Published in final edited form as:

Phys Chem Chem Phys. 2010 July 7; 12(25): 6579–6582. doi:10.1039/b926889h.

Determination of equilibrium free energy from nonequilibrium work measurements

L. Y. Chen, D. A. Bastien, and H. E. Espejel

Department of Physics, University of Texas at San Antonio, One UTSA Circle, San Antonio, Texas 78249, USA

Abstract

The fluctuation–dissipation theorem (FDT) of Brownian dynamics (BD) is applied to extract the equilibrium free-energy profile from the nonequilibrium, irreversible work measured in single-molecule pulling experiments. Two sets of *in silico* experiments are performed to explore the free-energy landscape of deca-alanine peptide as a function of its end-to-end distance and to determine the free-energy profile of water permeation through the channels of aquaglyceroporin GlpF. With a small number of pulling paths sampled, the BD-FDT is shown to produce accurate estimates of the free-energy profiles for both systems.

Accurately extracting equilibrium free energy from nonequilibrium work done to a system is important to our understanding of single-molecule pulling experiments^{1–6} and to the computation of free energy in general.^{4,7–10} The formula for this important task in the current literature is the Jarzynski equality (JE),⁴

$$\exp(-\beta\Delta G) = \left\langle \exp\left(-\beta W_{A \rightarrow B}^t\right) \right\rangle_F. \quad (1)$$

JE relates the free energy difference between State A and State B, $\Delta G = G_B - G_A$, to the thermodynamic work,

$$W_{A \rightarrow B}^t = \int_{A \rightarrow B} dt \lambda \frac{\partial H}{\partial \lambda}, \quad (2)$$

done to the system when it is driven/pulled from State A to State B. Here $\lambda(t)$ is the control parameter of the pulling experiments. When the system is pulled with a spring (spring constant k), the total Hamiltonian takes the following form,⁵

$$H(x, t) = H_0(x) + 1/2k[z(x) - z(x_A) - \lambda(t)]^2 \quad (3)$$

where $H_0(x)$ is the Hamiltonian of the biomolecule itself without the external control term. $z(x)$ is the “order parameter” of the biomolecule as a function of x , the collection of the coordinates of all the atoms of the system. For constant-velocity pulling experiments or steered molecular dynamics (SMD) simulations,⁶ $\lambda(t) = vt$ where v is the pulling speed. The order parameter $z(x)$ is simply the center-of-mass coordinate of the pulled atom/molecule. x_A is the set of the coordinates of all the atoms when the system is in State A. The brackets in eqn (1) represent the statistical average over the pulling paths sampled in a given experimental or computational study,

$$\langle f(W) \rangle_F = \frac{1}{N_F} \sum_{p=1}^{N_F} f(W_{A \rightarrow B}^{(p)}) \quad (4)$$

where the index p indicates the p -th of the N_F forward paths from State A to State B.

It is interesting to note the symmetry of the JE. Eqn (1) relates the free-energy difference $G_B - G_A$ to the work $W_{A \rightarrow B}^f$ along the forward paths from A to B. If it is valid, $G_A - G_B$ must be related to the work $W_{B \rightarrow A}^r$ along the reverse paths from B to A in the same manner. Namely,

$$\exp(+\beta\Delta G) = \left\langle \exp(-\beta W_{B \rightarrow A}^r) \right\rangle_R, \quad (5)$$

in which the statistical average is over N_R reverse paths from State B to State A,

$$\langle f(W) \rangle_R = \frac{1}{N_R} \sum_{p=1}^{N_R} f(W_{B \rightarrow A}^{(p)}). \quad (6)$$

The definition of $W_{B \rightarrow A}^r$ is given in eqn (2) with A and B exchanging their places. Combining eqn (1) and eqn (5) and defining the following self-consistency factor,¹¹

$$C_{FR} \triangleq \langle \exp(-\beta W_{A \rightarrow B}^f) \rangle_F \langle \exp(-\beta W_{B \rightarrow A}^r) \rangle_R, \quad (7)$$

we observe that the validity of JE requires $C_{FR} = 1$. This self-consistency condition can be used to check the accuracy of the JE estimate of the free-energy difference. The magnitude of $C_{FR} - 1$ gives a measure of the deviations of the JE estimates from the true free energy in a given experimental or computational study.

