Free energy profiles from single-molecule pulling experiments

Gerhard Hummer[1] and Attila Szabo[1]

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

Contributed by Attila Szabo, October 19, 2010 (sent for review August 16, 2010)

Nonequilibrium pulling experiments provide detailed information about the thermodynamic and kinetic properties of molecules. We show that unperturbed free energy profiles as a function of molecular extension can be obtained rigorously from such experiments without using work-weighted position histograms. An inverse Weierstrass transform is used to relate the system free energy obtained from the Jarzynski equality directly to the underlying molecular free energy surface. An accurate approximation for the free energy surface is obtained by using the method of steepest descent to evaluate the inverse transform. The formalism is applied to simulated data obtained from a kinetic model of RNA folding, in which the dynamics consists of jumping between linker-dominated folded and unfolded free energy surfaces.

atomic force microscope | force spectroscopy | free energy calculation | nonequilibrium work relations | steered molecular dynamics

Mechanical forces exerted on single molecules by laser optical tweezers or atomic force microscopes can induce structural transitions such as the unfolding of a protein or nucleic acid or the dissociation of a complex. In these experiments, the molecular system can be driven far from equilibrium, permitting the exploration of metastable states and rare molecular processes. Nonequilibrium work relations, most notably the Jarzynski equality (1) and the Crooks fluctuation theorem (2), form the basis for the rigorous extraction of equilibrium thermodynamic information from nonequilibrium single-molecule pulling experiments (3–9). The reconstruction of molecular free energy surfaces was made possible by extending the Jarzynski equality from a relation for the system free energy that depends on an experimentally controllable parameter to a relation for a molecular free energy surface that depends on a fluctuating coordinate (3). Here we show that, for a harmonic pulling apparatus, the molecular surface can be directly obtained from an inverse Weierstrass transform of the system free energy. With the help of a Hubbard–Stratonovich transformation of this formal relation, we derive an accurate approximation for the free energy surface that is exact if the surface is locally harmonic with positive curvature.

Single-molecule pulling experiments can be described by a time-dependent Hamiltonian $\mathscr{H}_0(\mathbf{x}) + V[\mathbf{x},z(t)]$ in which the Hamiltonian $\mathscr{H}_0(\mathbf{x})$ of the molecular system (where $\mathbf{x}$ is a phase-space coordinate including the solvent) is coupled to a pulling apparatus whose anchoring position moves along a prescribed path $z(t)$ as a function of time t (Fig. 1) (3). For a harmonic spring attached to the molecular system at an extension $q(\mathbf{x})$, the spring Hamiltonian is $V[\mathbf{x},z(t)] = k[q(\mathbf{x}) - z(t)]^2/2$, where k is the spring constant. The difference $A(z)$ between the Helmholtz free energies of the systems at control parameters $z(t)$ and $z(0)$ is given by the Jarzynski equality (1)

$$e^{-\beta A(z)} \equiv \frac{\int d\mathbf{x} e^{-\beta\mathscr{H}_0(\mathbf{x})-\beta V[\mathbf{x},z(t)]}}{\int d\mathbf{x} e^{-\beta\mathscr{H}_0(\mathbf{x})-\beta V[\mathbf{x},z(0)]}} = \langle e^{-\beta W(z)}\rangle, \qquad [1]$$

where the average is over all trajectories $\mathbf{x}(t)$ starting from an equilibrium Boltzmann distribution corresponding to Hamiltonian $\mathscr{H}[\mathbf{x},z(0)]$, and $\beta^{-1} = k_B T$ with k_B Boltzmann's constant and T the absolute temperature. Jarzynski's work is

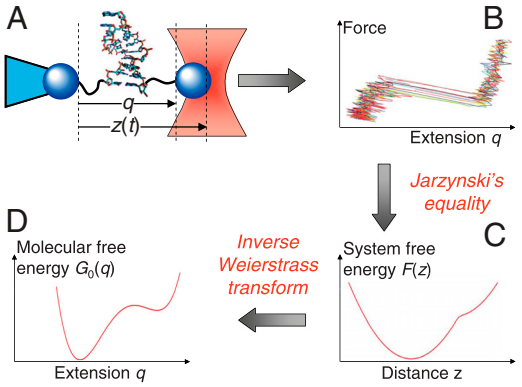

Fig. 1. Schematic of a single-molecule pulling experiment and its analysis. (*A*) A laser optical tweezer traps a bead in its focus (red). As the focus is moved, a force is exerted on a molecule, such as an RNA hairpin, attached to the moving bead. The arrows indicate the experimentally controlled distance $z(t)$ between the anchor and the center of the laser focus and the fluctuating molecular extension q, which includes contributions from the flexible linkers and the fixed bead radii (black lines). (*B*) Measured forces as a function of extension q from different nonequilibrium pulling experiments. Jumps in extension, associated with a slight drop in force, indicate unfolding of the molecule. (*C*) Nonequilibrium force-extension curves transformed into equilibrium system free energies $A(z)$ as a function of the experimentally controlled distance z by using the Jarzynski equality (1). (*D*) System free energy transformed into the molecular free energy surface $G_0(q)$ by using an inverse Weierstrass transform.

$$W(z) \equiv W[z = z(t)] = \int_0^t \frac{\partial V}{\partial z}\frac{dz}{dt'}dt' = \int_{z(0)}^{z(t)} F dz. \qquad [2]$$

The last identity in terms of a force-distance integral holds for pulling with a spring, $V = V[q - z(t)]$, with $F = -\partial V/\partial q = \partial V/\partial z$ the measured pulling force. Note that the integration here is along the position $z(t)$ of the spring, not the molecular extension q.

