Force-dependent hopping rates of RNA hairpins can be estimated from accurate measurement of the folding landscapes

Changbong Hyeon[†‡], Greg Morrison[‡§¶], and D. Thirumalai[§‖††]

[†]Department of Chemistry, Chung-Ang University, Seoul 156-756, Republic of Korea; and [§]Biophysics Program, Institute For Physical Science and Technology, and Departments of [¶]Physics and [‖]Chemistry, University of Maryland, College Park, MD 20742

Edited by Harold A. Scheraga, Cornell University, Ithaca, NY, and approved May 1, 2008 (received for review March 12, 2008)

The sequence-dependent folding landscapes of nucleic acid hairpins reflect much of the complexity of biomolecular folding. Folding trajectories, generated by using single-molecule force-clamp experiments by attaching semiflexible polymers to the ends of hairpins, have been used to infer their folding landscapes. Using simulations and theory, we study the effect of the dynamics of the attached handles on the handle-free RNA free-energy profile $F^o_{eq}(z_m)$, where z_m is the molecular extension of the hairpin. Accurate measurements of $F^o_{eq}(z_m)$ requires stiff polymers with small L/l_p, where L is the contour length of the handle, and l_p is the persistence length. Paradoxically, reliable estimates of the hopping rates can only be made by using flexible handles. Nevertheless, we show that the equilibrium free-energy profile $F^o_{eq}(z_m)$ at an external tension f_m, the force (f) at which the folded and unfolded states are equally populated, in conjunction with Kramers' theory, can provide accurate estimates of the force-dependent hopping rates in the absence of handles at arbitrary values of f. Our theoretical framework shows that z_m is a good reaction coordinate for nucleic acid hairpins under tension.

force spectroscopy | generalized Rouse model | role of handles | reaction coordinate

A molecular understanding of how proteins and RNA fold is needed to describe the functions of enzymes (1) and ribozymes (2), interactions between biomolecules, and the origins of misfolding that is linked to a number of diseases (3). The energy-landscape perspective has provided a conceptual framework for describing the mechanisms by which unfolded molecules navigate the large conformational space in search of the native state (4–6). Recently, single-molecule techniques have been used to probe features of the energy landscape of proteins and RNA that are not easily accessible in ensemble experiments (7–18). It is possible to construct the shape of the energy landscape, including the energy scales of ruggedness (19, 20), by using dynamical trajectories that are generated by applying a constant force (f) to the ends of proteins and RNA (14, 15, 21, 22). If the observation time is long enough for the molecule to sample the accessible conformational space, then the time average of an observable X recorded for the αth molecule $[\langle X\rangle = \lim_{t\to\infty}(\frac{1}{t})\int_0^t d\tau X_\alpha(\tau)]$ should equal the ensemble average ($\langle X\rangle = \lim_{t\to\infty}(\frac{1}{N})\sum_{i=1}^{N} X_i$), and the distribution $P(X)$ should converge to the equilibrium distribution function $P_{eq}(X)$. By using this strategy, laser optical tweezer (LOT) experiments have been used to obtain the sequence-dependent folding landscape of a number of RNA and DNA hairpins (8, 14, 15, 23), by using $X = R_m$, the end-to-end distance of the hairpin that is conjugate to f, as a natural reaction coordinate. In LOT experiments, the hairpin is held between two long handles [DNA (15) or DNA/RNA hybrids (8)], whose ends are attached to polystyrene beads (Fig. 1*a*). The equilibrium free-energy profile $\beta F_{eq}(R_m) = -\log P_{eq}(R_m)$ ($\beta \equiv 1/k_BT$, k_B is the Boltzmann constant, and T is the absolute temperature) may be useful in describing the dynamics of the molecule, provided R_m is an appropriate reaction coordinate.

The dynamics of the RNA extension in the presence of f ($z_m = z_{3'} - z_{5'} \approx R_m$, provided transverse fluctuations are small) is indirectly obtained in an LOT experiment by monitoring the distance between the attached polystyrene beads ($z_{sys} = z_p - z_o$), one of which is optically trapped at the center of the laser focus (Fig. 1*a*). The goal of these experiments is to extract the folding landscape [$\beta F^o_{eq}(z_m)$] and the dynamics of the hairpin in the absence of handles by using the f-dependent trajectories $z_{sys}(t)$. To achieve these goals, the fluctuations in the handles should minimally perturb the dynamics of the hairpin to probe the true dynamics of a molecule of interest. However, depending on L and l_p (L is the contour length of the handle, and l_p is its persistence length), the intrinsic fluctuations of the handles cannot only distort the signal from the hairpin, but also directly affect its dynamics. The first is a problem that pertains to the measurement process, whereas the second is a problem of the coupling between the instruments and the dynamics of RNA.

Here, we use coarse-grained molecular simulations of a RNA hairpin and theory to show that, to obtain accurate $\beta F_{eq}(R_m)$, the linkers used in the LOT have to be stiff, i.e., the value of L/l_p has to be small. To investigate the handle effects on the energy landscape and hopping kinetics, we simulated the hairpin dynamics under force-clamp conditions by explicitly modeling the linkers as polymers with varying L and l_p. Surprisingly, the force-dependent folding and unfolding rates that are directly measured by using the time traces, $z_m(t)$, are closest to the ideal values (those that are obtained by directly applying f, without the handles, to the 3′ end with a fixed 5′ end) only when the handles are flexible. Most importantly, accurate estimates of the f-dependent hopping rates over a wide range of f values, in the absence of handles, can be made by using $\beta F_{eq}(R_m)$, in the presence of handles obtained at $f = f_m$, the transition midpoint at which the native basin of attraction (NBA) and the unfolded basin of attraction (UBA) of the RNA are equally populated. The physics of a hairpin attached to handles is captured by using a generalized Rouse model (GRM), in which there is a favorable interaction between the two noncovalently linked ends. The GRM gives quantitative agreement with the simulation results. The key results announced here provide a framework for using the measured folding landscape of nucleic acid hairpins at $f \approx f_m$ to obtain f-dependent folding and unfolding times and the transition state movements as f is varied (24–29).

