t)}{\partial t} = \mathscr{L}_t C(\mathbf{x},t) - \beta\frac{\partial H(\mathbf{x},t)}{\partial t}C(\mathbf{x},t) \tag{24}$$

which has the same form as eq 14 when $k(\mathbf{x}) \to \beta\partial H(\mathbf{x},t)/\partial t$. Therefore, by comparing eqs 13 and 21, one can see that

$$e^{-\beta\Delta G(t)} = \langle e^{-\beta\int_0^t \frac{\partial H[\mathbf{x}(\tau),\tau]}{\partial\tau}d\tau}\rangle \equiv \langle e^{-\beta W_t}\rangle \tag{25}$$

where the averages are taken over all trajectories starting from an equilibrium initial distribution and evolving according to the time-dependent equation of motion corresponding to the operator $\mathscr{L}_t$. This is Jarzynski's celebrated theorem[20] relating the difference between the free energies at two different times to the average of the Boltzmann factor, $e^{-\beta W_t}$, of the external work done on the system.

Because of the formal similarity of Jarzynski's identity with Zwanzig's less general one in eq 5, many results obtained in the context of free energy perturbation theory can be immediately generalized. For example, in analogy with eq 6, the cumulant expansion of the free energy is[20,31]

$$\Delta G(t) = \langle W_t\rangle - \frac{\beta}{2}(\langle W_t^2\rangle - \langle W_t\rangle^2) \pm \ldots \tag{26}$$

2.4. Reconstruction of Free Energy Surfaces. Finally, we are in a position to consider the central problem of this article, namely, how to extract the underlying free energy surface from force-spectroscopy experiments using laser tweezers, atomic force microscopes, or steered molecular dynamics simulations.[1−18] In both experiment and simulations, a spring is attached to a molecular system, and this spring is then moved. Both these situations are described by the time-dependent Hamiltonian $H(\mathbf{x},t) = H_0(x) + V[q(\mathbf{x}),t]$ where $q(\mathbf{x})$ is the pulling coordinate. If the pulling is done with a harmonic spring moving at constant velocity v, then $V(q,t) = k_s(q - vt)^2/2$. We would like to extract the free energy of the system along the pulling coordinate q in the absence of externally applied forces. We define the free energy along q as

$$e^{-\beta[G_0(q)+\delta G]} = \langle\delta[q - q(\mathbf{x})]\rangle_0 = \frac{\int\delta[q - q(\mathbf{x})]\,e^{-\beta H_0(\mathbf{x})}\,d\mathbf{x}}{\int e^{-\beta H_0(\mathbf{x})}\,d\mathbf{x}} \tag{27}$$

where δG is an arbitrary constant. Since Jarzynski's identity relates free energy changes to a certain average of the external work done, it is clear that it has something to do with this problem. In fact, in the introduction of a paper addressing this problem in the context of steered molecular dynamics,[32] Schulten and co-workers used the Jarzynski identity to show that reconstruction of the free energy surface must be possible in principle, but then proposed several model-dependent procedures based on solving the Smoluchowski equation for diffusion in the presence of a potential. The difficulty in reconstructing free energy surfaces is that Jarzynski's identity is *not*[15,32]

$$e^{-\beta G_0(q)} = \langle e^{-\beta W(q)}\rangle$$

This is because Jarzynski's identity, eq 7, deals with free energy differences between *states* described by different values of the coupling parameters λ (here, times t), not different molecular positions. In a pulling experiment, only the anchor $z(t)$ of the pulling spring is a controllable

coupling parameter (Figure 1). The point of attachment to the molecular system (i.e., the "tip" of the spring), q, can take on a variety of positions at a given time t. The free energy along that fluctuating position q is thus a *potential of mean force*.

After we found the simple derivation of Jarzynski's identity presented above,[26] it became clear that what is needed is $C(\mathbf{x},t)$ in eq 22. So, when $H(\mathbf{x},t) = H_0(\mathbf{x}) + V[q(\mathbf{x}),t]$, one has

$$\frac{\mathrm{e}^{-\beta\{H_0(\mathbf{x})+V[q(\mathbf{x}),t]\}}}{\int \mathrm{e}^{-\beta\{H_0(\mathbf{x}')+V[q(\mathbf{x}'),0]\}}\,\mathrm{d}\mathbf{x}'} = \left\langle \delta[\mathbf{x}-\mathbf{x}(t)]\,\exp\left[-\beta\int_0^t \frac{\partial V[\mathbf{x}(\tau),\tau]}{\partial\tau}\,\mathrm{d}\tau\right]\right\rangle = \langle \delta[\mathbf{x}-\mathbf{x}(t)]\,\mathrm{e}^{-\beta W_\mathrm{t}}\rangle \quad (28)$$

This is the key result upon which our procedure is built. Operationally, this amounts to simply sticking Dirac δ functions into eqs 21 and 25 in the right places. Integrating eq 28 over $\mathbf{x}$ gives the Jarzynski identity in eq 25.