Some time ago, we recognized that the Feynman–Kac theorem of quantum mechanics can be used to reduce the problem of evaluating the path average on the right-hand side of Eq. **1** into the solution of a Schrödinger-like equation (3). Remarkably, this equation could be trivially solved and the Jarzynski equality established. This insight immediately allowed us to show that the unperturbed free energy profile $G_0(q)$, up to an arbitrary constant, can be expressed in terms of position histograms reweighted by the nonequilibrium work (3):

Author contributions: G.H. and A.S. designed research, performed research, analyzed data, and wrote the paper.

The authors declare no conflict of interest.

[1]To whom correspondence may be addressed. E-mail: gerhard.hummer@nih.gov or attilas@nih.gov.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1015661107/-/DCSupplemental.

$$\exp[-\beta G_0(q)] \equiv \int \delta[q - q(\mathbf{x})] \exp[-\beta \mathcal{H}_0(\mathbf{x})] d\mathbf{x}$$
$$= \langle \delta\{q - q[\mathbf{x}(t)]\} e^{-\beta W(z) + \beta V[q(t), z(t)]} \rangle$$
$$= \langle \delta\{q - q[\mathbf{x}(t)]\} e^{-\beta \int F dq + \beta V[q(0), z(0)]} \rangle, \quad [3]$$

where $\delta(\ldots)$ is the Dirac delta function. The last line involves the more conventional definition of work, $\int F dq$, where F is the pulling force, and the integral is along the contour of the trajectories of the fluctuating molecular extension $q(t)$. By using Eq. **3**, such nonequilibrium trajectories $q(t)$ can be transformed into free energy profiles $G_0(q)$ through work-weighted position distributions (by going directly from B to D in Fig. 1). In practice, this transformation can be accomplished by using a generalization (3) of the weighted-histogram method used in equilibrium statistical mechanics (10).

In this paper, we introduce an alternate simpler approach that avoids the position histograms of Eq. **3**. Specifically, we obtain $G_0(q)$ in the important special case that $V[\mathbf{x}, z(t)] = k[q(\mathbf{x}) - z(t)]^2/2$, where k is the spring constant of a harmonic pulling spring moving along $z(t)$. We first derive a remarkably simple exact relation expressing $G_0(q)$ in terms of $A(z)$. Starting with this formal result, we then obtain an approximate analytic expression for $G_0(q)$ that involves only $A(z)$ and its first two derivatives. The free energy $A(z)$ can be obtained from the Jarzynski equality in Eq. **1** or from the Crooks–Minh–Adib relation (2, 11) when both forward and reverse trajectories are available.

Although this procedure for extracting the intrinsic free energy surface has a wider range of validity than the related "stiff-spring" approximation (12) and is considerably easier to use than our original work-weighted-histogram method (3), it is not exact. In particular, barriers sharper than the curvature of the pulling spring cannot be fully resolved. So, in practice, if the curvature of the reconstructed barrier turns out to be similar to that of the spring, $-G_0''(q) \approx k$, one should implement the more complicated work-weighted-histogram method, which may require running additional experiments to improve the statistics.

Theory

Consider the following manipulations of the definition (up to a constant of dimension length) of $A(z)$ in terms of the one-dimensional free energy surface $G_0(q)$:

$$e^{-\beta A(z)} = \int dq e^{-\beta G_0(q) - \beta k(q-z)^2/2} = \int dq e^{-\beta k q^2/2 - \beta G_0(q+z)}$$
$$= \int dq e^{-\beta k q^2/2} e^{q\frac{d}{dz}} e^{-\beta G_0(z)} = (2\pi/\beta k)^{1/2} e^{\frac{1}{2\beta k}\frac{d^2}{dz^2}} e^{-\beta G_0(z)}, \quad [4]$$

where in succession we changed the variables of integration, introduced the translation operator $\exp(qd/dz)g(z) = g(q+z)$, and used $\int dq \exp(-aq^2/2 + bq) = (2\pi/a)^{1/2} \exp(b^2/2a)$, where b is the differential operator. Eq. **4** is not unexpected if one notices that $\exp[-\beta k(q-z)^2/2]$ has the same structure as the Green's function or propagator of the free diffusion equation or the imaginary-time Schrödinger equation for a free particle and that this Green's function can be formally written as the exponential of the second-derivative operator.

The last line of Eq. **4** can be formally inverted to yield

$$e^{-\beta G_0(q=z)} = (2\pi\epsilon)^{-1/2} e^{-\frac{\epsilon}{2}\frac{d^2}{dz^2}} e^{-\beta A(z)}, \quad [5]$$

where we have defined $\epsilon = (\beta k)^{-1}$. This equation, together with the Jarzynski identity, determines the free energy surface at a point $q = z$ in terms of all derivatives of $\langle \exp[-\beta W(z)] \rangle$ at z. In the mathematical literature, the first line of Eq. **4** defines a Weierstrass transform, in which $\exp[-\beta G_0(q)]$ is convoluted with a Gaussian to obtain $\exp[-\beta A(z)]$, and Eq. **5** is its inverse (13). This equation can be generalized to a certain class of anharmonic springs by using the Hirschman–Widder convolution transform and its inverse (13).