Results and Discussions

Modeling the LOT Experiments. To extract the folding landscape from LOT experiments, the time scales associated with the

Author contributions: C.H., G.M., and D.T. designed research, performed research, analyzed data, and wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

[‡]C.H. and G.M. contributed equally to this work.

[††]To whom correspondence should be addressed. E-mail: thirum@umd.edu.

This article contains supporting information online at www.pnas.org/cgi/content/full/0802484105/DCSupplemental.

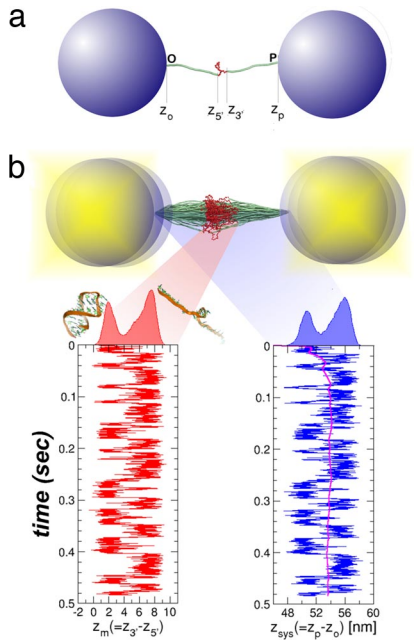

Fig. 1. A schematic diagram of the optical-tweezers setup used to measure the hairpin's folding landscape. (*a*) Two RNA/DNA hybrid linkers are attached to the 5′ and 3′ ends of the RNA hairpin, and a constant force is applied to one end through the bead. (*b*) Ensemble of sampled conformations of the H–RNA–H system during the hopping transitions obtained by using $L = 25$ nm and $l_p = 70$ nm. The illustration is created by using the simulated structures collected every 0.5 ms. An example of the time trace of each component of the system, at $f = 15.4$ pN, given L for both linkers is 25 nm. z_m ($= z_{5'} - z_{3'}$) and z_{sys}($= z_p - z_o$) measure the extension dynamics of the RNA hairpin and of the entire system, respectively. The time-averaged value $\bar{z}_r(t) = (1/\tau)\int_{d\tau}^{t} Z_{sys}(\tau)$ for the time trace of z_{sys} is shown as the bold line. The histograms of the extension are shown on top of each column.

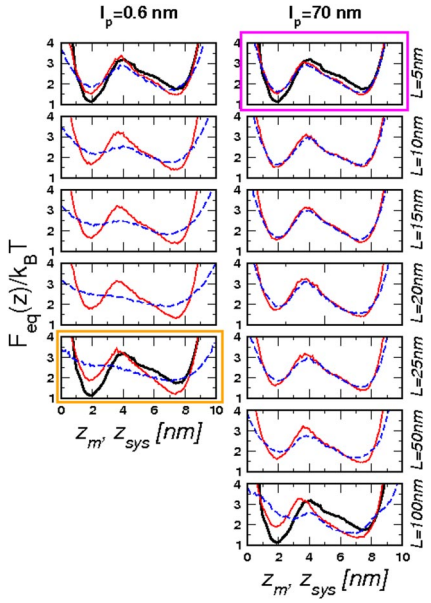

Fig. 2. The free-energy profiles, $F_{eq}(z_{sys})$ (dashed line in blue) and $F_{eq}(z_m)$ (solid line in red), calculated by using the histograms obtained from the time traces $z_{sys}(t)$ and $z_m(t)$ for varying L and l_p. $F_{eq}(z_{sys})$ and $F_{eq}(z_m)$ for a given l_p and different L are plotted in the same graph to highlight the differences. The intrinsic free energy $F^o_{eq}(z_m)$, the free-energy profile in the absence of handles, is shown in black. The condition that produces the least deviation ($l_p = 70$ nm, $L = 5$ nm) and the condition of maximal difference ($l_p = 0.6$ nm, $L = 25$ nm) between $F_{eq}(z_m)$ and $F_{eq}(z_{sys})$ are enclosed in the magenta and orange boxes, respectively.

dynamics of the beads, handles, and the hairpin have to be well separated (30–33). The bead fluctuations are described by the overdamped Langevin equation $\gamma dx_p/dt = -kx_p + F(t)$, where k is the spring constant associated with the restoring force, and the random white-noise force $F(t)$ satisfies $\langle F(t)\rangle = 0$, and $\langle F(t)F(t')\rangle = 2\gamma k_B T\delta(t - t')$. The bead relaxes to its equilibrium position on a time scale $\tau_r = \gamma/k$. In terms of the trap stiffness, k_p, and the stiffness k_m associated with the Handle–RNA–Handle (H–RNA–H; see Fig. 1), $k = k_p + k_m$. With $\gamma = 6\pi\eta a, a = 1\mu$m, $\eta \approx 1cP, k_p \approx 0.01$ pN/nm (34), and $k_m \approx 0.1$ pN/nm, we find $\tau_r \lesssim 1$ ms. In LOT experiments (30, 32, 33), separation in time scales is satisfied such that $\tau^o_U \approx \tau^o_F \gg \tau_r$ at $f \approx f_m$, where τ^o_U and τ^o_F are the intrinsic values of the RNA (un)folding times in the absence of handles.

Because z_m is a natural reaction coordinate in force experiments, the dispersion of the bead position may affect the measurement of $F_{eq}(z_m)$. At equilibrium, the fluctuations in the bead positions satisfy $\langle \delta x^2_{eq}\rangle \sim k_BT/(k_p + k_m) \sim k_BT/k_m$, and hence k_m should be large enough to minimize the dispersion of the bead position. The force fluctuation, $\langle \delta f^2_{eq}\rangle \sim k_BTk^2_p/(k_p + k_m)$, is negligible in the LOT because $k_p \ll k_m$, and as a result $\delta f_{eq}/f_m \approx 0$, because $\delta f_{eq} \approx 0.1$ pN, whereas $f_m \approx 15$ pN. Thus, we model the LOT setup by assuming that the force and position fluctuations due to the bead are small and exclusively focus on the effect of handle dynamics on the folding landscape and hopping kinetics of RNA (Fig. 1 *a* and *b*).