In retrospect, this result is implicit in the work of Jarzynski[20,21] and Crooks.[33] In a deep and insightful paper analyzing all aspects of the relationship between free energies and irreversible work, Crooks presented the following identity in an unnumbered equation:[33]

$$\langle f[x(\tau)]\,\mathrm{e}^{-\beta W_\mathrm{d}}\rangle_\mathrm{neq} = \langle f[x(\tau)]\rangle_\mathrm{eq} \quad (29)$$

which he attributed to Jarzynski. This identity relates a nonequilibrium path average of an arbitrary function $f(x)$ (on the left-hand side) to an equilibrium average over the ensemble corresponding to the Hamiltonian at time τ (on the right-hand side). W_d is the dissipated work, $W_\mathrm{d} = W_\mathrm{t} - \Delta G(\tau)$. If one chooses $f[\mathbf{x}(\tau)] = \delta[\mathbf{x} - \mathbf{x}(\tau)]$ and uses eq 21 for $\Delta G(\tau)$, one gets eq 28.

To relate the general relation, eq 28, to single-molecule pulling, let us consider the special case that the perturbation depends only on the pulling coordinate q and time t, that is, $V(\mathbf{x},t) = V[q(\mathbf{x}),t]$. Multiplying both sides of eq 28 by $\mathrm{e}^{\beta V(q,t)}\delta[q - q(\mathbf{x})]$, integrating with respect to $\mathbf{x}$, and using the definition in eq 27, we have[26]

$$\mathrm{e}^{-\beta G_0(q)} = \left\langle \delta[q - q\,(\mathbf{x}(t))]\,\exp\left[-\beta\left(\int_0^t \frac{\partial V\{q[\mathbf{x}(\tau)],\tau\}}{\partial\tau} - V\{q[\mathbf{x}(t)],t\}\right)\right]\right\rangle \quad (30)$$

where $\langle ...\rangle$ is an average over all trajectories starting from a Boltzmann distribution for the Hamiltonian $H(\mathbf{x},0) = H_0(\mathbf{x}) + V[q(\mathbf{x}),0]$ and evolving according to the time-dependent $\mathscr{L}_t$. The constant δG in the definition eq 27 is the free energy difference between the entire system at δG and the molecular system, $\mathrm{e}^{-\beta\delta G} = \int \mathrm{e}^{-\beta\{H_0(\mathbf{x})+V[q(\mathbf{x}),0]\}}\,\mathrm{d}\mathbf{x}/\int \mathrm{e}^{-\beta H_0(\mathbf{x})}\,\mathrm{d}\mathbf{x}$. Because δG is independent of time, eq 30 allows us to find the free energy along the pulling coordinate from trajectories of position versus time.

More commonly, force-versus-position curves are reported from pulling experiments. Integrating the identity $\mathrm{d}V = (\partial V/\partial q)\,\mathrm{d}q + (\partial V/\partial t)\,\mathrm{d}t$ from $[t = 0, q = q(0)]$ to $[t = t, q = q(\mathrm{t})]$ and noting that the restoring force is $F(q,t) = -\partial V/\partial q$, we have

$$W_\mathrm{t} = \int_0^t \frac{\partial V[q(\tau),\tau]}{\partial\tau}\,\mathrm{d}\tau = \int_C F\,\mathrm{d}q + V[q(t),t] - V[q(0),0] \quad (31)$$

where the integral[34] over q is along the position-versus-time contour connecting $q(0)$ and $q(t)$. Using this in eq 30, we have[26]

$$\mathrm{e}^{-\beta G_0(q)} = \langle \delta[q - q(\mathbf{x}(t))]\,\mathrm{e}^{-\beta(\int_C F\mathrm{d}q - V[q(0),0])}\rangle \quad (32)$$

This and eq 30 are our central results. When the work is calculated as a time integral, as in eq 30, to get the free energy along the pulling coordinate, one must subtract the energy stored in the pulling apparatus at time t. Alternatively, when the work is obtained as a force-extension contour integral, as in eq 32, it is the initial energy, $V[q(0),0]$, that must be subtracted. Schurr and Fujimoto,[35] in rederiving some of these relations, called W_t the "accumulated work," as distinguished from the "transferred work" $\int_C F\,\mathrm{d}q$.

3. Practical Implementation

Based on the rigorous relations eqs 30 and 32, we can now develop practical approaches to estimate the free energy surface $G_0(q)$ from repeated pulling experiments. But before proceeding to the actual analysis, we briefly discuss how the pulling measurements (or steered molecular dynamics simulations) should be performed to conform with eqs 30 and 32. We then describe a histogram-based and a moment-based approach to analyze the results.

3.1. Initialization, Execution, and Noise Corrections. For eqs 30 and 32 to be valid, the trajectories used to obtain the averages must be generated in a specific way. For simplicity, we discuss the requirements for a harmonic pulling spring moving at a velocity v so that $V(q,t) = k_\mathrm{s}[q - z(t)]^2/2$ where $z(t) = z(0) + vt$ for $t \geq 0$ is the position of the anchor of the pulling spring and $q(\mathbf{x})$ is the point of attachment of the molecular system and the pulling spring projected onto the pulling direction (see Figure 1).