If one expands the exponential of the differential operator in Eq. **5** and then takes the logarithm on both sides, one obtains to order ϵ^2 (and up to a constant)

$$G_0(q = z) = A(z) + \frac{\epsilon}{2}(\beta \dot{A}^2 - \ddot{A})$$
$$+ \frac{\epsilon^2}{8}(\dddot{A} + 4\beta^2 \ddot{A}\dot{A}^2 - 4\beta \dddot{A}\dot{A} - 2\beta \ddot{A}^2), \quad [6]$$

where $\dot{A} \equiv dA/dz$, and $A(z)$ and its derivatives are evaluated at z. The first two terms constitute the stiff-spring approximation of Park et al. (12), which is accurate if the spring constant is much larger than the maximum of $|G_0''(q)|$. The third term has little practical value because it involves higher derivatives of $A(z)$, which are difficult to obtain accurately in computer simulations or pulling experiments.

We now obtain an alternate expression for the free energy surface $G_0(q)$ that contains only $A(z)$ and its first and second derivatives but has a wider range of validity than the stiff-spring approximation. Our strategy is essentially the same as that used to show that mean-field theory of critical phenomena is exact for an infinite-range spin model (14) (i.e., the partition function is transformed into an integral that is evaluated by the method of steepest descent). By using the Hubbard–Stratonovich transform of the exponential operator

$$e^{-\frac{\epsilon}{2}\frac{d^2}{dz^2}} = (2\pi\epsilon)^{-1/2} \int d\zeta e^{-\zeta^2/2\epsilon} e^{i\zeta\frac{d}{dz}} \quad [7]$$

and $\exp(i\zeta d/dz)g(z) = g(z + i\zeta)$, Eq. **5** becomes

$$e^{-\beta G_0(z)} = (2\pi\epsilon)^{-1} \int d\zeta e^{-\zeta^2/2\epsilon - \beta A(z + i\zeta)}, \quad [8]$$

which is an alternate inversion formula (13). To evaluate the integral in the auxiliary variable ζ, we expand the argument of the exponential to second order in ζ about a stationary point $\zeta_0 = -i\epsilon\beta\dot{A}(z + i\zeta_0)$ defined by $\partial[\zeta^2/2\epsilon + \beta A(z + i\zeta)]/\partial\zeta|_{\zeta=\zeta_0} = 0$. After integrating the Gaussian, we obtain

$$e^{-\beta G_0(z)} \approx [2\pi\epsilon(1 - \epsilon\beta\ddot{A})]^{-1/2} e^{-\beta A + \epsilon\beta^2\dot{A}^2/2}, \quad [9]$$

where $A, \dot{A}$, and $\ddot{A}$ are evaluated at $z + i\zeta_0$. Thus it appears that one must evaluate the free energy and its derivatives at complex values of the extensions. However, ζ_0 turns out to be purely imaginary, which can be seen by rewriting the equation that determines ζ_0 as $z = z' - \epsilon\beta\dot{A}(z')$, where $z' = z + i\zeta_0$. We then obtain G_0 at a location that is shifted from the instantaneous minimum of the moving harmonic potential in terms of $A(z')$ and its derivatives at real z'. After taking the logarithm on both sides of Eq. **9**, we find within a constant

$$G_0\left(q = z - \frac{\dot{A}(z)}{k}\right) \approx A(z) - \frac{\dot{A}(z)^2}{2k} + \frac{1}{2\beta}\ln\left(1 - \frac{\ddot{A}(z)}{k}\right), \quad [10]$$

where we have dropped the primes.

Thus to find the free energy profile $G_0(q)$, we need $A(z)$ and its first and second derivatives. The derivatives can be obtained by numerical differentiation or directly from the work-weighted trajectory averages:

$$\dot{A}(z) = -k\langle\langle q - z \rangle\rangle = \langle\langle F \rangle\rangle, \quad [11]$$

$$1 - \ddot{A}(z)/k = \beta(\langle\langle F^2 \rangle\rangle - \langle\langle F \rangle\rangle^2)/k. \quad [12]$$

If only forward trajectories are available, the equilibrium averages at fixed z can be evaluated as $\langle\langle\ldots\rangle\rangle = \langle(\ldots)\exp[-\beta W(z)]\rangle / \langle\exp[-\beta W(z)]\rangle$; if reverse trajectories are also available, Eq. **6** of ref. 11 can be used. Eq. **12** shows that the argument of the logarithm in Eq. **10** is always positive.

Because Eq. **10** is exact if $G_0(q)$ is locally quadratic with positive curvature, "quasi-harmonic" seems an appropriate term to describe this approximation. It can derived in a less elegant way by approximately evaluating the integral in the first line of Eq. **4** and its derivatives (15). This approximation is closely related to the one we proposed (6) on the basis of simply assuming that the distribution of $\langle\langle\delta\{q - q[\mathbf{x}(t)]\}\rangle\rangle$ is Gaussian.