Short, Stiff Handles Are Required for Accurate Free-Energy Profiles. For purposes of illustration, we used the self-organized polymer (SOP) model of the P5GA hairpin (35) and applied a force $f = f_m \approx 15.4$ pN. The force is exerted on the end of the handle attached to the 3′ end of the RNA (P in Fig. 1*a*) while fixing the other end (O in Fig. 1*a*). Simulations of P5GA with handles of length $L = 25$ nm and persistence length $l_p = 70$ nm show that the extension of the entire system ($z_{sys} = z_p - z_0$) fluctuates between two limits centered around $z_{sys} \approx 50$ nm and $z_{sys} \approx 56$ nm (Fig. 1*b*). The time-dependent transitions in z_{sys} between 50 and 56 nm correspond to the hopping of the RNA between the NBA and UBA. Decomposition of z_{sys} as $z_{sys} = z_H{}^{5'} + z_m + z_H{}^{3'}$, where $z_H{}^{5'}$ ($= z_{5'} - z_o$) and $z_H{}^{3'}$ ($= z_p - z_{3'}$) are the extensions of the handles parallel to the force direction (Fig. 1*a*), shows that $z_{sys}(t)$ reflects the transitions in $z_m(t)$ (Fig. 1*b*). Because the simulation time is long enough for the hairpin to ergodically explore the conformations between the NBA and UBA, the histograms collected from the time traces amount to the equilibrium distributions $P_{eq}(X)$, where $X = z_{sys}, z_H{}^{5'}, z_m$, or $z_H{}^{3'}$ [Fig. 1*b*; for $P_{eq}(z_H{}^{5'})$ and $P_{eq}(z_H{}^{3'})$, see the supporting information (SI) Fig. S1*a*]. To establish that the time traces are ergodic, we show that $\bar{z}_T(t) = (1/\tau)\int_{d\tau}^{t} Z_{sys}(\tau)$ reaches the thermodynamic average ($\approx \int_{-\infty}^{\infty} z_{sys}P_{eq}(z_{sys})dz_{sys} = 53.7$ nm) after $t \gtrsim 0.1$ sec (the magenta line on $z_{sys}(t)$ in Fig. 1*b*).

Fig. 1*b* shows that the positions of the handles along the f direction fluctuate, even in the presence of tension, which results in slight differences between $P_{eq}(z_{sys})$ and $P_{eq}(z_m)$. Comparison between the free-energy profiles obtained from the $z_{sys}(t)$ and $z_m(t)$ can be used to investigate the effect of the characteristics of the handles on the free-energy landscape. To this end, we repeated the force-clamp simulations by varying the contour length ($L = 5$–100 nm) and persistence length ($l_p = 0.6$ and 70 nm) of the handles. Fig. 2 shows that the discrepancy between the measured free energy $F_{eq}(z_{sys})$ (dashed lines in blue) and the molecular free energy $F_{eq}(z_m)$ (solid lines in red) increases for the more flexible and longer handles (see *SI Text* and Fig. S1 for further discussion of the dependence of the handle fluctuations on L and l_p). For small l_p and large L, the basins of attraction in $F_{eq}(z_m)$ are not well resolved. The largest deviation between $F_{eq}(z_{sys})$ and $F_{eq}(z_m)$ is found when $l_p = 0.6$ nm and $L = 25$ nm

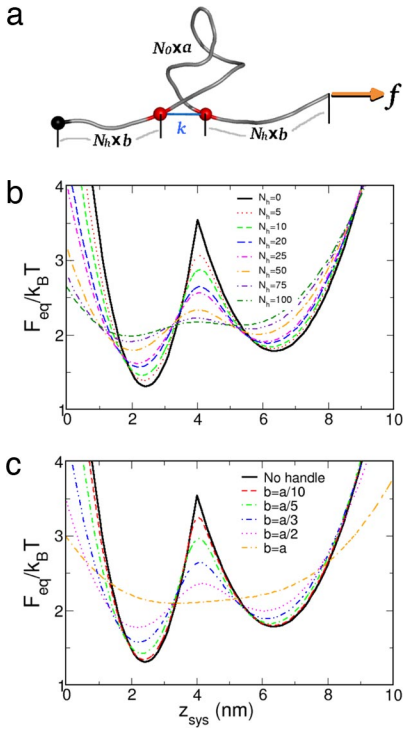

Fig. 3. Free-energy profiles for the GRM. (*a*) A schematic diagram of the GRM, showing the number of monomers (N_0 and N_h) and Kuhn lengths (*a* and *b*) in each region of the chain, the harmonic interaction between the ends of the RNA mimic, and the external tension. (*b*) The free-energy profile for a fixed $b = (a/3)$ and increasing N_h as a function of $R_{sys} \approx z_{sys}$. The barrier heights decrease, and the well depths increase for increasing N_h. (*c*) The free-energy profile for fixed $N_h = 20$ and varying *b*. The barrier heights decrease, and the well depths increase for increasing *b*. In both (*b*) and (*c*), the profiles are shifted so that the positions of the local maxima and minima coincide with those of the intrinsic free energy (with $N_h = 0$).

$(L/l_p \approx 40)$ (the graph enclosed by the orange box in Fig. 2*a*). In contrast, the best agreement between $F_{eq}(z_{sys})$ and $F_{eq}(z_m)$ is found for $l_p = 70$ nm and $L = 5$ nm (the graph inside the magenta box in Fig. 2), which corresponds to $L/l_p \approx 0.07$. In the LOT experiments, $L/l_p \approx 6$–7 (8, 14, 15).