Equations 30 and 32 require that the initial conditions of pulling trajectories $[\mathbf{x}(t{=}0)]$ are chosen from an equilibrium distribution corresponding to the Hamiltonian $H(\mathbf{x},0) = H_0(\mathbf{x}) + k[q(\mathbf{x}) - z(0)]^2/2$. In an experiment, the pulling apparatus should thus be equilibrated at a fixed position $z(0)$ in all trajectories. Furthermore, the control parameter must have the same time dependence, $z(t)$, in each trajectory. One of the resulting difficulties in an experiment (but not in a computer simulation) is that one has to know the absolute pulling apparatus position $z(t)$ in a given trajectory. Instrument noise will likely broaden the observed work distributions, thereby lowering the estimated free energy differences. By performing analogous experiments without a load (i.e., pulling at velocity v but without an attached molecule), one can correct for this effect. Without load, the free energy profile should be flat. To correct for instrument noise, the $G_{0,\mathrm{nl}}(q)$ estimated from the no-load experiments should be sub-

tracted from the $G_0(q)$ obtained from experiments with attached molecules.

3.2. Analysis Using Weighted Histograms. If one has a sufficient number of properly sampled pulling traces, the path averages in eqs 30 and 32 can be evaluated with a histogram technique. In principle (i.e., for an infinite number of pulling traces), the entire free energy surface $G_0(q)$ could be reconstructed from observations at a single time $t \geq 0$. In practice, at any time t the trajectories will likely be clustered near the location of the pulling spring. Consequently, from observations collected at time t, one can find reliable estimates of $G_0(q)$ only near $q \approx z(t)$.

However, we can combine multiple histograms obtained for different times t to improve our estimate of $G_0(q)$. This is analogous to the weighted histogram approach of Ferrenberg and Swendsen.[36] Here, we have to unbias the observations not only with respect to the potential of the pulling spring, $V(q,t)$, as in usual umbrella sampling, but also with respect to the nonequilibrium work, W_t. This leads to the following expression for the free energy profile:[26]

$$e^{-\beta G_0(q)} = \frac{\sum_t \frac{\langle \delta[q - q(t)] \exp(-\beta W_t)\rangle}{\langle \exp(-\beta W_t)\rangle}}{\sum_t \frac{\exp(-\beta V(q,t))}{\langle \exp(-\beta W_t)\rangle}} \qquad (33)$$

where the sums are over the histograms collected at different times t. In this expression, we have used the Jarzynski identity, $e^{-\beta \Delta G(t)} = \langle e^{-\beta W_t}\rangle$. A self-consistent formula that does not use this $\Delta G(t)$ has also been developed.[26] Implementation of eq 33 is discussed in ref 26 (but note that in eq 10 of this paper, u_{ik} should be u_{il}).

3.3. Analysis Using Moments. In many practical situations, one may not have enough trajectories to use a histogram analysis. One can then use a simple moment-based approach. If the pulling spring is relatively stiff, then most trajectories will be clustered near $q[\mathbf{x}(t)] \approx z(t)$ at any given time t. Given the mean

$$\overline{q_t} = \frac{\langle q(t)\, e^{-\beta W_t}\rangle}{\langle e^{-\beta W_t}\rangle} \qquad (34)$$

and variance

$$\sigma_t^2 = \frac{\langle q^2(t)\, e^{-\beta W_t}\rangle}{\langle e^{-\beta W_t}\rangle} - \overline{q_t}^2 \qquad (35)$$

one can approximate the (weighted!) distribution of positions by a Gaussian,

$$e^{-\beta[G_0(q)+V(q,t)]} = \langle \delta[q - q(t)]\, e^{-\beta W_t}\rangle \approx e^{-(q-\overline{q}_t)^2/(2\sigma_t^2)} \frac{\langle e^{-\beta W_t}\rangle}{(2\pi\sigma_t^2)^{1/2}} \qquad (36)$$

We note that near a sharp, cusp-like feature in $G_0(q)$ corresponding to a transition point, the distribution of q's may be multimodal,[14] so the Gaussian approximation in eq 36 may be poor. Using the Gaussian distribution, eq 36, results in the following approximation

$$G_0(q) \approx k_B T \frac{(q - \overline{q_t})^2}{2\sigma_t^2} - V(q,t) - \beta^{-1} \ln \frac{\langle e^{-\beta W_t}\rangle}{(2\pi\sigma_t^2)^{1/2}} \qquad (37)$$

where the logarithmic constant in the last term "matches" the locally quadratic expansions of the free energy surface $G_0(q)$ obtained at different times t. Equation 37 is a convenient starting point for an expansion in cumulants of the work W_t. Here, we simply take the derivative with respect to q and obtain

$$G_0'(\overline{q_t}) \approx -\frac{\partial V(q,t)}{\partial q}\Big|_{q=\overline{q_t}} = F(\overline{q_t}, t) \qquad (38)$$

The right-hand side is the restoring force measured by the pulling apparatus at the weighted average position, $\bar{q}_t$, at time t. If the variance at σ_t^2 can be estimated reliably, one can also estimate the second derivative

$$G_0''(\bar{q}_t) \approx F'(\overline{q_t}, t) + \frac{k_B T}{\sigma_t^2} \qquad (39)$$

To implement the moment-based formula eq 38, one needs to average the position q according to eq 34, where each observation is weighted with the Boltzmann factor of the accumulated work, W_t, given in eq 31. For a harmonic pulling apparatus, $V(q,t) = k_s[q - z(t)]^2/2$, we have $F(t) = k_s[z(t) - q]$ such that

$$G_0'[z(t) - \bar{F}_t/k_s] \approx \bar{F}_t = \langle F(t)\, e^{-\beta W_t}\rangle / \langle e^{-\beta W_t}\rangle \qquad (40)$$

Within the Gaussian approximation, the restoring force averaged with the Boltzmann-weighted work thus gives the first derivative of the potential of mean force, $G_0(q)$, that is, the mean force.