Results and Discussion

Bistable Molecular Free Energy Surface. We now investigate the range of validity of the quasi-harmonic approximation. Fig. 2 compares different reconstructions of a bistable molecular free energy surface $G_0(q)$ from corresponding $A(z)$ curves. Reconstructions are shown for different spring constants k of the pulling spring obtained with the stiff-spring approximation (Eq. **6** truncated after the second term) and the quasi-harmonic approximation (Eq. **10**). As expected, when the spring is very stiff, larger than the curvatures at the well bottom and barrier top, both approximations work well. As the pulling spring is softened (Fig. 2, *Middle* and *Bottom*), the stiff-spring approximation rapidly deteriorates. In contrast, the quasi-harmonic approximation recovers the free energy surface, with only small deviations near the barrier. Of course, if the force constant k of the pulling spring is further reduced, the quasi-harmonic approximation will break down in the barrier region, and in such cases the work-weighted-histogram method (3) should be used. Nevertheless, we expect that the quasi-harmonic approximation should work not only in the limit of stiff pulling springs (such as those of atomic force microscopes) but also for softer springs (such as those in laser optical tweezers). As a test, we now apply the quasi-harmonic approximation to reconstruct the free energy surface for RNA folding, as probed with laser tweezers.

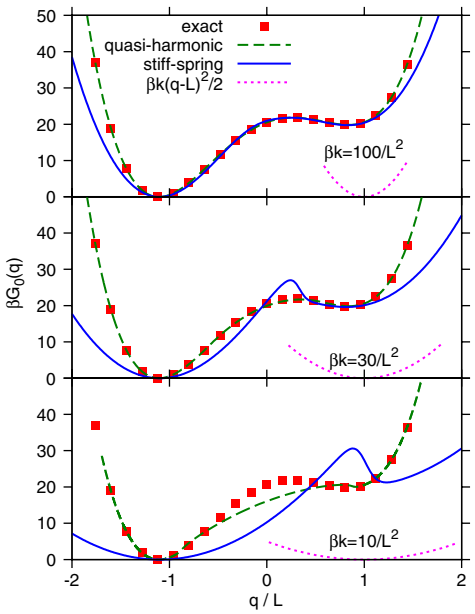

Fig. 2. Free energy reconstruction by using three pulling spring constants $\beta kL^2 = 100$ (*Top*), 30 (*Middle*), and 10 (*Bottom*), with L the unit of length. Filled squares: Free energy surface $\beta G_0(q) = 10[(q/L)^2 - 1]^2 + 10q/L + 10$; green dashed line: quasi-harmonic approximation, Eq. **10**; solid blue line: stiff-spring approximation (12), Eq. **6** truncated after the second term; purple dotted line: pulling potential $\beta k(q - L)^2/2$.

Forced Unfolding of RNA with Soft Tweezers. The above formalism allows one to extract the free energy $G_0(q)$ as a function of the extension q of the system. In many practical applications, the extension q will include not only the molecule of interest, but also connecting linkers through which it is attached to the pulling apparatus. Ideally, one would like to know the free energy $G_M(x)$ of the molecule as a function of its end-to-end extension x. For linkers that do not interact strongly with the molecule of interest, except through the covalent attachment, G_M is related to G_0 and the free energy G_L of the linker by (6)

$$e^{-\beta G_0(q)} = \int dx e^{-\beta G_M(x)} e^{-\beta G_L(q-x)}. \qquad [13]$$

Now, by measuring the free energy of first the linker and then the entire construct, one could obtain $G_M(x)$ after deconvolution (8, 16). However, this procedure becomes increasingly difficult and may even be virtually impossible when the linkers are too soft.

Here we present an illustrative application of our formalism to such a case, where $G_0(q)$ is completely determined by linkers and the only information it contains about the molecular surface are differences in free energy and extension between intact and ruptured states. We will first construct a simple model of the classic experiments of Liphardt et al. (17) on the forced unfolding of RNA. The parameters of this model will be obtained from experiment (see *Methods*), and then simulated data will be analyzed by using our formalism. We will make the simplifying assumption that the linkers are harmonic with force constants determined from the slope of the experimental force-extension curves near rupture. It turns out that in the relevant regime from about 5 to 20 pN force, with the unfolding transition at ≈15 pN, these curves are linear (Fig. 2*A* in ref. 17) and the harmonic approximation is actually better than what one might first think.

We shall assume that the system diffuses on the multidimensional free energy surface $G = G_M(x,Q) + \frac{1}{2}k_L(x-q)^2 + \frac{1}{2}k[q - z(t)]^2$, where x is the RNA extension and q is the extension of the whole construct. The RNA free energy surface explicitly includes some other molecular reaction coordinate Q that captures the difference between intact and ruptured states (e.g., the fraction of native base pairs) (18, 19) because x is not always a good reaction coordinate. k_L is the effective force constant of the linkers near rupture. For simplicity, k_L is assumed to be unchanged upon unfolding, which is reasonable for long linkers where the released contour length has a negligible effect (see, e.g., Fig. 2*A* of ref. 17). The dynamics of Q, x, and q are described by diffusion coefficients D_Q, D_x, and D_q, respectively. The dynamics of q is normally dominated by the diffusive motion of the optically trapped bead used for pulling (or the tip of the atomic force microscope). Typically, the bead dynamics is much slower than the intrabasin dynamics of the molecule but faster than interbasin dynamics associated with molecular unfolding and refolding. Under this condition, the multidimensional problem can be reduced to coupled reaction-diffusion equations (20) for the probabilities $p_1(q,t)$ and $p_2(q,t)$ of the folded and unfolded states, respectively, at total extension q and time t:

