Experimental Free Energy Surface Reconstruction From Single-Molecule Force Spectroscopy Using Jarzynski’s Equality

Nolan C. Harris, Yang Song, and Ching-Hwa Kiang*
Department of Physics and Astronomy, Rice University, Houston, TX 77005

We used the atomic force microscope to manipulate and unfold individual molecules of the titin I27 domain and reconstructed its free energy surface using Jarzynski’s equality. The free energy surface for both stretching and unfolding was reconstructed using an exact formula that relates the nonequilibrium work fluctuations to the molecular free energy. In addition, the unfolding free energy barrier, i.e. the activation energy, was directly obtained from experimental data for the first time. This work demonstrates that Jarzynski’s equality can be used to analyze nonequilibrium single-molecule experiments, and to obtain the free energy surfaces for molecular systems, including interactions for which only nonequilibrium work can be measured.

One way to probe molecular properties is to drive a system out of equilibrium and to observe the response. Interpretation of data from dynamic measurements allows one to reconstruct both the equilibrium properties of molecules and responses to external perturbations [1]. Equilibrium parameters are usually deduced from kinetic measurements, and it remains challenging to relate nonequilibrium distribution data to equilibrium properties [2]. Advances in single-molecule manipulation and measurement techniques have made it possible to directly probe the dynamics of molecular interactions [3, 4]. The nonequilibrium work theorem, i.e. Jarzynski’s equality [5], relates nonequilibrium measurements of nanoscale systems to equilibrium properties [6, 7, 8]. It promises to extract thermodynamic parameters such as free energies from single-molecule measurements.

Forced unfolding of single molecules, now achievable using the atomic force microscope (AFM) and laser optical tweezers, has been used to probe the molecular interactions and mechanical properties of individual molecules [3, 9]. In these experiments, single molecules are held at both ends and stretched while the cantilever spring restoring force ($F_s$) is measured. Applying an external force drives the system out of equilibrium, and transitions between states are directly observed as the system settles to a new equilibrium state. However, interpretation of these results and deduction of equilibrium properties from these nonequilibrium measurements remains controversial [10, 11, 12, 13].

It has been widely anticipated that equilibrium free energy differences can be derived from nonequilibrium measurements using Jarzynski’s equality [5]. The difference in equilibrium free energy, $\Delta G$, is related to the fluctuations of work performed during a nonequilibrium process, $W_\lambda$, by [5, 8]

$$\langle e^{-\beta W_\lambda} \rangle_N \equiv \int dW_{\lambda} \rho(W_{\lambda}) e^{-\beta W_{\lambda}} = e^{-\beta \Delta G} \tag{1}$$

where $\beta = (1/k_B T)$, $k_B$ is the Boltzmann constant, and $T$ is the temperature of the thermal bath. The $\langle \ldots \rangle_N$ represents an average over $N$ realizations of the process, and the equality is exact in the limit $N \to \infty$. The nonequilibrium work distribution, $\rho(W_\lambda)$, depends on the schedule for varying the work parameter $\lambda$, which is the external perturbation. The equality is simple; however, its application to interpreting single-molecule results is not straightforward. The equation involves the thermodynamic work done on the system and the controlled work parameter with $W_\lambda = \int F \cdot d\lambda$. In AFM experiments, the system includes the cantilever spring and the molecule.

FIG. 1: (color). Single-molecule pulling experiments using AFM. (a) One end of the molecule is attached to the cantilever tip and the other end to a gold substrate, whose position is controlled by a piezolectric actuator. An analogue of the single-molecule force measurements is illustrated. The cantilever spring obeys Hooke’s law, whereas the protein molecular spring follows the worm-like chain model (illustrated using rubber bands). (b) A representative force versus time trace, taken at 1.00 $\mu$m/s using a cantilever with a spring constant of 0.04 N/m. Each force peak represents unfolding of an individual titin I27 domain, with the final peak resulting from the detachment of the molecule from the AFM tip. (c) Corresponding force-extension curve. The tip force baseline was determined using the part of the force curve where the molecule is completely detached from the tip, when the cantilever spring is at its equilibrium position.
plus water, and $\lambda$ refers to the change in cantilever anchor to stage distance (see Fig. 1), not the tip-to-sample distance, which measures the molecule end-to-end distance $z$, i.e. the order parameter, or reaction coordinate.