3.4. Error and Efficiency: Many Fast versus Few Slow Experiments? Given a certain amount of measurement time, is it better to run many fast and short pulling experiments or few slow and long ones? In a simple error analysis, one can assume that the amount of dissipated work, $W_t - \Delta G(t)$, grows linearly with time, as one would expect from linear response theory.[37] As was shown in ref 31, for sufficiently slow pulling the most accurate estimate of the free energy difference is then obtained by a single slow experiment. However, the expected error of the estimate increases very little if instead of the single long experiment of time τ, N short ones of time τ/N are performed, as long as the standard deviation of the accumulated work is around $k_B T$. Because multiple measurements allow us to put error bars on the calculated result, the "optimum" allocation is to run as many such runs as possible.

4. Analysis of an Experiment: Forced Unfolding of RNA

In a landmark paper, Liphardt et al.[15] used a simplified version of our formalism to obtain free energy profiles for RNA unfolding from force-versus-extension curves. Their

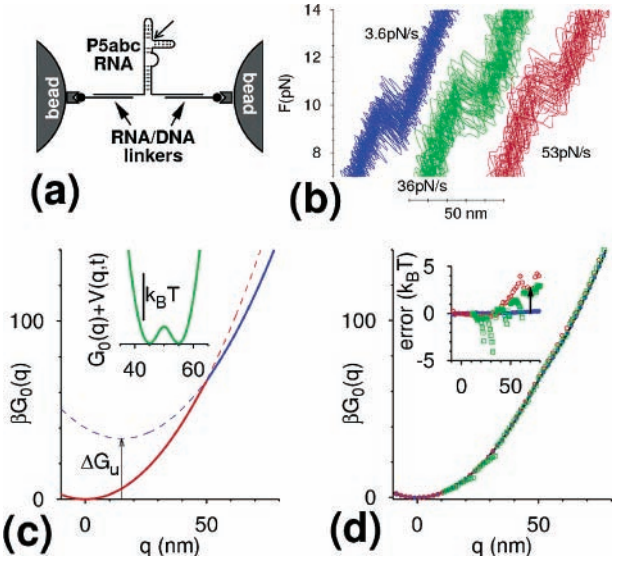

FIGURE 2. Analysis of RNA unfolding experiment of Liphardt et al.[15] Panel a gives a schematic representation of the setup. P5abc RNA is attached to two beads through long RNA/DNA hybrid linkers. The beads are pulled apart using a laser optical tweezer. An arrow indicates the position up to which we conclude that the RNA unfolds in the first unfolding transition at forces around 9.5 pN, the remaining structure unfolding at a higher force of ~12 pN (not shown). Panel b shows pulling traces from simulations, smoothed by convolution with a triangular window function, similar to those seen in the experiments near the first unfolding transition (Figure 1B of ref 15). Panel c gives the potential of mean force, $G_0(q)$ (thick solid red and blue lines), comprised of a folded branch (red) and unfolded branch (blue). The arrow indicates the free energy, ΔG_u, of the RNA unfolding transition. The inset shows the combined potential, $G_0(q) + V(q,t)$, at a time t where it is bistable. Panel d presents reconstructions of $G_0(q)$, which for our one-dimensional model is identical to $G_0(q)$ (black line). The blue solid and red open spheres are reconstructions using the weighted histogram formula, eq 33, for 250 trajectories, each at force loading rates of 3.6 and 53 pN/s, respectively, starting from an equilibrium distribution around $q = 0$. The open green squares use the approximate force-extension integral, eq 45, for the same trajectories but starting the integration at an extension of $q_{start} = 10$ nm. The inset shows the difference of the reconstructed potential. The arrow indicates the bias in the estimator of $-\ln \langle e^{-\beta W_t} \rangle$ at $\bar{q}_t$ expected for a Gaussian of the same variance in βW_t. The bias was estimated by repeatedly drawing samples of 250 Gaussian random numbers, calculating the logarithm of the average of the exponential of these numbers, and comparing the average of that to the exact result given by eq 26.

experiment is illustrated schematically in Figure 2a. The ends of a folded RNA are attached to two beads by long flexible linkers. The beads are moved apart by an optical tweezer, thus unfolding the RNA. Typical force versus extension curves are shown schematically in Figure 2b for different force loading rates. There are two regimes in which the force depends almost linearly on distance (with almost the same slope). These regimes are connected by random, sudden jumps in the distance that occur when the applied force becomes sufficiently large (about 9−12 pN).

