Extracting Kinetics from Single-Molecule Force Spectroscopy: Nanopore Unzipping of DNA Hairpins

Olga K. Dudko,* Jérôme Mathé,† Attila Szabo,‡ Amit Meller,§ and Gerhard Hummer‡
*Mathematical and Statistical Computing Laboratory, Division of Computational Bioscience, Center for Information Technology, and ‡Laboratory of Chemical Physics, National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health, Bethesda, Maryland USA; †Laboratoire MPI, Université d'Evry Val d'Essonne, Evry Cedex, France; and §Department of Physics and Biomedical Engineering, Boston University, Boston, Massachusetts

ABSTRACT Single-molecule force experiments provide powerful new tools to explore biomolecular interactions. Here, we describe a systematic procedure for extracting kinetic information from force-spectroscopy experiments, and apply it to nanopore unzipping of individual DNA hairpins. Two types of measurements are considered: unzipping at constant voltage, and unzipping at constant voltage-ramp speeds. We perform a global maximum-likelihood analysis of the experimental data at low-to-intermediate ramp speeds. To validate the theoretical models, we compare their predictions with two independent sets of data, collected at high ramp speeds and at constant voltage, by using a quantitative relation between the two types of measurements. Microscopic approaches based on Kramers theory of diffusive barrier crossing allow us to estimate not only intrinsic rates and transition state locations, as in the widely used phenomenological approach based on Bell's formula, but also free energies of activation. The problem of extracting unique and accurate kinetic parameters of a molecular transition is discussed in light of the apparent success of the microscopic theories in reproducing the experimental data.

INTRODUCTION

In single-molecule (SM) experiments forces can be exerted directly on individual molecules and their response can be followed as a function of time. These experiments are beginning to reveal fundamentally novel and unique information on the structure, dynamics, and interactions of individual biomolecules. The unfolding of nucleic acids, for instance, has been probed extensively using optical tweezers (1,2). Nanopore force spectroscopy, unlike other single-molecule methods, is a linker-free method to probe DNA-DNA and DNA-protein interactions. Instead of grafting the polynucleotides to macroscopic beads or surfaces, an electric field is used to drive the charged polymer through a nanopore (3–7). Whereas the experimental observables in SM experiments are typically clean and unambiguous, their interpretation in terms of the underlying molecular interactions and structures often turns out to be challenging. The phenomenological approach pioneered by Evans and co-workers (8) offers a simple framework for the interpretation of SM results, based on the assumption that the external force accelerates the rupture rate exponentially (9), $k(F) = k_0 e^{\beta F x^{\ddagger}}$, where $\beta = 1/k_B T$ with k_B being Boltzmann's constant and T the absolute temperature. This approach is the most commonly used procedure to extract the intrinsic rate k_0 (i.e., the rate at zero force) and the transition state location $x^{\ddagger}$ from SM force experiments. The Bell relation is valid only for low forces, and is based on the assumption that any applied force F reduces the instantaneous barrier height by the quantity $Fx^{\ddagger}$ and thus does not alter the location of the transition state (10). However, as the force increases, the distance between the minimum and the transition state in any one-dimensional potential necessarily decreases, thus weakening the dependence of the rupture rate on the force. This force-induced shift in the transition state has recently been called the ''Hammond effect'' by Thirumalai and co-workers (11,12). By using Kramers theory (13) of diffusive barrier crossing in the presence of force (14–16), alternative microscopic models have been formulated that remove the limitations of the phenomenological Bell relation (17,18). These microscopic theories still lead to analytical expressions for the rate of rupture at constant force and for the rupture force distribution in the presence of time-varying external forces. The resulting expressions are only slightly more complex than those in the phenomenological theory.

In a recent article (19), a unified approach was developed that casts the phenomenological and microscopic theories into a common framework. To illustrate the use of this approach and to establish a systematic procedure for extracting kinetic information from SM force experiments, we here analyze an extensive set of DNA unzipping measurements (7,20). In these experiments, individual DNA hairpins were unzipped by threading them through a nanopore, either at constant voltage or using voltage ramps at constant speeds. To extract the parameters of the theoretical models, we perform a global maximum-likelihood (ML) analysis of experimental data for low-to-intermediate ramp speeds. With the ML estimates, we explore to what degree the different theories can reproduce the data from the two types of measurements both in the regime used in the fit and outside that regime. In addition, we use the data to test a recently proposed quantitative relation between the constant voltage and constant ramp speed measurements (19). A critical comparison of the different

Submitted December 8, 2006, and accepted for publication January 31, 2007.
Address reprint requests to Gerhard Hummer, E-mail: gerhard.hummer@nih.gov.

0006-3495/07/06/4188/08 $2.00
doi: 10.1529/biophysj.106.102855

theories, and of earlier analyses of the data (20) highlights the challenges in extracting unique and accurate kinetic parameters from SM force experiments.