We briefly review the experimental setup to which Eq. (1) applies. Consider at time $t = 0$, the system is at an equilibrium state $\lambda(0) = \lambda_A$. We perform external work on the system by controlling the work parameter following a $pre-determined$ schedule, $\lambda(t)$, from an initial state $\lambda_A$ to a final state $\lambda_B$. The system is then allowed to relax to equilibrium while $\lambda$ is held constant at $\lambda_B$. Since we do not perform external work on the system during relaxation, we can omit this last step and the equality still holds. Hence Jarzynski’s equality allows us to determine the $G(\lambda)$ from an equilibrated state $A$ to an arbitrary state $B$.

A proof-of-principle experiment and molecular dynamics simulations testing the Jarzynski estimator have been performed [14, 15]. The experimental test involved stretching individual RNA molecules reversibly and irreversibly using optical tweezers, and the free energy of unfolding, $i.e.$ the stability of the molecules, was determined. However, the usefulness of Jarzynski’s equality lies with its ability to obtain directly the entire free energy landscape, which could only be estimated using kinetic approaches to date [16]. We will show that Jarzynski’s equality can be used to determine directly the free energy profile of molecular stretching and unfolding, including the free energy barrier of unfolding.

Our system of interest is the mechanical unfolding of the I27 domain of human cardiac titin [17]. The mechanical properties of the immunoglobulin (Ig)-like domains are directly correlated with the protein’s biological function in the muscles [8]. The kinetic barrier of these mechanical proteins is important in determining the dynamic behavior of the proteins during the stretch-release process. Therefore, the titin free energy surface, including the unfolding barrier height, is useful for quantification of titin’s function in the heart muscle.

We used AFM to stretch individual molecules of eight serially linked repeats of the titin I27 domain, as illustrated in Fig. 1. The protein was stretched when the substrate stage was moved by $\lambda$, which was set at a constant velocity $v$, $i.e.$ $\lambda = vt$. The cantilever displacement from its equilibrium position $\Delta z$ was recorded, and the molecular end-to-end distance as a function of time was calculated using $z = \lambda - \Delta z$. The force curves are aligned using the best worm-like chain (WLC) fit of the force below the unfolding force. This method has been shown to minimize the effect of instrument drift that affects the measured values [18]. To correctly calculate $\Delta G$ as a function of the molecular end-to-end distance, we used an exact expression that connects the nonequilibrium fluctuations of work to the Gibbs free energy $G(z)$ [2]:

$$e^{-\beta G(z)} = \langle \delta (z - z_t)e^{-\beta W_z(t) - U_0(z_0, \lambda_A)} \rangle_N$$

where $z_0$ and $z_t$ are the end-to-end distances of the molecule at times 0 and $t$ during one realization of the process, $F_m$ is the force on the molecule, $W_z(t)$ is the mechanical work done on the molecule up to time $t$, $\delta (z - z_t)$ is the Dirac $\delta$ function, and $U_0$ is the potential energy stored in the cantilever spring at time 0.

To calculate $G(z)$ using Eq. (2), we divided each of the $N$ trajectories of duration $\tau$ into discrete time steps $\delta t$ so $T = \tau / \delta t$, where $T$ is the total number of time steps.
the width $\delta$ error of 0.07 nm uncertainty in $z$. As an example, if we use 15 nm or 19 nm, $\Delta z$ of unfolding is insensitive to the distance of reconstruction. The integration starts from the beginning of the curve, where the cantilever is close to its resting position, $z = 0$, at $t = 0$. This initial condition is required for using Jarzynski estimator, which states that the process needs to start from an equilibrated state. It is also advantageous when using Eq. (3) that the initial energy stored in the cantilever spring, $U_0(z_0, \lambda_\Lambda)$, is close to 0. We compare this result to the approximate free energy surface derived from

$$ e^{-\beta G'} \approx \langle e^{-\beta \int F_m dz - U_0(z_0, \lambda_\Lambda)} \rangle_N \tag{4} $$

The unfolding free energy surface of titin I27 determined from Eq. (4) and Eq. (3) are very similar, perhaps due to the relatively stiff cantilever used in AFM. However, it is physically and theoretically more meaningful to use Eq. (3), since determination of the entire free energy surface relies on converting the coordinate from $t$ to $z$. Figure 2 displays the free energy surface measured at three different velocities, determined using Eq. (3).