To fully understand the analysis in ref 15 and to test the validity of the approximations, we started by generating "data" from simulations of a simple model that captured what seemed to us to be the essential physics of the problem. Then we analyzed the simulated data using various versions of our formalism to see how well they worked. The measurable potential of mean force, $G_0(q)$, will be a convolution of contributions, $G_{linker}(q)$, from the molecular linkers (curved black lines in Figure 1) and the molecule of interest, $G_{mol}(q)$ (protein in Figure 1):

$$e^{-\beta G_0(q)} = \int e^{-\beta G_{mol}(q')}\, e^{-\beta G_{linker}(q-q')}\, dq' \qquad (41)$$

By reconstructing the free energy surface twice, once using measurements on the linker, without RNA, and a second time for the combined system, one could, in principle, obtain $G_{mol}(q)$ after deconvolution. Here, we can greatly simplify this analysis if we assume that the RNA unfolding is instantaneous on the time scale of the pulling experiment[15] and it simply leads to an increase of Δq in the RNA end-to-end distance. Then, the integral in eq 41 separates into a sum of two terms, and the potential of mean force along q of the combined system is given by $e^{-\beta G_0(q)} = e^{-\beta G_f(q)} + e^{-\beta G_u(q)}$ where G_f and G_u, are the free energies of the entire system with folded and unfolded RNAs, respectively. Near the relevant region of the unfolding transition, the measured force-extension curves are essentially linear, due to effectively "harmonic" linkers. Thus, the free energies of the folded (G_f) and unfolded (G_u) states of the system are

$$G_f(q) = \frac{k_0}{2} q^2 \qquad (42)$$

$$G_u(q) = \frac{k_0}{2}(q - \Delta q)^2 + \Delta G_u \qquad (43)$$

where k_0 is the effective force constant of the linkers, Δq is the apparent increase in length due to RNA unfolding, and $\Delta G_u > 0$ is the unfolding free energy. For this model, the unfolding free energy ΔG_u can be obtained by measuring the average force $F_* = G_0'(q^*)$ on a reversible (slow) path at the midpoint of the transition where $G_f(q^*) = G_u(q^*)$ using

$$\Delta G_u = F_* \Delta q \qquad (44)$$

At finite pulling speed, in analogy to eqs 6 and 26, one can approximate ΔG_u as $\bar{F}_* \Delta q - \beta(\overline{F_*^2} - \bar{F}_*^2)\Delta q^2/2$.

The total free energy is $G(q,t) = G_0(q) + k_s(q - vt)^2/2$ where the spring constant of the pulling apparatus is taken to be $k_s = 0.1$ pN/nm, which is in the range of values reported for an earlier experiment.[14] The parameters $k_0 = 0.22$ pN/nm and $\Delta q = 15$ nm were chosen by eye to reproduce the slopes and the distance between the linear regimes of the experimental force extension curves (Figure 2b of ref 15). The unfolding free energy ΔG_u was chosen so that at the transition, the mean force would agree with experiments at corresponding pulling speeds.

The resulting potential of mean force $G_0(q)$ is shown in Figure 2c. The folded potential (red) is a parabola centered at $q = 0$. The unfolded potential (blue) is a parabola centered at $\Delta q = 15$ nm. The two curves cross

at $q^* = 50$ nm, which corresponds to an unfolding free energy of $\Delta G_u = 34 k_B T$. This apparent potential of mean force $G_0(q)$ is not bistable, and the RNA unfolding merely leads to a small change in slope at $q^* = 50$ nm. This is because the potential of mean force is dominated by the behavior of the linker molecules. However, the combined potential of mean force that includes the pulling spring, $V(q,t)$, can be bistable at a certain position of the pulling apparatus, as shown in the inset of Figure 2c. This explains why, even though the intrinsic potential of mean force is nearly featureless, reversible jumps between two states can be seen in quasi-equilibrium experiments where RNA is held close to the point of rupture.[14]

On the above free energy surface, we performed Brownian dynamics simulations at force loading rates $k_s v = 3.6$, 36, and 53 pN/s. The diffusion constant $D = 1200$ nm^2/s in the Brownian dynamics simulations was chosen such that the mean rupture force would agree with the measurements at different force loading rates. The resulting pulling traces shown in Figure 2b are very similar to the experimental ones.[15] We analyzed these simulated curves using our formally exact reconstruction formulas given in eqs 30 and 32. In addition, we used an expression involving integration of the force over distance, omitting the bias $V(q,0)$. This approximation was suggested by us in our original paper[26] and was adopted by Liphardt et al.[15] for the analysis of their data. Such an approximation is reasonable when the path integral eq 32 can be factorized,

$$e^{-\beta G_0(q)} \approx \langle \exp[-\beta (\int_{q(0)}^{q_{start}} F\, dq - V[q(0),0])] \rangle \times \langle \exp[-\beta \int_{q_{start}}^{q} F\, dq] \rangle \quad (45)$$

so that the first term is a constant. In this somewhat casual notation, the force-extension integral is taken over all trajectories crossing the boundaries q_{start} and q for the first time. It can be seen from Figure 2d that eq 45 is an excellent approximation for the present system when q_{start} is in the linear pretransition regime. Results from eq 45 are comparable in accuracy to those obtained from the histogram expression eq 33. A free energy reconstruction based on the average force, eq 40, also performs with comparable accuracy (not shown) even though the distributions of q are not always Gaussian. At the highest pulling speed and at the end of the pulling interval (~80 nm), the reconstructed free energy exceeds the actual one by about $3 k_B T$, or about 2%. This small difference can be explained quantitatively by the expected bias of the estimator. Thus, essentially the same free energy surface is obtained for all pulling speeds, and the approximate expression eq 45 appears to give accurate results for our model system of the RNA unfolding experiment.