$$\frac{\partial p_1}{\partial t} = -k_1(q)p_1 + k_2(q)p_2 + D_q \frac{\partial}{\partial q} e^{-\beta V_1} \frac{\partial}{\partial q} e^{\beta V_1} p_1, \qquad [14]$$

$$\frac{\partial p_2}{\partial t} = k_1(q)p_1 - k_2(q)p_2 + D_q \frac{\partial}{\partial q} e^{-\beta V_2} \frac{\partial}{\partial q} e^{\beta V_2} p_2, \qquad [15]$$

where $V_i(q,t)$ is the potential of mean force (PMF) along q in state i, $\exp(-\beta V_i) \propto \int_i dQdx \exp(-\beta G)$, and the integral is over the configuration space of state i. The "surface hopping" rate $k_1(q)$ $[k_2(q)]$ is the rate of unfolding (folding) at extension q. For the sake of simplicity, we assume that the folded and unfolded mole-

cules are much stiffer than the linkers; then, to a good approximation, $V_i(q,t) = \frac{1}{2}k_L(x_i - q)^2 + \frac{1}{2}k[q - z(t)]^2 + G_i^0$. Here G_i^0 and x_i are the free energy and extension of state i, respectively. More generally, the extensions x_i can depend on q, particularly for the unfolded state. Although this dependence can be treated explicitly in a straightforward way, here it is handled implicitly by determining x_i from a fit to experimental data (see *Methods*).

It should be emphasized that in this model, apart from x_i and ΔG_i^0, all information about the free energy surface of the RNA, $G_M(x,Q)$, is packed into the rate constants $k_i(q)$ that describe surface hopping. These rates can, and in fact should, be calculated by applying the Kramers–Langer theory to the potential $G_M(x,Q) + \frac{1}{2}k_Lx^2 - k_Lxq \approx G_M(x,Q) - k_Lxq$ for fixed q. Although such calculations have been done for simple one-dimensional (21) and two-dimensional (22) potentials (when $F = k_Lq$), here as a first step we will approximate the rates by the simple transition state theory-like formula $k_i(q) = k_i^0 \exp\{-\frac{\beta}{2}k_L[(x_\cap - q)^2 - (x_i - q)^2]\}$, where $x_\cap$ is the location of the barrier separating x_1 and x_2. The intrinsic rates are defined so that $k_1^0/k_2^0 = \exp[-\beta(G_2^0 - G_1^0)] \equiv \exp(-\beta\Delta G^0)$, with ΔG^0 the unfolding free energy. The resulting kinetic model is similar to that developed by Manosas and Ritort (23).

In free energy reconstruction experiments, the objective is to determine the PMF $G_0(q)$ in the absence of pulling ($k = 0$). In our model, $G_0(q)$ satisfies

$$e^{-\beta G_0(q)} = e^{-\beta G_1(q)} + e^{-\beta G_2(q)}, \qquad [16]$$

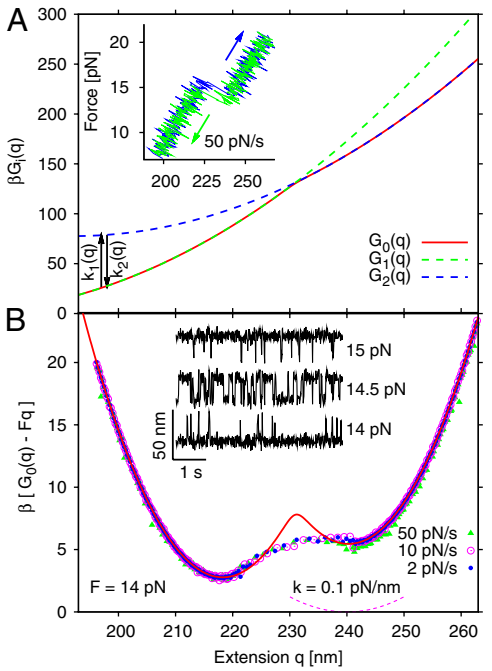

Fig. 3. Unfolding of RNA. (*A*) Free energy surfaces $G_1(q)$ (green) and $G_2(q)$ (blue) of the RNA-linker construct and the corresponding PMF $G_0(q)$ (red). The *Inset* shows the force-extension curves of two simulated pulling trajectories, a forward trajectory in the unfolding direction (blue) and a reverse trajectory in the folding direction (green). (*B*) Free energy surfaces reconstructed from 100 trajectories each in the forward and reverse directions at force loading rates of $R = 2$ (filled circles), 10 (open circles), and 50 pN/s (triangles), analyzed with the quasi-harmonic approximation to the inverse Weierstrass transform Eq. **10**. The solid curve is the exact $G_0(q)$. For clarity, an artificial force of $F = 14$ pN is imposed on all curves. The dashed curve at the bottom right indicates the pulling potential with a spring constant of $k = 0.1$ pN/nm. The *Inset* shows three extension-versus-time trajectories obtained at constant forces of 14, 14.5, and 15 pN.