The unfolding free energy barrier $\Delta G^u_0$ can be calculated from the reconstructed free energy curve. Using 0.6 nm as the distance between the native and the transition state ($x_u$), we calculated the unfolding free energy barrier $\Delta G^u_0$ for pulling velocities of 0.05, 0.10, and 1.00 nm/s, to be 11.0, 11.7, and 11.4 kcal/mol, from Eq. (3) and 11.5, 11.5, and 10.7 kcal/mol from Eq. (4), respectively. The uncertainty in the averaged $\Delta G^u_0 = 11.4$ and 11.2 kcal/mol, calculated using the bootstrap method, is 0.4 and 0.3 kcal/mol for Eqs. (3) and (4), respectively. This result compares favorably to an estimated value of 10–16 kcal/mol [12, 20, 21]. A major source of error for $\Delta G^u_0$ from Jarzynski estimator comes from the uncertainty in $x_u$. Using the largest estimated error of 0.07 nm uncertainty in $x_u$, the estimated uncertainty of $\Delta G_0^u$ is 1.2 kcal/mol.

The free energy surface is accurately reconstructed from $z = 0$ to 17 nm, the transition state. The free energy of unfolding is insensitive to the distance of reconstruction. As an example, if we use 15 nm or 19 nm, $\Delta G^u_0$ changes by 2.5 and 4.4%, respectively, for the pulling velocity of 0.05 $\mu$m/s. The vast majority of the proteins in the ensemble are in the folded state (99.9997% using the free energy $\Delta G^u_0$ from Ref. [22]) so the contribution from the initially unfolded proteins is negligible. To minimize the contribution from other unfolded domains to the measured free energy, we analyzed only the first domain stretching event. Using all domain unfolding events in the analysis changes $\Delta G^u_0$ by less than 2%.

Note that it is not possible to compare our results directly to published values, since $\Delta G^u_0$ has not been determined, and only kinetic information is available. Chemical denaturant studies gave an estimated unfolding rate constant, $k_u^0$, of $6 \times 10^{-4} \text{s}^{-1}$, while forced-unfolding studies gave an estimated $k_u^0$ of $10^{-3} \text{--} 10^{-6} \text{s}^{-1}$ [3, 10, 11, 12], and the $\Delta G^u_0$ was then calculated using $k_u^0 = k_0 e^{-\beta \Delta G^u_0}$. Since the prefactor $k_0$ of protein unfolding is unknown, the free energy barrier can only be estimated by this procedure [24, 25]. However, combining our free energy determination with the kinetic information, we can determine the prefactor $1/k_0$ to be 6 $\mu$s, which lies within the expected range [24, 26].

The free energy surface immediately past the transition state cannot be reconstructed with high accuracy from constant velocity unfolding experimental data. This is because the force exerted on the molecule is discontinuous when the domain ruptures and expands against the cantilever. In the region where the domain ruptures and the cantilever snaps back to its equilibrium posi-
tion, the assumption that the force on the molecule ($F_m$) is balanced by the cantilever spring restoring force ($F_s$) no longer holds. Therefore, using the measured $F_s$ gave rise to an overestimate of the free energy. Note that even though the snapping process is almost instantaneous (small change in $t$, hence $\lambda$), the change in $z$ is significant because $\Delta z = F_s/k_s$, where $k_s$ is the cantilever spring constant (see Fig. 3). A lower pulling velocity and larger spring constant will reduce the size of the snapping region. However, we can estimate the folding free energy barrier, $\Delta G_f^\dagger$, from the equilibrium unfolding free energy determined from chemical denaturation studies. 22. Using $\Delta G_u = 7.5$ kcal/mol, we obtained $\Delta G_f^\dagger = 3.9$ kcal/mol, in the expected range for titin I27. Figure 3 summarizes the reconstructed free energy surface and its relation to pulling experiments.

One requirement for using Jarzynski’s equality is that the schedule for varying the work parameter $\lambda$ must be pre-determined, which means that constant force ramp is not an appropriate schedule. A constant $dF_s/dt$ requires force feedback and, since the measured force $F_s$ fluctuates from one pull to another, the result is a different schedule of $\lambda$ for each realization. On the other hand, the dynamic force spectroscopy method commonly used in AFM pulling of proteins is particularly suitable for such analysis because the schedule for pulling is pre-determined and remains the same for all experiments performed at the same velocity.