Many researchers believe that the JE is exact in the limit of $N_F \rightarrow \infty$. But for a finite N_F , the JE produces large deviations from the true free energy when the pulling is irreversible.⁴⁻⁷ In this paper, we employ the fluctuation-dissipation theorem (FDT) of Brownian dynamics (BD) for extracting equilibrium free energy from irreversible work.¹² In terms of mechanical work, the BD-FDT relates the equilibrium free energy to the nonequilibrium work as follows:

$$\exp(-\beta\Delta G) = \langle \exp(-\beta W_{A \rightarrow B}^f/2) \rangle_F \langle \exp(-\beta W_{B \rightarrow A}^r/2) \rangle_R. \quad (8)$$

Here, $W_{A \rightarrow B}$ is the force-displacement integral along a forward path from State A to State B,

$$W_{A \rightarrow B} = \int_{A \rightarrow B} k [\lambda(t) + z(x_A) - z(x)] dz. \quad (9)$$

$W_{B \rightarrow A}$ is a similarly defined force-displacement integral along a reverse path from State B to State A. Note that the mechanical work defined in eqn (9) is not equal to the thermodynamic work defined in eqn (2) other than in the limit of $k \rightarrow \infty$. Moreover, even in

the $k \rightarrow \infty$ limit, BD-FDT in eqn (8) and JE in eqn (1) do not agree with one another unless the pulling is nearly reversible. In the near-reversible regime, all the relevant paths fall in the close vicinity of one single path connecting States A and B. Then the statistical averages in eqn (1) and (8) become trivial. The free-energy difference is simply approximately equal to the reversible work.

In the rest of this paper, we apply BD-FDT [eqn (8)] along with JE [eqn (1)] to two biophysical systems to explore their free-energy profiles and to illustrate the efficiency and accuracy of BD-FDT. We have performed *in silico* experiments of mechanically unfolding deca-alanine peptide helix and pulling water molecules through the pores of aquaglyceroporin, the glycerol uptake facilitator (GlpF). Both JE and BD-FDT are used to extract the free-energy difference from the irreversible work measured. All simulations are implemented with the NAMD molecular dynamics software package.¹³ Atomistic interactions are described by the all-atom CHARMM 27 force field.¹⁴ Pulling experiments are done in the limit of $k \rightarrow \infty$, by fixing the center-of-mass coordinate, $z(x)$, of the pulled atom/molecule to the control parameter $\lambda(t)$ at each time step. In this limit, $dz = vdt$ and the pulling force then is equal to the net force on the pulled atom/molecule exerted by the rest of the entire system, namely, $k[\lambda(t) - z(x) + z(x_A)] \rightarrow \partial H_0 / \partial z$. Therefore, the thermodynamic work in eqn (2) is equal to the mechanical work in eqn (9).

In the first set of *in silico* experiments, we study the free-energy landscape of a deca-alanine peptide in vacuum that has been thoroughly investigated in ref. 6. This system, consisting of 104 atoms, is simple enough for which an exact solution is achievable by unfolding and refolding it at a reversibly slow speed, yet complex enough to serve as a prototype for proteins. We follow ref. 6, choosing the z -coordinate along the helix axis. We fix the nitrogen atom on one end at the origin and pull the nitrogen atom on the other end with a constant speed v along the z -axis, sampling one unfolding (forward) path from State A ($z_A = 13.42 \text{ \AA}$) to State B ($z_B = 33.42 \text{ \AA}$). Then from State B, we pull it back with the same speed, sampling one refolding (reverse) path. In this case, we choose the end-to-end distance as the order parameter. It is simply equal to the z -coordinate of the pulled nitrogen atom.