GRM Captures the Physics of H–RNA–H Under Tension. To establish the generality of the relationship between the free-energy profiles as measured by z_m and those measured by z_{sys}, we introduce an exactly solvable model that minimally represents the RNA and handles (Fig. 3*a*). We mimic the hairpin using a Gaussian chain with N_0 monomers and Kuhn length *a*. The endpoints of the RNA mimic are harmonically trapped in a potential with stiffness *k* as long as they are within a cutoff distance $c = 4$ nm. Two handles, each with N_h monomers and Kuhn length *b*, are attached to the ends of the RNA (see *Methods*). We fix one endpoint of the entire chain at the origin and apply a force $f_m \approx 15.4$ pN to the other end. The free energies as a function of both the RNA's extension, $R_m = |\mathbf{r}_{3'} - \mathbf{r}_{5'}|$ ($\approx z_m$ at high *f*) and the system's extension $R_{sys} = |\mathbf{r}_P - \mathbf{r}_0|$ ($\approx z_{sys}$ at high *f*) are exactly solvable in the continuum representation. We choose *k* such that f_m is near the midpoint of the transition, so that $\int_0^c d^3\mathbf{r} P_{eq}(\mathbf{r}) \approx \int_c^\infty d^3\mathbf{r} P_{eq}(\mathbf{r})$. We tune N_0 so that the barrier heights for the GRM and P5GA are similar at $f = f_m$. These requirements give $N_0 = 20$ and $k \approx 0.54$ pN/nm. Although the stiffness in the handles of the simulated system (Fig. 1) cannot be accurately modeled by using a Gaussian chain, the primary effect of attaching the handles is to alter the fluctuations of the endpoints of the RNA. By equating the longitudinal fluctuations for the WLC, $\langle \delta \mathbf{R}_\parallel^2 \rangle_{WLC} \sim L l_p^{-1/2} (\beta f)^{-3/2}$, with the fluctuations for the Gaussian handles, $\langle \delta \mathbf{R}_\parallel^2 \rangle_G \sim Lb$, we estimate that the effective persistence length of the handles scales as $l_p^{eff} \sim b^{-2} f^{-3}$ (see *SI Text* for details). Thus, smaller spacing in the Gaussian handles in the GRM will mimic stiffer handles in the H–RNA–H system. The free energies computed for the GRM, shown in Fig. 3 *b* and *c*, are consistent with the results of the simulations. The free-energy profiles deviate significantly from $F_{eq}^o(z_m)$ as N_h increases or stiffness decreases. The relevant variable that determines the accuracy of $F_{eq}(z_{sys})$ is $N_h b^2 \sim L/l_p^{eff}$, with the free energies remaining unchanged if $N_h b^2$ is kept constant. The barrier height and well depths as a function of z_m are unchanged as a function of *L* and *b*. However, the apparent free energy of activation is decreased as measured by z_{sys} (seen in Fig. 2 as well). The GRM confirms that accurate measurement of the folding landscape by using z_{sys} requires stiff handles.

Accurate Estimates of the Hopping Kinetics Requires Short and Flexible Handles. Because LOT experiments can also be used to measure the force-dependent rates of hopping between the NBA and the UBA, it is important to assess the influence of the dynamics of the handles on the intrinsic hopping kinetics of the RNA hairpin. In other words, how should the structural characteristics of the linkers be chosen so that the measured hopping rates using the time traces $z(t)$ and the intrinsic rates are as close as possible?

Folding and unfolding rates of P5GA and the free-energy profile without handles. We first performed force-clamp simulations of P5GA in the absence of handles to obtain the intrinsic (or ideal) folding (τ_F^o) or unfolding (τ_U^o) times, that serve as a reference for the H–RNA–H system. To obtain the boundary conditions for calculating the mean refolding and unfolding times, we collected the histograms of the time traces and determined the positions of the minima of the NBA and UBA, $z_F = 1.9$ nm, and $z_U = 7.4$ nm (Fig. 4*a*). The analysis of the time traces provides the transition times in which z_m reaches $z_m = z_U$ starting from $z_m = z_F$. The mean unfolding time τ_U is obtained by using either $\tau_U = 1/N\Sigma_i \tau_U(i)$ or from the fits to the survival probability $P_F(t) = e^{-t}/\tau_U$ (Fig. S3). The mean folding time is similarly calculated, and the two methods give similar results. The values of τ_U^o and τ_F^o computed from the time trace of $z_m(t)$ are 2.9 and 1.9 ms, respectively. At $f_m = 15.4$ pN and $L = 0$ nm, the equilibrium constant $K_{eq} = \tau_F^o/\tau_U^o = 0.67$, which shows that the bare molecular free energy is slightly tilted toward NBA at $f = 15.4$ pN.

Hopping times depend on the handle characteristics. The values of the folding (τ_F^m) and unfolding(τ_U^m) times were also calculated for the P5GA hairpin with attached handles (Fig. 1). As the length of the handles increase both τ_U^m and τ_F^m increase gradually, and the equilibrium distribution shifts toward the UBA, i.e., $K_{eq} = \tau_F^m/\tau_U^m$ increases (Fig. 4*b*). Strikingly, the use of flexible handles results in minimal deviations of τ_U^m and τ_F^m from their intrinsic values (Fig. 4*b*). Attachment of handles (stiff or flexible) to the 5′ and 3′ ends restricts their movement, which results in a decrease in the number of paths to the NBA and UBA. Thus, both τ_U^m and τ_F^m increase (Fig. 4*b*). As the stiffness of the handle increases the extent of pinning increases. These arguments show that flexible and short handles, that have the least restriction on the fluctuations of the 5′ and 3′ ends of the hairpin, cause minimal perturbation to the intrinsic RNA dynamics and, hence, the hopping rates. Because the experimentally accessible quantity is the extension of the H–RNA–H, it is important to show that the transition times can be reliably obtained by using $z_{sys}(t)$. Although $z_{sys}(t)$ differs from $z_m(t)$ in amplitude, the "phase" between the two quantities track each other reliably throughout the simulation, even when the handles are long and flexible (see Fig. S2). We calculated τ_U^{sys} and τ_F^{sys} by analyzing the trajectories $z_{sys}(t)$ using the same procedure used to compute their intrinsic values. Comparison of τ_U^{sys} (τ_F^{sys}) and τ_U^m (τ_F^m) for both stiff and