Using nonequilibrium single-molecule measurements and Jarzynski’s equality, we have reconstructed the free energy surface of both mechanical stretching and unfolding of the I27 domain of human cardiac titin. Since the profile is an equilibrium property, the reverse of the free energy surface of stretching is equivalent to that of protein folding from an extended state. The unfolding free energy barrier and the prefactor were determined directly from experimental measurements without having to assume either a two- or a three- state model, which are major sources of error in the event of populated intermediate states. In fact, with adequate resolution and accuracy, an intermediate state should be directly resolvable in the free energy curve. The topography and the roughness of the folding free energy landscape can also be determined. Reconstruction of free energy surfaces directly from experimental data is valuable to obtain fundamental thermodynamic properties such as the free energy barrier of unfolding, to understand the mechanical properties of the molecule, and to compare with theory and simulation results. 27. With a complete characterization of the free energy surface of molecular processes, questions such as whether thermal, chemical, and mechanical unfolding probe the same process may be resolved. Moreover, since the free energy surface is determined in a particular environment, how the free energy surface changes with environmental parameters such as temperature, solution ionic concentration, and acidity may now be evaluated. Quantification of the molecular response to external parameters should lead to a better understanding of molecular behavior in the complex cellular environment.

We thank C. Jarzynski, D. Thirumulai, K. W. Plaxco, and S. S. Plotkin for helpful discussions. We also thank NSF DMR-0505814, NIH 1T90DK70121-01, Hamill Innovation Fund, and Welch Foundation C-1632 for support.

*To whom correspondence should be addressed.

Electronic address: chkang@rice.edu

[1] H. Frauenfelder, P. G. Wolynes, and R. H. Austin, Rev. Mod. Phys. 71, S419 (1999); J. M. Fernandez, S. Chu, and A. F. Oberhauser, Science 292, 653 (2001); P. L. Geissler et al., Science 291, 2121 (2001).
[2] R. F. Fox, Proc. Natl. Acad. Sci. USA 100, 12537 (2003).
[3] M. Rief et al., Science 276, 1109 (1997).
[4] D. K. Lubensky and D. R. Nelson, Phys. Rev. Lett. 85, 1572 (2000).
[5] C. Jarzynski, Phys. Rev. Lett. 78, 2690 (1997).
[6] G. Hummer and A. Szabo, Proc. Natl. Acad. Sci. USA 98, 3658 (2001).
[7] G. Hummer and A. Szabo, Acc. Chem. Res. 38, 504 (2005).
[8] C. Jarzynski, in Supplement of Progress in Theoretical Physics (2006), 165, pp. 1–17.
[9] B. L. Smith et al., Nature 399, 761 (1999).
[10] M. Carrión-Vázquez et al., Proc. Natl. Acad. Sci. USA 96, 3694 (1999).
[11] P. M. Williams et al., Nature 422, 446 (2003).
[12] G. Hummer and A. Szabo, Biophys. J. 85, 5 (2003).
[13] M. S. Li et al., Proc. Natl. Acad. Sci. USA 103, 93 (2006).
[14] J. Liphardt et al., Science 296, 1832 (2002).
[15] S. Park et al., J. Chem. Phys. 119, 3559 (2003).
[16] M. Schlierf and M. Rief, Biophys. J. 90, L33 (2006); M. Manosas, D. Collin, and F. Ritort, Phys. Rev. Lett. 96, 218301 (2006); M. T. Woodside et al., Science 314, 1001 (2006).
[17] K. Wang, J. McClure, and A. Tu, Proc. Natl. Acad. Sci. USA 76, 3698 (1979).
[18] D. Collin et al., Nature 437, 231 (2005).
[19] C. Bustamante, Y. R. Chemla, N. R. Forde, and D. Izhaky, Annu. Rev. Biochem. 73, 705 (2004).
[20] O. K. Dudko, G. Hummer, and A. Szabo, Phys. Rev. Lett. 96, 108101 (2006).
[21] C. F. Wright, A. Steward, and J. Clarke, J. Mol. Biol. 338, 445 (2004).
[22] V. Grantcharova, E. J. Alm, D. Baker, and A. L. Horwich, Curr. Opin. Struct. Biol. 11, 70 (2001).
[23] J. E. Kohn et al., Proc. Natl. Acad. Sci. USA 101, 12491 (2004).
[24] B. Schuler, E. A. Lipman, and W. A. Eaton, Nature 419, 743 (2002).
[25] R. B. Best, S. B. Fowler, J. L. Toca-Herrera, and J. Clarke, Proc. Natl. Acad. Sci. USA 99, 12143 (2002).
[26] M. S. Li, D. Klimov, and D. Thirumalai, Polymer 45, 573 (2004).
[27] S. S. Plotkin and J. N. Onuchic, Proc. Natl. Acad. Sci. USA 97, 6509 (2000).