Fig. 1 shows five unfolding paths along with five refolding paths at the pulling speed of $10 \text{ \AA ns}^{-1}$. Clearly, at this speed, $W_{A \rightarrow B}(z) + W_{B \rightarrow A}(z) \neq 0$, the pulling is irreversible. Fig. 2 shows the work done to the polypeptide when it is unfolded at a speed of $0.1 \text{ \AA ns}^{-1}$ along with the work done to the system when it is refolded at the same speed. The negligible difference between the two work curves suggests that unfolding/refolding at $0.1 \text{ \AA ns}^{-1}$ is indeed slow enough. The processes can be regarded as near-reversible [$W_{A \rightarrow B}(z) + W_{B \rightarrow A}(z) \approx 0$] and the work as a function of the end-to-end distance is a good approximation to the exact solution of free energy, [$G(z) - G_A \approx W_{A \rightarrow B}(z)$].

Fig. 2 also shows the results of BD-FDT and the results of JE in comparison with the exact solution. The free energy of this research work is obtained through eqn (8) with 10 unfolding and 10 refolding paths at three irreversible speeds of 10, 20, and $40 \text{ \AA ns}^{-1}$. The free energy so obtained is very close to the exact value with deviations less than $0.2 \text{ kcal mol}^{-1}$ for $v < 20 \text{ \AA ns}^{-1}$. The JE free energy is obtained through eqn (1) from 20 unfolding paths sampled at the same irreversible speeds. In ref. 6, Park *et al.* give a thorough study of the same system based on JE. Sampling 10 blocks of 10 unfolding paths at pulling speed of $10 \text{ \AA ns}^{-1}$, they find that the deviation of the JE value of the free energy is as large as $1.9 \text{ kcal mol}^{-1}$ from the exact solution. Our simulations confirm their results. Park *et al.* also employ the second and third order moments of work in place of the exponential average of JE. They conclude that the second order moment gives the best result and that the third order moment causes greater deviations from the exact solution and greater error bars, indicating to us the non-Gaussianness of the system's stochastic dynamics.

It is worth noting that, in the study of deca-alanine, the nonequilibrium BD-FDT approach significantly speeds up the computation of the free-energy profile as compared with equilibrium approaches. An equilibrium approach such as thermodynamic integration would require a computing time similar to the near-reversible pulling. In the present work, the near-reversible pulling over a distance of 20 Å at the speed of 0.1 Å ns⁻¹ takes 200 ns of MD simulation from State A to State B. In contrast, the BD-FDT approach (sampling 10 forward and 10 reverse paths at the speed of 20 Å ns⁻¹) takes only 20 ns of MD simulation but produces a free-energy curve that nearly completely overlaps with the exact results (Fig. 2b). This is clear evidence for the efficiency of the BD-FDT approach.

In the second set of *in silico* experiments, we explore the free-energy landscape of water permeation through aquaglyceroporin GlpF.15⁻¹⁸ GlpF possesses a homotetrameric structure. Each monomer forms an independently functional pore (channel). The core of this protein consists of two half-membrane spanning repeats related by an essentially two-fold symmetry. The N-termini of the α -helical repeats convene approximately at the center of the channel, where the Asn-Pro-Ala (NPA) motifs are located. The selectivity filter (SF), the narrowest section of the pore, is about 8 Å from the center of the channel, namely, from the central plane of the lipid bilayer. In this work, the *z*-axis of the Cartesian coordinates is chosen to be normal to the water-membrane interface, pointing from periplasm to cytoplasm. The origin of the coordinates is set so that *z* = 7.7 Å at the midpoint between the N-termini of Asn 68 and Asn 203.

This work uses the all-atom model of the GlpF-membrane system studied in ref. 19, consisting of GlpF homotetramer embedded in a fully hydrated palmitoylcholinephosphatidylethanolamine (POPE) bilayer. The GlpF-POPE-bilayer system is explicitly solvated with TIP3 water. Sodium and chlorine ions are added and the whole system is neutralized. The system consists of 106 189 atoms. The pressure and the temperature are maintained at 1 bar and 300 K respectively. The Langevin damping coefficient is chosen to be 5 ps⁻¹. The particle-mesh Ewald method is utilized to compute electrostatic interactions. The time step of 1 fs is used for short-range interactions and 4 fs for long-range forces. Covalent bonds of all hydrogen atoms are fixed to their equilibrium length. The periodic boundary conditions are applied.