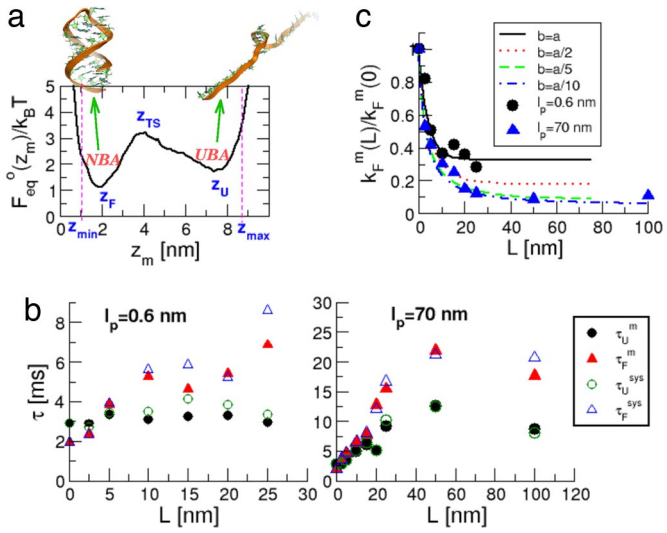

Fig. 4. Hopping rates from free energy profile. (*a*) The free-energy profile for P5GA with $L = 0$ nm. (*b*) The transition times at $f = f_m$, obtained by using $z_m(t)$ (filled symbols) and $z_{sys}(t)$ (open symbols). The ratio $\tau^m_{U(F)}/\tau^{sys}_{U(F)} \approx 1$, which shows that $z_{sys}(t)$ mirrors the hopping of P5GA. (*c*) Folding rate $k^m_F(L)/k^m_F(0)$ as a function of L for varying b, by using the GRM. The plots show $b/a = 1$, 1/2, 1/5, and 1/10. The simulation results for P5GA are also shown as symbols to emphasize that the GRM accounts for the hopping kinetics in the H–RNA–H system accurately.

flexible handles shows excellent agreement at all L values (Fig. 4*b*). Thus, it is possible to infer the RNA dynamics $z_m(t)$ by measuring $z_{sys}(t)$.

Theoretical predictions using the GRM are consistent with the simulations. The simulation results can be fully understood by using the GRM (Fig. 3*a*), for which we can exactly solve the overdamped Langevin equation using the discrete representation of the Gaussian chain (see *Methods*). By assuming that transverse fluctuations are small (which is reasonable under the relatively high tension of $f = 15.4$ pN), we use the Wilemski and Fixman (WF) theory (36) to determine an approximate time of contact formation ($\tau^m_F = (k^m_F)^{-1}$) as a function of b (i.e., increasing handle stiffness) and N_h. The refolding rate of the RNA hairpin under tension is analogous to k^m_F. A plot of $k^m_F(L)/k^m_F(0)$ versus L (Fig. 4*c*) illustrates that smaller deviations from the handle-free values occur when l_p is small. Moreover, Fig. 4*c* shows that the refolding rate decreases for increasing N_h regardless of the stiffness of the chain. The saturating value of k^m_F as $N_h \to \infty$ depends on b, with stiffer handles having a much larger effect on the folding rate. Although the handles used in LOT experiments are significantly longer than the handle lengths considered here, the saturation of the folding rate suggests that $L \approx 100$ nm is sufficiently long for finite-size effects to be negligible. We also find the dependence of k_F on L agrees well with the behavior observed in the simulation of P5GA. The ratio $k^m_F(L)/k^m_F(0)$ for $b = a$ agrees well with the trends of the flexible linker ($l_p = 0.6$ nm) for all of the simulated lengths, with both saturating at $k_F(L) \approx 0.35k_F(0)$ for large L. The trends for stiffer chains (smaller b) in the GRM qualitatively agree with the P5GA simulation with stiff handles ($l_p = 70$ nm), with remarkably good agreement for $0.1 \leq b/a \leq 0.2$ over the entire range of L. The GRM, which captures the physics of both the equilibrium and kinetic properties of the more complicated H–RNA–H, provides a theoretical basis for extracting kinetic information from experimentally (or computationally) determined folding landscapes.

Free-Energy Landscapes and Hopping Rates. Stiff handles are needed to obtain $F_{eq}(z_{sys})$ (15) that resembles $F^o_{eq}(z_m)$, whereas