where $G_1(q) = \frac{1}{2}k_L(x_1 - q)^2$ and $G_2(q) = \frac{1}{2}k_L(x_2 - q)^2 + \Delta G^0$ are the free energy surfaces of the folded and unfolded state, respectively, as shown in Fig. 3*A*. These surfaces are harmonic because of our simplifying assumption that the linker is harmonic. Note that $G_0(q)$ is not double-welled because it is dominated by the linker. We simulated RNA pulling according to Eqs. **14** and **15** at three different speeds typical of the experiments (4, 5, 17), with force loading rates ranging from $R = 2$ to 50 pN/s. One hundred trajectories each were generated in the forward (unfolding) and reverse (folding) directions, with two typical trajectories shown in the *Inset* in Fig. 3*A*. Although $A(z)$ can be found analytically in this case, we calculated this free energy from all the trajectories by using the Minh–Adib formalism (11), which is based on the Crooks fluctuation theorem (2) (see *Methods*). The required derivatives $\dot{A}(z)$ and $\ddot{A}(z)$ were obtained by numerical differentiation. We then used these derivatives in the quasi-harmonic approximation, Eq. **10**, to estimate the PMF $G_0(q)$. If the resulting curve were plotted on the same scale as Fig. 3*A*, differences would not be clearly visible. Therefore in Fig. 3*B* we artificially impose a constant force of $F = 14$ pN and plot $G_0(q) - Fq$. We find that the reconstructed PMFs generally agree well with the exact $G_0(q)$, even at the highest pulling speed of 50 pN/s. The stiff-spring approximation is so poor in this case that it is not shown. However, the barrier is underestimated by the quasi-harmonic approximation because the absolute value of the curvature of $G_0(q)$ at the barrier is greater than the force constant of the pulling spring. Not fully resolving sharp barriers is the major limitation of this method and is the price one has to pay for its simplicity. The more sophisticated work-weighted-histogram method (3) is exact and limited only by inadequate sampling. In our RNA unfolding example, the histogram method describes the barrier accurately by using the same data (see *SI Text* and Fig. S1*B*). However, for stiff springs the quasi-harmonic approximation performs better than the histogram method (Fig. S1*A*). The two methods are thus complementary and, in practice, should be implemented together.

We also simulated the hopping dynamics described by Eqs. **14** and **15** in the presence of a constant force by replacing $\frac{1}{2}k[q - z(t)]^2$ with $-Fq$ in the free energy surfaces V_i. The resulting trajectories of extension versus time (Fig. 3*B*, *Inset*) are very similar to those observed experimentally (Fig. 2*C* of ref. 17), thus confirming that our model did capture the essential physics in a quantitative manner. If these trajectories were analyzed with the equilibrium method of Woodside et al. (8), the resulting free energy would be $G_0(q) - Fq$. Although these surfaces do have a significant barrier for $F \approx 14.5$ pN, it is important to emphasize that the dynamics is not simply diffusive barrier crossing. Indeed, diffusion on this surface with $D_q = 250$ nm^2/ms would produce incorrect rates of folding and unfolding. Instead, the folding and unfolding occurs because of surface hopping governed by the q-dependent rates, $k_i(q)$. However, after a transition has occurred, the linker extension and bead position do relax because of diffusion on the appropriate q-dependent free energy surfaces. So, although in this example $G_0(q)$ does not contain information about the molecular surface other than ΔG^0 and $x_2 - x_1$, the force-dependent kinetics does because the hopping rates are determined by the dynamics on the molecular surface (18).

Concluding Remarks

We have shown that the system free energies $A(z)$ obtained from single-molecule pulling experiments can be converted rigorously into molecular free energy surfaces $G_0(q)$ with the help of an inverse Weierstrass transform, when the pulling spring is harmonic. By first applying a Hubbard–Stratonovich transform to the formal relation between $A(z)$ and $G_0(q)$ and approximating the resulting integral by the method of steepest descent, we derived a simple practical approximation for $G_0(q)$ in terms of $A(z)$, $\dot{A}(z)$, and $\ddot{A}(z)$. We showed that this quasi-harmonic formula performs

better than the stiff-spring approximation (12), which requires the same input and is widely used in the analysis of steered molecular dynamics simulations. Nevertheless, in contrast to the work-weighted-histogram method (3), the quasi-harmonic formalism is approximate and cannot fully resolve barriers that are sharper than the curvature of the pulling spring. Our work highlights the connections between nonequilibrium work relations and inverse integral transforms, and leads to a simple procedure to extract free energy profiles from single-molecule force spectroscopy.