In addition to showing that the same free energy profile can be extracted from experiments performed at different pulling speeds, Liphardt et al.[15] estimated a free energy difference between the folded and unfolded states of the entire system. Specifically, they found the free energy difference between two points on the surface that bracket the transition region. Their distance of 30 nm was chosen to correspond to the estimated increase in the end-to-end distance of the RNA upon fully unfolding. In this way, they obtained a free energy difference of $60 k_B T$.[15] If in Figure 2, panels c or d, we choose the interval $25 \le q \le 55$ nm, we also find that the free energy difference is $60 k_B T$. Moreover, integrated over this interval, the mean and variance of the work at the various pulling speeds are close to the experimental values (Supporting Information). However, in our model the free energy difference between $q = 55$ nm and 25 nm has a significant contribution from the linkers, $60 k_B T - \Delta G_u \approx 26 k_B T \approx {}^1/_2 k_s((55\text{ nm} - \Delta q)^2 - (25\text{ nm})^2)$.

Why is our estimate of the RNA unfolding free energy, $\Delta G_u = 34 k_B T$, lower than the equilibrium experimental value of $60 k_B T$?[14] We think that this is because under the conditions of the experiment, the RNA chosen for the experiments unfolds in two steps. Evidence for a two-step transition was reported in an earlier paper by Liphardt et al.[14] It is only the first step that is modeled here, and this step may correspond to unfolding of RNA up to the arrow in Figure 2a. There is a hint in the experimental data (Figure 2B of ref 15) that a second transition does indeed occur at a higher force of about 12 pN with a jump width of about $\Delta q' = 8-10$ nm. This is about the right length to account for unfolding of the remaining RNA secondary structure, and combined with the first transition, the free energy of unfolding on the "reversible" path is in the expected range, 12 pN × 9 nm + $\Delta G_u \approx 60 k_B T$, which agrees with the value reported near equilibrium.[14]

5. Concluding Remarks

So far, we have discussed how *equilibrium* properties can be obtained from nonequilibrium single-molecule force spectroscopy. Can one also find kinetic rate coefficients? For forced unfolding or dissociation, we could start by first extracting the potential of mean force, $G_0(q)$, underlying the molecular rupture event. This may require a deconvolution step to remove the contribution of a molecular linker. In addition, we would also need to approximate the dynamics on this surface, for instance by diffusion, and then estimate a position-dependent diffusion coefficient, $D(q)$. Given $G_0(q)$ and $D(q)$, the Smoluchowski equation could be solved to estimate the intrinsic rate of molecular dissociation or unfolding. Considering the challenges already encountered in obtaining an accurate molecular free energy surface, this approach seems premature.

A simple and widely used phenomenological theory seems to provide a much easier route to rate coefficients. According to Bell,[38] a constant external force F should accelerate rupture as $k(F) = k_0\, e^{\beta x^\ddagger F}$ where k_0 is the intrinsic rate of rupture, and $x^\ddagger$ is the location of the transition state relative to the free energy minimum. This simple relation can be generalized to explicitly time-dependent forces

(e.g., see refs 7 and 39−41), resulting in a time-dependent rate coefficient of rupture,

$$k(t) = k_0\, e^{\beta x^{\ddagger} F(t)} \tag{46}$$

This time-dependent rate, combined with eq 10 for the survival probability, gives us the fraction $S(t)$ of molecules that have not ruptured up to time t.[27] The distribution of rupture forces can then be found using $p(F)\, dF = -\partial S(t)/\partial t\, dt$. However, when tested against even the simplest model of rupture, Brownian dynamics in a harmonic well with a cusplike barrier, this phenomenological theory performs poorly, leading to intrinsic rate coefficients k_0 that can be off by orders of magnitude.[27]

As a simple alternative, we developed a model-dependent formalism to estimate rate coefficients from single-molecule pulling.[27] An analytic solution for Brownian dynamics in a harmonic well with a cusplike barrier is fitted globally to the experimental data measured over a broad range of force loading rates. In this approach, we also use the survival probability of eq 10 but with a time-dependent rate coefficient calculated using Kramers theory,

$$k(t) \approx k_0\left(1 - \frac{x^{\ddagger} F(t)}{2\Delta G^{\ddagger}}\right) e^{\beta x^{\ddagger} F(t)[1 - x^{\ddagger} F(t)/(4\Delta G^{\ddagger})]} \tag{47}$$