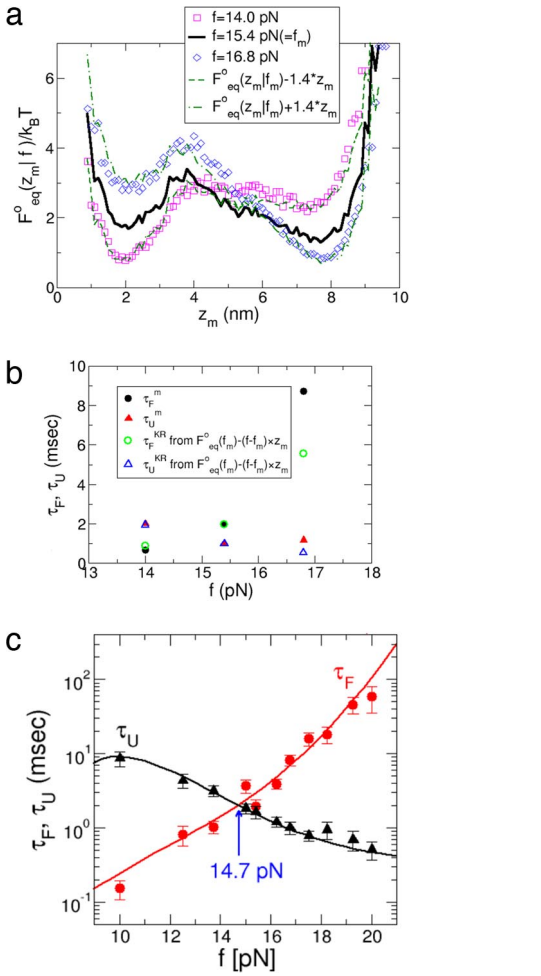

Fig. 5. Force-dependent hopping rates from free energy profile at $f = f_m$. (*a*) Comparison of the measured free-energy profiles (symbols) with the shifted free-energy profiles $\beta F^o_{eq}(z_m|f_m) - \beta(f - f_m)\cdot z_m$. (*b*) Folding and unfolding times as a function of force $f = 14$ pN $< f_m$, $f = 15.4$ pN $\approx f_m$, and $f = 16.8$ pN $> f_m$. $\tau^m_{F(U)}$ is obtained from the time trace in figure 2*B* in ref 35 at each force, whereas $\tau^m_{F(U)}$ is computed by using the tilted profile $\beta F^o_{eq}(z_m|f) = \beta F^o_{eq}(z_m|f_m) - \beta(f - f_m)\cdot z_m$ in equation (1). (*c*) Folding and unfolding times using the GRM. Symbols are a direct simulation of the GRM (error bars are standard deviation of the mean). The solid lines are obtained by using the Kramers theory (Eq. **1**). We choose $D_U \approx 3D_0$, so that that the simulated and Kramers times agree at $f = f_m$. The position of each basin of attraction as a function of force for the GRM is given by $z_U \approx N_0a^2\,\beta f/3$ and $z_F \approx N_0a^2\,\beta f/(3 + 2N_0a^2\beta k)$.

the flexible handles produce hopping rates that are close to their handle-free values. These two findings appear to demand two mutually exclusive requirements in the choice of the handles in LOT experiments. However, if z_m is a good reaction coordinate, then it should be possible to extract the hopping rates by using accurately measured $F_{eq}(z_{sys})$ [$\approx F_{eq}(z_m) \approx F^o_{eq}(z_m)$] at $f \approx f_m$, by using handles with small L/l_p. The (un)folding times can be calculated by using the mean first passage time (Kramers' rate expression) with appropriate boundary conditions (37),

$$\tau^{KR}_U = \int_{z_F}^{z_U} dy e^{\beta F_{eq}(y)} \frac{1}{D_U} \int_{z_{min}}^{y} dx e^{-\beta F_{eq}(x)}, \qquad [1]$$

$$\tau^{KR}_F = \int_{z_F}^{z_U} dy e^{\beta F_{eq}(y)} \frac{1}{D_F} \int_{y}^{z_{max}} dx e^{-\beta F_{eq}(x)},$$

where z_{min}, z_{max}, z_U, and z_F are defined in Fig. 4*a*. The effective diffusion coefficient $D_F(D_U)$ is obtained by equating τ_F^{KR} (τ_U^{KR}) in Eq. **1**, with $F_{eq}(z_m) = F^o_{eq}(z_m)$, to the simulated τ_F^o (τ_U^o). We calculated the *f*-dependent $\tau_U^m(f)$ and $\tau_F^m(f)$ by evaluating Eq. **1** using $F^o_{eq}(z_m|f) = F^o_{eq}(z_m|f_m) - (f - f_m)\cdot z_m$. The calculated and simulated results for P5GA are in good agreement (Fig. 5 *a* and *b*). At the higher force ($f = 16.8$ pN), the statistics of hopping transitions within our simulation time is not sufficient to establish ergodicity. As a result, the simulation results are not as accurate at high forces (see Fig. S4). To further show that the use of $F^o_{eq}(z_m|f)$ in Eq. **1** gives accurate hopping rates, we calculated $\tau_U^o(f)$ for the GRM and compared the results with direct simulations of the handle-free GRM, which allows the study of a wider range of forces (see *Methods*). The results in Fig. 5*c* show that $F^o_{eq}(z_m|f)$ indeed gives very accurate values for the transition times from the UBA and NBA over a wide force range.

Conclusions

The self-assembly of RNA and proteins may be viewed as a diffusive process in a multidimensional folding landscape. To translate this physical picture into a predictive tool, it is important to discern a suitable low-dimensional representation of the complex energy landscape, from which the folding kinetics can be extracted. Our results show that, in the context of nucleic acid hairpins, precise measurement of the sequence-dependent folding landscape of RNA is sufficient to obtain good estimates of the *f*-dependent hopping rates in the absence of handles. It suffices to measure $F_{eq}(z_{sys}) \approx F_{eq}(z_m) \approx F^o_{eq}(z_m)$ at $f = f_m$ by using stiff handles, whereas $F_{eq}(z_m|f)$ at other values for *f* can be obtained by tilting $F_{eq}(z_m|f_m)$. The accurate computation of the hopping rates by using $F_{eq}(z_m)$ show that z_m is an excellent reaction coordinate for nucleic acid hairpins under tension. Further theoretical and experimental work is needed to test whether the proposed framework can be used to predict the force-dependent hopping rates for other RNA molecules that fold and unfold through populated intermediates.