Finally, we cannot resist noting that the problem of obtaining the molecular free energy surface $G_0(q)$ from the system free energy $A(z)$ is closely related to that of reconstructing an image blurred by Gaussian noise. As can be seen from Eq. **4**, $\exp[-\beta A(z)]$ is the convolution of the Boltzmann distribution $\exp[-\beta G_0(q)]$ with a Gaussian of variance βk. The positive definite $\exp[-\beta G_0(q)]$ can be thought of as the intensity of an original one-dimensional image and $\exp[-\beta A(z)]$ as the Gaussian blurred image. Image reconstruction is a highly challenging inverse problem, for which numerous procedures have been developed. We found that the Lucy–Richardson method (24, 25), widely used in astronomy, works well also for free energy reconstruction as long as the free energy range is not too large (because one effectively inverts populations) and the pulling spring is not too soft. In the notation of this paper, the Lucy–Richardson algorithm can be written as

$$G_0^{(n+1)}(q) = G_0^{(n)}(q) - \beta^{-1} \ln \int dz \frac{e^{-\beta A(z) - \beta k(q-z)^2/2}}{\int dq' e^{-\beta G_0^{(n)}(q') - \beta k(q'-z)^2/2}}, \quad [17]$$

where $G_0^{(n)}(q)$ is the nth iteration of $G_0(q)$, starting from $G_0^{(0)}(q) = 1$. An advantage of this procedure is that the noise need not be Gaussian, and the Lucy–Richardson method can thus be adapted to deconvolve Eq. **4** when the linker has a nonlinear force-extension curve, such as that of a worm-like chain.

Methods

The nonequilibrium unfolding of the P5ab RNA hairpin at 10 mM Mg^{2+} concentration (17) was simulated at different force loading rates of $R = 2$, 10, and 50 pN/s. The parameters of the model were chosen as follows. The force constant of the trap was $k = 0.1$ pN/nm (4). The diffusion coefficient of the trapped bead was obtained from the Stokes–Einstein formula for a ≈1-μm bead in water, $D_q = 250$ nm^2/ms. The linker force constant of $k_L = 0.29$ pN/nm and the parameters $x_1 = 170$ nm and $x_2 = 192$ nm were obtained from the slopes and the intercepts at $F = 0$, respectively, of linear fits to the experimental force-extension curves before and after rupture in Fig. 2*A* of ref. 17. Note that, because the actual linker is not harmonic at all extensions, the x_i are only fitting parameters, but their difference $x_2 - x_1$ is the gain in length upon RNA unfolding at forces near 15 pN. The remaining parameters were obtained by noting (*i*) that in the presence of a constant force F, the effective folding and unfolding rates of our model are given by Bell's formula (i.e., $\langle k_i(q)\rangle = \int dq k_i(q) \exp[-\beta G_i(q) + \beta F q] / \int dq \exp[-\beta G_i(q) + \beta F q] = k_i^0 \exp[\beta F(x_\cap - x_i)]$) and (*ii*) that this formula was used to fit the experimental rates over a limited force range (17) (see Fig. 2 and Table 1 of ref. 17). In this way, we find that $x_\cap - x_1 = 11.9$ nm, $k_1^0 = e^{-39}$ s^{-1}, and $\Delta G^0 = 193$ kJ/mol.

We simulated trajectories corresponding to Eqs. **14** and **16** by using the following algorithm. The positions $q(n\Delta t) \equiv q_n$ are determined by $q_{n+1} = q_n - D_q \Delta t \beta \partial V_i / \partial q_n + (2 D_q \Delta t)^{1/2} g_n$, where g_n are uncorrelated Gaussian random numbers with zero mean and unit variance. The trajectory then stays on surface i with probability $p = \exp\{-[k_i(q_n) + k_i(q_{n+1})]\Delta t/2\}$ and jumps to the other surface with probability $1 - p$. The time step was $\Delta t = 10^{-4}$ ms. The beginning and end points of the trajectories had spring locations $z_0 = 270$ nm and $z_1 = 470$ nm, respectively. On the forward path $z(t) = z_0 + vt$, $0 \le t \le \tau$, the system is first equilibrated with the spring at z_0, which is then moved from z_0 to z_1 at a constant velocity $v = (k_L^{-1} + k^{-1})R$ over a time $\tau = (z_1 - z_0)/v$. On the reverse path $\underline{z}(t) = z_1 - vt$, the system is equilibrated with the spring at z_1, and the spring is moved from z_1 to z_0 at a velocity $-v$.

If only forward trajectories are available, the free energy difference $A(z)$ between systems with spring locations z and z_0 can be calculated in a straightforward way by using the Jarzynski equality (1) $\exp[-\beta A(z)] = \langle \exp[-\beta W(z)] \rangle_f$, where $W(z) = \int_{z_0}^{z} F(z') dz'$ and $F(z) = k[z - q(z)]$. Although for the present model, with 100 trajectories, this procedure is satisfactory, we here combine the results from forward and reverse trajectories by using the Minh–Adib formalism (11). The work along the reverse paths is defined as $\underline{W}(z) = \int_{z_1}^{z} F(z') dz'$. We first determine the free energy difference $\Delta A \equiv A(z_1)$ between spring locations z_1 and z_0 by numerically solving the Bennett acceptance ratio formula (2, 26):

$$\left\langle \frac{1}{n_r + n_f e^{\beta[W(z_1) - \Delta A]}} \right\rangle_f = \left\langle \frac{1}{n_f + n_r e^{\beta[\underline{W}(z_0) + \Delta A]}} \right\rangle_r, \quad [18]$$

for ΔA with the Newton–Raphson method. The averages are over $n_f = 100$ and $n_r = 100$ forward and reverse trajectories, respectively. With this ΔA, we then use the formalism of Minh and Adib (11) to obtain $A(z)$ for intermediate values of z, $z_0 \le z \le z_1$:

$$e^{-\beta A(z)} = \left\langle \frac{n_f e^{-\beta W(z)}}{n_f + n_r e^{-\beta[W(z_1) - \Delta A]}} \right\rangle_f + \left\langle \frac{n_r e^{-\beta[\underline{W}(z) + \Delta A]}}{n_r + n_f e^{-\beta[\underline{W}(z_0) + \Delta A]}} \right\rangle_r. \quad [19]$$

Trajectory data were recorded at z steps of 1 nm, at experimentally achievable frequencies of less than 1 kHz. From the resulting $A(z)$, the first and second derivatives were obtained by using centered differences.