$k(t)$ is given here only for the limit of a stiff molecular coordinate and a soft pulling spring with a force constant k_s that includes softening linker contributions,[27] so $F(t) = k_s v t$ is the average force at time t. $\Delta G^{\ddagger}$ is the free energy barrier to rupture, which is not considered in the phenomenological model. In the limit of a high barrier, $\Delta G^{\ddagger} \to \infty$, our theory approaches the phenomenological limit, eq 46. In ref 27, we tested this formalism against computer simulations and showed that it works well even for multimodule constructs, such as titin, that are connected by anharmonic linkers. We also applied the formalism to the unfolding of the I27 titin module[11] to compare rate coefficients from force-induced and chemical denaturation, and the extracted free energy barrier $\Delta G^{\ddagger}$ with an estimate from recent measurements using single-molecule fluorescence spectroscopy.[42]

Single-molecule force spectroscopy is still a relatively new technique, but its utility has already been demonstrated in both experiments[1−15] and simulations.[9,6−19] Its major strength, the ability to induce and monitor mechanical transitions in single molecules, is also a source of difficulties because these transitions occur under nonequilibrium conditions and are coupled to a pulling apparatus. As we showed in this and an earlier paper,[26] an extension of Jarzynski's identity[20,21] nevertheless makes it possible to extract thermodynamic information rigorously from repeated nonequilibrium pulling experiments. Even though the accuracy required for such measurements may seem daunting—the dissipated work should not exceed a few $k_B T$!—the first such measurement has already been performed by Liphardt et al.[15] As the resolution and accuracy of force spectroscopy further improves, we hope that the theories discussed in this paper[26,27] will enable routine measurements of thermodynamic and kinetic properties of single molecules.

Supporting Information Available: Free energy of stretching RNA from $q_0 = 25$ nm to $q_0 + \delta q$. This material is available free of charge via the Internet at http://pubs.acs.org.

References

(1) Perkins, T. T.; Smith, D. E.; Chu, S. Direct observation of tube-like motion of a single polymer chain. *Science* **1994**, *264*, 819−822.
(2) Florin, E. L.; Moy, V. T.; Gaub, H. E. Adhesion forces between individual ligand−receptor pairs. *Science* **1994**, *264*, 415−417.
(3) Strick, T. R.; Allemand, J. F.; Bensimon, D.; Bensimon, A.; Croquette, V. The elasticity of a single supercoiled DNA molecule. *Science* **1996**, *271*, 1835−1837.
(4) Smith, S. B.; Cui, Y. J.; Bustamante, C. Overstretching B-DNA. The elastic response of individual double-stranded and single-stranded DNA molecules. *Science* **1996**, *271*, 795−799.
(5) Kellermayer, M. S. Z.; Smith, S. B.; Granzier, H. L.; Bustamante, C. Folding-unfolding transitions in single titin molecules characterized with laser tweezers. *Science* **1997**, *276*, 1112−1116.
(6) Tskhovrebova, L.; Trinick, J.; Sleep, J. A.; Simmons, R. M. Elasticity and unfolding of single molecules of the giant muscle protein titin. *Nature* **1997**, *387*, 308−312.
(7) Rief, M.; Gautel, M.; Oesterhelt, F.; Fernandez, J. M.; Gaub, H. E. Reversible unfolding of individual titin immunoglobulin domains by AFM. *Science* **1997**, *276*, 1109−1112.
(8) Oberhauser, A. F.; Marszalek, P. E.; Erickson, H. P.; Fernandez, J. M. The molecular elasticity of the extracellular matrix protein tenascin. *Nature* **1998**, *393*, 181−185.
(9) Marszalek, P. E.; Lu, H.; Li, H. B.; Carrion-Vazquez, M.; Oberhauser, A. F.; Schulten,K.; Fernandez, J. M. Mechanical unfolding intermediates in titin modules. *Nature* **1999**, *402*, 100−103.
(10) Merkel, R.; Nassoy, P.; Leung, A.; Ritchie, K.; Evans, E. Energy landscapes of receptor−ligand bonds explored with dynamic force spectroscopy. *Nature* **1999**, *397*, 50−53.
(11) Carrion-Vazquez, M.; Oberhauser, A. F.; Fowler, S. B.; Marszalek, P. E.; Broedel, S. E.; Clarke, J.; Fernandez, J. M. Mechanical and chemical unfolding of a single protein. A comparison. *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96*, 3694−3699.
(12) Oesterhelt, F.; Oesterhelt, D.; Pfeiffer, M.; Engel, A.; Gaub, H. E.; Mfuller, D. J. Unfolding pathways of individual bacteriorhodopsins. *Science* **2000**, *288*, 143−146.
(13) Cui, Y.; Bustamante, C. Pulling a single chromatin fiber reveals the forces that maintain its higher-order structure. *Proc. Natl. Acad. Sci. U.S.A.* **2000**, *97*, 127−132.
(14) Liphardt, J.; Onoa, B.; Smith, S. B.; Tinoco, I., Jr.; Bustamante, C. Reversible unfolding of single RNA molecules by mechanical force. *Science* **2001**, *292*, 733−737.
(15) Liphardt, J.; Dumont, S.; Smith, S. B.; Tinoco, I.; Bustamante, C. Equilibrium information from nonequilibrium measurements in an experimental test of Jarzynski's equality. *Science* **2002**, *296*, 1832−1835.
(16) Grubmüller, H.; Heymann, B.; Tavan, P. Ligand binding molecular mechanics calculation of the streptavidin biotin rupture force. *Science* **1996**, *271*, 997−999.
(17) Isralewitz, B.; Izrailev, S.; Schulten, K. Binding pathway of retinal to bacterio-opsin. A prediction by molecular dynamics simulations. *Biophys. J.* **1997**, *73*, 2972−2979.
(18) Paci, E.; Karplus, M. Forced unfolding of fibronectin type 3 modules. An analysis by biased molecular dynamics simulations. *J Mol. Biol.* **1999**, *288*, 441−459.
(19) Jensen, M. O.; Park, S.; Tajkhorshid, E.; Schulten, K. Energetics of glycerol conduction through aquaglyceroporin Glpf. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 6731−6736.
(20) Jarzynski, C. Nonequilibrium equality for free energy differences. *Phys. Rev. Lett.* **1997**, *78*, 2690−2693.
(21) Jarzynski, C. Equilibrium free energy differences from nonequilibrium measurements. A master-equation approach. *Phys. Rev. E* **1997**, *56*, 5018−5035.
(22) Born, M. Volumen und Hydratationswärme der Ionen. *Z. Phys.* **1920**, *1*, 45−48.
(23) Zwanzig, R. W. High-temperature equation of state by a perturbation method. I. Nonpolar gases. *J. Chem. Phys.* **1954**, *22*, 1420−1426.
(24) Anderson, P. W. A mathematical model for the narrowing of spectral lines by exchange or motion. *J Phys. Soc. Jpn.* **1954**, *9*, 316−339.
(25) Roepstorff, G. *Path Integral Approach to Quantum Physics*; Springer: Berlin, Germany, 1994.