Methods

RNA Hairpin. The Hamiltonian for the RNA hairpin with *N* nucleotides, which is modeled by using the SOP model (35), is

$$H_{SOP} = -\frac{kR_0^2}{2}\sum_{i=1}^{N-1}\log\left(1-\frac{(r_{i,i+1}-r^o_{i,i+1})^2}{R_0^2}\right) + \sum_{i=1}^{N-3}\sum_{j=i+3}^{N}\varepsilon_h\left[\left(\frac{r^o_{i,j}}{r_{i,j}}\right)^{12} - 2\left(\frac{r^o_{i,j}}{r_{i,j}}\right)^{6}\right]\Delta_{i,j} + \sum_{i=1}^{N-3}\sum_{j=i+3}^{N}\varepsilon_l\left(\frac{\sigma}{r_{i,j}}\right)^6(1-\Delta_{i,j}) + \sum_{i=1}^{N-2}\varepsilon_l\left(\frac{\sigma^*}{r_{i,i+2}}\right)^6, \quad [2]$$

where $r_{i,j} = |\mathbf{r}_i - \mathbf{r}_j|$ and $r^o_{i,j}$ is the distance between monomers *i* and *j* in the native structure. The first term enforces backbone chain connectivity by using the finite extensible nonlinear elastic (FENE) potential, with $k \approx 1.4 \times 10^4$ pN·nm^{-1} and $R_0 = 0.2$ nm. The Lennard–Jones interaction (second term in Eq. **2**) describes interactions only between native contacts (defined as $r^o_{i,j} \leq 1.4$ nm for $|i - j| > 2$), with $\Delta_{i,j} = 1$ if monomers *i* and *j* are within 1.4 nm in the native state, and $\Delta_{i,j} = 0$ otherwise. Nonnative interactions are treated as purely repulsive (the third term in Eq. **2**) with $\sigma = 0.7$ nm. We take $\varepsilon_h = 4.9$ pN·nm and $\varepsilon_l = 7.0$ pN·nm for the strength of interactions. In the fourth term, the repulsion between the *i*th and $(i + 2)$th interaction sites along the backbone has $\sigma^* = 0.35$ nm to prevent disruption of the native helical structure.

Handle Polymers. The handles are modeled by using the Hamiltonian

$$H_{handles} = \frac{k_S}{2}\sum_{i=1}^{N-1}(r_{i,i+1} - r_0)^2 - k_A\sum_{i=1}^{N-2}\hat{\mathbf{r}}_{i,i+1}\cdot\hat{\mathbf{r}}_{i+1,i+2}. \quad [3]$$

The neighboring interaction sites, with an equilibrium distance $r_0 = 0.5$ nm, are harmonically constrained with a spring constant $k_S \approx 1.4 \times 10^4$ pN·nm^{-1}. In the second term of Eq. **3**, the strength of the bending potential, k_A, determines the handle flexibility. We choose two values, $k_A = 7.0$ pN·nm and $k_A = 561$ pN·nm to model flexible and semiflexible handles respectively and assign $k_A = 35$ pN·nm to the junction connecting two ends of the RNA and the handles. We determine the corresponding persistence length for the two k_A values as $l_p = 0.6$ and 70 nm (see *SI Text*). The contour length of each handle is varied from $n = 5$–200.

GRM. The Hamiltonian for the GRM (Fig. 3*a*) is

$$\beta H = \frac{3}{2b^2}\sum_{i=1}^{N_h}(\mathbf{r}_{i+1} - \mathbf{r}_i)^2 + \frac{3}{2b^2}\sum_{i=N_h+N_0+1}^{2N_h+N_0}(\mathbf{r}_{i+1} - \mathbf{r}_i)^2 - \beta\mathbf{f}\cdot(\mathbf{r}_N - \mathbf{r}_1) + \beta k_0\mathbf{r}_1^2 + \frac{3}{2a^2}\sum_{i=N_h+1}^{N_h+N_0}(\mathbf{r}_{i+1} - \mathbf{r}_i)^2 + \beta V[\mathbf{r}_{N-N_h+1} - \mathbf{r}_{N_h+1}], \quad [4]$$

where

$$V[\mathbf{r}] = \begin{cases} k\mathbf{r}^2 & |\mathbf{r}| \leq c \\ kc^2 & |\mathbf{r}| > c \end{cases}. \quad [5]$$

The first two terms in Eq. **4** are the discrete connectivity potentials for the two handles, each with N_h bonds ($N_h + 1$ monomers), and with Kuhn length *b*. The mechanical force **f** in the third term is applied along the *z* direction, with $|\mathbf{f}| = f_m = 15.4$ pN. We also fix the 5′ end of the system with a harmonic bond of strength $k_0 = 2.5 \times 10^4$ pN·nm^{-1} in the fourth term of Eq. **4**. The fifth term mimics the RNA hairpin with N_0 bonds and spacing $a = 0.5$ nm. Interactions between the two ends of the RNA hairpin are modeled as harmonic bond with strength $k \approx 0.54$ pN·nm^{-1} that is cut off at $c = 4$ nm (Eq. **5**). When $|\mathbf{R}_m| > 4$ nm, the bond is broken, mimicking the unfolded state.

The free energies as a function of both $R_m \approx z_m$ and $R_{sys} \approx z_{sys}$ are most easily determined in the continuum limit of the Hamiltonian in Eq. **4**, with $\sum_{i=1}^{N} \to \int_0^n ds$. Because of the relatively large value of the external tension ($f_m \gg k_BT/l_p$), we can neglect transverse fluctuations without significantly altering the equilibrium or kinetic properties of the GRM. The refolding time, τ_F^m of the RNA mimic (Fig. 3*a*), which is the WF closure time (36), can be determined by numerically diagonalizing the Rouse-like matrix with elements (38) $\mathbf{M}_{ij} = \frac{1}{2}\delta^2 H/\delta\mathbf{r}_i\delta\mathbf{r}_j$.

ACKNOWLEDGMENTS. We are grateful to Arthur Laporta and N. Toan for useful discussions. This work was supported in part by National Science Foundation Grant CHE 05-14056. C.H. is supported by the Korea Science and Engineering Foundation (KOSEF) grant (No. R01-2008-000-10920-0).