ACKNOWLEDGMENTS. We thank Prof. Olga Dudko and Drs. D. Minh and J. Chodera for discussions. This work was supported by the Intramural Research Program of the National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health.

1. Jarzynski C (1997) Nonequilibrium equality for free energy differences. *Phys Rev Lett* 78:2690–2693.
2. Crooks GE (2000) Path-ensemble averages in systems driven far from equilibrium. *Phys Rev E* 61:2361–2366.
3. Hummer G, Szabo A (2001) Free energy reconstruction from nonequilibrium single-molecule pulling experiments. *Proc Natl Acad Sci USA* 98:3658–3661.
4. Liphardt J, Dumont S, Smith SB, Tinoco I, Bustamante C (2002) Equilibrium information from nonequilibrium measurements in an experimental test of Jarzynski's equality. *Science* 296:1832–1835.
5. Collin D, et al. (2005) Verification of the Crooks fluctuation theorem and recovery of RNA folding free energies. *Nature* 437:231–234.
6. Hummer G, Szabo A (2005) Free energy surfaces from single-molecule force spectroscopy. *Acc Chem Res* 38:504–513.
7. Nome RA, Zhao JM, Hoff WD, Scherer NF (2007) Axis-dependent anisotropy in protein unfolding from integrated nonequilibrium single-molecule experiments, analysis, and simulation. *Proc Natl Acad Sci USA* 104:20799–20804.
8. Woodside MT, et al. (2006) Direct measurement of the full sequence-dependent folding landscape of a nucleic acid. *Science* 314:1001–1004.
9. Woodside MT, García-García C, Block SM (2008) Folding and unfolding single RNA molecules under tension. *Curr Opin Chem Biol* 12:640–646.
10. Ferrenberg AM, Swendsen RH (1989) Optimized Monte Carlo data analysis. *Phys Rev Lett* 63:1195–1198.
11. Minh DDL, Adib AB (2008) Optimized free energies from bidirectional single-molecule force spectroscopy. *Phys Rev Lett* 100:180602.
12. Park S, Khalili-Araghi F, Tajkhorshid E, Schulten K (2003) Free energy calculation from steered molecular dynamics simulations using Jarzynski's equality. *J Chem Phys* 119:3559–3566.
13. Zayed AI (1996) *Handbook of Function and Generalized Function Transformations* (CRC, Boca Raton, FL).
14. Kadanoff LP (2000) *Statistical Physics* (World Scientific, Singapore) p 226.
15. Hummer G, Szabo A (2008) *Theory and Evaluation of Single-Molecule Signals*, eds E Barkai, F Brown, M Orrit, and H Yang (World Scientific, Singapore), pp 139–180.
16. Gebhardt JCM, Bornschlögl T, Rief M (2010) Full distance-resolved folding energy landscape of one single protein molecule. *Proc Natl Acad Sci USA* 107:2013–2018.
17. Liphardt J, Onoa B, Smith SB, Tinoco I, Jr., Bustamante C (2001) Reversible unfolding of single RNA molecules by mechanical force. *Science* 292:733–737.
18. Dudko OK, Hummer G, Szabo A (2008) Theory analysis and interpretation of single-molecule force spectroscopy experiments. *Proc Natl Acad Sci USA* 105:15755–15760.
19. Best RB, Hummer G (2008) Protein folding kinetics under force from molecular simulation. *J Am Chem Soc* 130:3706–3707.
20. Berezhkovskii AM, Zitserman VY (1991) Activated rate-processes in the multidimensional case. Consideration of recrossings in the multidimensional Kramers problem with anisotropic friction. *Chem Phys* 157:141–155.
21. Dudko OK, Hummer G, Szabo A (2006) Intrinsic rates and activation free energies from single-molecule pulling experiments. *Phys Rev Lett* 96:108101.
22. Suzuki Y, Dudko OK (2010) Single-molecule rupture dynamics on multidimensional landscapes. *Phys Rev Lett* 104:048101.

23. Manosas M, Ritort F (2005) Thermodynamic and kinetic aspects of RNA pulling experiments. *Biophys J* 88:3224–3242.
24. Lucy LB (1974) An iterative technique for the rectification of observed distributions. *Astron J* 79:745–754.
25. Richardson WH (1972) Bayesian-based iterative method of image restoration. *J Opt Soc Am A* 62:55–59.
26. Shirts MR, Bair E, Hooker G, Pande VS (2003) Equilibrium free energies from nonequilibrium measurements using maximum-likelihood methods. *Phys Rev Lett* 91:140601.