Starting from the fully equilibrated structure of ref. 19, we run equilibrium Langevin dynamics for 4 ns after removing the four glycerols from the system. Then we perform constant-velocity SMD simulations to pull four waters through the four channels of GlpF simultaneously, one water molecule through each of them. Two sets of pulling paths are sampled. Each set has four forward paths (from periplasm to cytoplasm) and four reverse paths (from cytoplasm to periplasm). The work along these pulling paths is plotted in Fig. 3. In this case, the order parameter is chosen as the center-of-mass *z*-coordinate of one pulled water molecule. The pulling speed *v* = 50 Å ns⁻¹. Clearly, $W_{A \rightarrow B}(z) + W_{B \rightarrow A}(z) \neq 0$, the pulling is irreversible.

Fig. 4 shows the results of the free-energy estimates using the JE in eqn (1) and the BD-FDT in eqn (8). The application of the BD-FDT produces the free-energy profile (red curve) in agreement with the equilibrium simulation results (crosses connected by green lines).¹⁶ However, the JE results (magenta and blue curves) are far off. In keeping with Park *et al.*⁶ we also computed the free-energy profile using the second moment approximation of the JE. Even though the second moment results are closer to the correct free-energy profile than the full exponential form of JE, the deviation is still as large as 6 kcal mol⁻¹.

It is interesting to note that water permeation through the channels of GlpF is collective diffusion. Multiple water molecules form single files inside each pore and permeate through

a channel in a concerted manner. In our *in silico* experiments, only one water per channel is pulled. Yet our nonequilibrium SMD study is actually in full agreement with the equilibrium simulations of the highly correlated diffusion of multiple water molecules.¹⁶ It indicates that the path averages in the numerator and the denominator of BD-FDT [eqn (8)] cancel out the irreversible, nonequilibrium factors (the dissipative work terms in particular) in a very efficient manner even for a small number of pulling paths sampled. Note that the computing time required in this numerical study of the free-energy profile is mainly to sample the eight forward and eight reverse paths at the speed of $50 \text{ \AA ns}^{-1}$ over a distance of $20 \text{ \AA}$ each way. That is about 12 ns of MD simulation including the equilibrations. In contrast, an equilibrium approach such as Adaptive Biasing Force (ABF) would take about 120 ns of MD simulation to reach similar precision.¹⁹

In summary, we have applied BD-FDT [eqn (8)] along with JE [eqn (1)] to explore the free-energy landscapes of deca-alanine peptide in vacuum and water transport through aquaglyceroporin GlpF. Based on the accuracy and efficiency of the BD-FDT, it is reasonable to expect its applicability to various *in silico* and *in vitro* experiments of mechanically pulling biomolecules.

Acknowledgments

The authors thank Robert Renthall for sharing his biological insights. They acknowledge support from an NIH SC3 grant (GM084834) and the Texas Advanced Computing Center. DAB also acknowledges a UTSA CRSG grant and HEE, the UTSA MARC U*STAR Program supported by an NIH T34 grant (GM07717).