1. Fersht A (1998) *Structure and Mechanism in Protein Science: A Guide to Enzyme Catalysis and Protein Folding* (Freeman, New York).
2. Doudna JA, Cech TR (2002) The chemical repertoire of natural ribozymes. *Nature* 418:222–228.
3. Dobson CM (1999) Protein misfolding, evolution and disease. *Trends Biochem Sci* 24:329–332.
4. Dill KA, Chan HS (1997) From Levinthal to pathways to funnels. *Nat Struct Biol* 4:10–19.
5. Onuchic JN, Wolynes PG (2004) Theory of protein folding. *Curr Opin Struct Biol* 14:70–75.
6. Thirumalai D, Hyeon C (2005) RNA and protein folding: Common themes and variations. *Biochemistry* 44:4957–4970.
7. Fisher TE, Marszalek PE, Fernandez JM (2000) Stretching single molecules into novel conformations using the atomic force microscope. *Nat Struct Biol* 7:719–724.
8. 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.
9. Rhoades E, Gussakovsky E, and Haran G (2003) Watching proteins fold one molecule at a time. *Proc Natl Acad Sci* 100:3197–3202.
10. Li PTX, Collin D, Smith SB, Bustamante C, Tinoco I, Jr (2006) Probing the mechanical folding kinetics of TAR RNA by hopping, force-jump, and force-ramp methods. *Biophys J* 90:250–260.
11. Schuler B, Lipman EA, Eaton WA (2002) Probing the free-energy surface for protein folding with single-molecule fluorescence spectroscopy. *Nature* 419:743–747.
12. Zhuang X, Rief M (2003) Single-molecule folding. *Curr Opin Struct Biol* 13:86–97.
13. Merkel R, Nassoy P, Leung A, Ritchie K, Evans E (1999) Energy landscapes of receptor–ligand bonds explored with dynamic force spectroscopy. *Nature* 397:50–53.

14. Woodside MT, *et al.* (2006) Nanomechanical measurements of the sequence-dependent folding landscapes of single nucleic acid hairpins. *Proc Natl Acad Sci*, 103:6190–6195.
15. Woodside MT, *et al.* Direct measurement of the full, sequence-dependent folding landscape of a nucleic acid. *Science* 314:1001–1004.
16. Li PTX, Bustamante C, Tinoco I, Jr (2007) Real-time control of the energy landscape by force directs the folding of RNA molecules. *Proc Natl Acad Sci* 104:7039–7044.
17. Dietz H Rief M (2004) Exploring the energy landscape of GFP by single-molecule mechanical experiments. *Proc Natl Acad Sci*, 101:16192–16197.
18. Mickler M, *et al.* (2007) Revealing the bifurcation in the unfolding pathways of GFP by using single-molecule experiments and simulations. *Proc Natl Acad Sci* 104:20268–20273.
19. Hyeon C, Thirumalai D (2003) Can energy landscape roughness of proteins and RNA be measured by using mechanical unfolding experiments? *Proc Natl Acad Sci*, 100:10249–10253.
20. Nevo R, Brumfeld V, Kapon R, Hinterdorfer P, Reich Z (2005) Direct measurement of protein energy landscape roughness. *EMBO Rep* 6:482.
21. Schlierf M, Li H, Fernandez JM (2004) The unfolding kinetics of ubiquitin captured with single-molecule force-clamp techniques. *Proc Natl Acad Sci*, 101:7299–7304.
22. Brujić J, Hermans RI, Walther KA, Fernandez JM (2006) Single-molecule force spectroscopy reveals signatures of glassy dynamics in the energy landscape of ubiquitin. *Nat Phys* 2:282–286.
23. Onoa B, *et al.* (2003) Identifying kinetic barriers to mechanical unfolding of the *T. thermophila* ribozyme. *Science* 299:1892–1895.
24. Klimov DK, Thirumalai D (1999) Stretching single-domain proteins: Phase diagram and kinetics of force-induced unfolding. *Proc Natl Acad Sci* 96(11):6166–6170.
25. Hyeon C, Thirumalai D (2005) Mechanical unfolding of RNA hairpins. *Proc Natl Acad Sci*, 102:6789–6794.
26. Manosas M, Collin D, Ritort F (2006) Force-dependent fragility in RNA hairpins. *Phys Rev Lett* 96:218301.
27. West DK, Brockwell DJ, Olmsted PD, Radford SE, Paci E (2006) Mechanical resistance of proteins explained using simple molecular models. *Biophys J* 90:287–297.
28. Dudko OK, Hummer G, Szabo A (2006) Intrinsic rates and activation free energies from single-molecule pulling experiments. *Phys Rev Lett* 96:108101.
29. Hyeon C, Thirumalai D. (2007) Measuring the energy landscape roughness and the transition state location of biomolecules using single molecule mechanical unfolding experiments. *J Phys Condens Matter* 19:113101.
30. Manosas M, Ritort F (2005) Thermodynamic and kinetic aspects of RNA pulling experiments. *Biophys J* 88:3224–3242.
31. Hyeon C, Thirumalai D (2006) Forced-unfolding and force-quench refolding of RNA hairpins. *Biophys J* 90:3410–3427.
32. Manosas M, *et al.* (2007) Force unfolding kinetics of RNA using optical tweezers. II. Modeling experiments. *Biophys J* 92:3010–3021.
33. West DK, Paci E, Olmsted PD (2006) Internal protein dynamics shifts the distance to the mechanical transition state. *Phys Rev E* 74:061912.
34. Svoboda K, Block SM (1994) Biological applications of optical forces. *Annu Rev Biophys Biomol Struct* 23:247–285.
35. Hyeon C, Thirumalai D (2007) Mechanical unfolding of RNA : From hairpins to structures with internal multiloops. *Biophys J* 92:731–743.
36. Wilemski G, Fixman M (1974) Diffusion-controlled intrchain reactions of polymers. II Results for a pair of terminal reactive groups. *J Chem Phys* 60:878–890.
37. Zwanzig R (2001) *Nonequilibrium Statistical Mechanics* (Oxford Univ Press, New York).
38. Doi M, Edwards SF (1988) *The Theory of Polymer Dynamics* (Clarendon, Oxford).