(26) Hummer, G.; Szabo, A. Free energy reconstruction from nonequilibrium single-molecule pulling experiments. *Proc. Natl. Acad. Sci. U.S.A.* **2001**, *98*, 3658−3661.
(27) Hummer, G.; Szabo, A. Kinetics from nonequilibrium single-molecule pulling experiments. *Biophys. J.* **2003**, *85*, *5*−*15*.
(28) Bolhuis, P. G.; Chandler, D.; Dellago, C.; Geissler, P. L. Transition path sampling. Throwing ropes over rough mountain passes in the dark. *Annu. Rev. Phys. Chem.* **2002**, *53*, 291−318.
(29) Yeh, I.-C.; Hummer, G. Peptide loop-closure kinetics from microsecond molecular dynamics simulations in explicit solvent. *J. Am. Chem. Soc.* **2002**, *124*, 6563−6568.
(30) Lapidus, L. J.; Eaton, W. A.; Hofrichter, J. Measuring the rate of intramolecular contact formation in polypeptides. *Proc. Natl. Acad. Sci. U.S.A.* **2000**, *97*, 7220−7225.
(31) Hummer, G. Fast-growth thermodynamic integration. Error and efficiency analysis. *J. Chem. Phys.* **2001**, *114*, 7330−7337.
(32) Gullingsrud, J. R.; Braun, R.; Schulten, K. Reconstructing potentials of mean force through time series analysis of steered molecular dynamics simulations. *J. Comput. Phys.* **1999**, *151*, 190−211.
(33) Crooks, G. E. Path-ensemble averages in systems driven far from equilibrium. *Phys. Rev. E* **2000**, *61*, 2361−2366.
(34) The Riemann sum corresponding to the force integral is $\int_C F\,dq \approx \sum_{i=1}^{N}(q_i - q_{i-1})(F_i + F_{i-1})/2$, where i labels positions consecutive in time, $q_i = q(t_i)$, and $F_i = F(q_i,t_i)$, with $t_0 = 0$ and $t_N = t$. In general, F is a multivalued function of q, and the sum contains both positive and negative contributions.
(35) Schurr, J. M.; Fujimoto, B. S. Equalities for the nonequilibrium work transferred from an external potential to a molecular system. Analysis of single-molecule extension experiment. *J. Phys. Chem. B* **2003**, *107*, 14007−14019.
(36) Ferrenberg, A. M.; Swendsen, R. H. Optimized Monte Carlo data analysis. *Phys. Rev. Lett.* **1989**, *63*, 1195−1198.
(37) Wood, R. H. Estimation of errors in free energy calculations due to the lag between the Hamiltonian and the system configuration. *J. Phys. Chem.* **1991**, *95*, 4838−4842.
(38) Bell, G. I. Models for the specific adhesion of cells to cells. *Science* **1978**, *200*, 618−627.
(39) Evans, E.; Ritchie, K. Dynamic strength of molecular adhesion bonds. *Biophys. J.* **1997**, *72*, 1541−1555.
(40) Evans, E.; Berk, D.; Leung, A. Detachment of agglutinin-bonded red blood cells. I. Forces to rupture molecular-point attachments. *Biophys. J.* **1991**, *59*, 838−848.
(41) Rief, M.; Fernandez, J. M.; Gaub, H. E. Elastically coupled two-level systems as a model for biopolymer extensibility. *Phys. Rev. Lett.* **1998**, *81*, 4764−4767.
(42) Schuler, B.; Lipmann, E.; Eaton, W. A. Probing the free-energy surface for protein folding with single-molecule fluorescence spectroscopy. *Nature (London)* **2002**, *419*, 743−747.

AR040148D