References

1. For a recent review, see: Neuman KC, Nagy A. Nat. Methods. 2008; 5:491. [PubMed: 18511917] .
2. Liphardt J, Dumont S, Smith SB, Tinoco I Jr, Bustamante C. Science. 2002; 296:1832. [PubMed: 12052949]
3. Harris NC, Song Y, Kiang CH. Phys. Rev. Lett. 2007; 99:068101. [PubMed: 17930869]
4. Jarzynski C. Phys. Rev. Lett. 1997; 78:2690.
5. Hummer G, Szabo A. Proc. Natl. Acad. Sci. U. S. A. 2001; 98:3658. [PubMed: 11274384]
6. Park S, Khalili-Araghi F, Tajkhorshid E, Schulten K. J. Chem. Phys. 2003; 119Park S, Schulten K. J. Chem. Phys. 2004; 120:5946. [PubMed: 15267476]
7. Kosztin I, Barz B, Janosi L. J. Chem. Phys. 2006; 124:064106.
8. Huang L, Makarov DE. J. Chem. Phys. 2006; 124:064108.
9. Ytreberg FM, Swendsen RH, Zuckerman DM. J. Chem. Phys. 2006; 125:184114. [PubMed: 17115745]
10. Meirovitch H. Curr. Opin. Struct. Biol. 2007; 17:181. [PubMed: 17395451]
11. Chen LY. J. Chem. Phys. 2008; 129:091101. [PubMed: 19044857]
12. Chen LY. J. Chem. Phys. 2008; 129:144113. [PubMed: 19045140]
13. Phillips JC, Braun R, Wang W, Gumbart J, Tajkhorshid E, Villa E, Chipot C, Skeel RD, Kale L, Schulten K. J. Comput. Chem. 2005; 26:1781. [PubMed: 16222654]
14. MacKerell AD, Banavali N, Foloppe N. Biopolymers. 2000; 56:257. [PubMed: 11754339]
15. Heller KB, Lin EC, Wilson TH. J. Bacteriol. 1980; 144:274. [PubMed: 6998951]
16. de Groot BL, Grubmuller H. Science. 2001; 294:2353. [PubMed: 11743202]
17. Stroud RM, Nollert P, Miercke L. Adv. Protein Chem. 2003; 63:291. [PubMed: 12629974]
18. Stroud RM, Miercke LJW, O'Connell J, Khademi S, Lee JK, Remis J, Harries W, Robles Y, Akhavan D. Curr. Opin. Struct. Biol. 2003; 13:424. [PubMed: 12948772]
19. Henin J, Tajkhorshid E, Schulten K, Chipot C. Biophys. J. 2008; 94:832. [PubMed: 17921212]

Fig. 1.

$W_{A \rightarrow B}(z)$ (kcal mol⁻¹) as a function of the end-to-end distance z (Å) along five forward (unfolding) paths (green) and $-W_{B \rightarrow A}(z)$ along five reverse (refolding) paths (blue). For forward paths, $W_{A \rightarrow B}(z)$ is the work done to the system when the end-to-end distance is pulled from 13.42 Å to z . For reverse paths, $W_{B \rightarrow A}(z)$ is the work done to the system when the end-to-end distance is reversed from z to 13.42 Å. The pulling velocity is 10 Å ns⁻¹. The peptide is shown in the left when the end-to-end distance is 13.42 Å and in the right when that distance is 33.42 Å.

Fig. 2.

Free energy (kcal mol^{-1}) as a function of the end-to-end distance ($\text{\AA}$). The BD-FDT results of this work nearly overlap with the exact results. The JE results deviate more at higher pulling speed. The zero point of the free energy is set to where the end-to-end distance of deca-alanine is equal to $13.42 \text{ \AA}$.

Fig. 3.

Work done along the forward (red) and reverse (green) pulling paths. For forward paths, the work value, $W_{A \rightarrow B}(z)$, is what is done to the system when the center of mass of a water molecule is pulled from $-8 \text{ \AA}$ to z . For reverse paths, the work value, $W_{B \rightarrow A}(z)$, is what is done to the system when the center of mass of a water molecule is pulled from z to $-8 \text{ \AA}$. The pulling speed is $50 \text{ \AA ns}^{-1}$. The pulling is clearly irreversible because $W_{A \rightarrow B}(z) + W_{B \rightarrow A}(z) \neq 0$.

Fig. 4.

(Top) GlpF monomer forming one of the four pores with Asn-68 and Asn-203 shown in green and Trp-48 and Arg-206 shown in blue (all four in licorice representation), indicating the locations of the NPA motifs and the SF respectively. (Bottom) Free energy estimates from the full JE, the second moment expansion of the JE, the Brownian Dynamics FDT, and the equilibrium simulations of de Groot and Grubmuller (points taken from the free-energy curve in Fig. 3 of ref. 16). The green lines connecting the crosses are only an aid for the eyes). The midpoint between the N_{δ} atoms of Asn-68 and Asn-203 of the NPA motifs is located at $z = 7.7 \text{ \AA}$. The SF is located around $z = 0$ and the NPA is around $z = 7.7 \text{ \AA}$.