EXPERIMENT

The experimental procedures and materials were described in detail in (7,20). Briefly, PAGE-purified ssDNA and DNA hairpins (Eurogentec, San Diego, CA) were buffered in 10 mM Tris, 1 mM EDTA, pH 8.5 solution and, before the measurements, were heated to 75°C for 10 min, and then quenched to 4°C. The hairpin molecules consisted of a 3′ single-stranded overhang (50 mer poly-dA), and a 10-bp helical part containing an intervening six-base loop with the following sequence (the self-complementary parts are underlined):

HP1: 5′-GCTCTGTTGCTCTCTCGCAACAGAGC$(A)_{50}$.

The basic apparatus and experimental method used for reconstituting the α-HL channel in a horizontally supported planar lipid bilayer has been described previously (21). The temperature of the system was maintained at 15.0 ± 0.1°C, using a custom cell design (22). The buffer solution was 1 M KCl, 10 mM Tris-Hcl, with a pH of 8.5. The ion current was measured using a patch-clamp amplifier (Axopatch 200B, Axon Instruments, Union City, CA) and the signal was filtered using a 100-kHz low-pass four-poles Butterworth filter (Krohn Hite 3302, Avon, MA). The signal was digitized at 1 MHz/12 bits using a DAQ card. Our apparatus incorporates a feedback loop used to control the applied transmembrane voltage, as described earlier (22). The response time of the membrane potential to a step in the control voltage was 4 ± 1 μs. In each experiment (performed at given conditions set by the voltage or the voltage ramp) we typically collected over 1000 unzipping events. Our software and hardware combination permits high-throughput unattended data acquisition, such that the total acquisition time for each experiment was ~10 min.

THEORY

We now briefly describe a recently introduced approach (19) to analyzing SM force experiments that casts the microscopic and the phenomenological theories into a unified framework. Voltage-driven unzipping of a DNA hairpin in the nanopore is viewed as an irreversible molecular transition induced by a force F, during which the molecule moves along the reaction coordinate x on a free-energy surface, $U(x) = U_0(x) - Fx$. The bare free-energy surface $U_0(x)$ is assumed to have a single well, a barrier at a distance $x^\ddagger$ from the well center, and an activation free energy $\Delta G^\ddagger$. The initial state, in the well of the free energy surface, represents the single-stranded overhang of DNA threaded into the pore with the hairpin closed (Fig. 1 *a*). Escape over the barrier involves unzipping

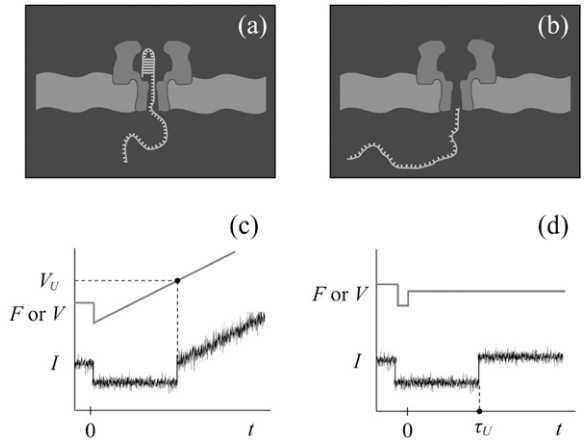

FIGURE 1 Voltage-induced nanopore unzipping of individual DNA hairpin molecules. (*a*) Single-stranded DNA overhang threaded into the pore constriction with the folded hairpin trapped in the pore vestibule. (*b*) The double-stranded part of the DNA is unzipped after voltage-driven passage through the pore. (*c*) Unzipping at a constant voltage ramp speed $\dot{V}$: applied force F (or voltage V) and resulting pore current I as a function of time. The unzipping voltage V_U is obtained from the jump in ionic current from the ''blocked-pore level'' to the ''open-pore level''. A collection of unzipping events at different loading rates $\dot{V}$ yields a distribution $p(V_U/\dot{V})$. (*d*) Unzipping at constant force F (or voltage V): the unzipping time τ_U is obtained from the abrupt increase in the pore current. The characteristic unzipping rate $k(V)$ is obtained as the inverse of the average unzipping time for constant voltage V.

of the double-stranded part of the DNA and the pore being cleared (Fig. 1 *b*).

We first consider the case of a constant force accelerating the rate of molecular rupture (Fig. 1 *d*). The theory (19) predicts that for a single-well free energy surface, the escape rate $k(F)$ for irreversible rupture under a constant external force F is

$$k(F) = k_0\left(1-\frac{\nu F x^\ddagger}{\Delta G^\ddagger}\right)^{1/\nu-1}\exp\left\{\beta\Delta G^\ddagger\left[1-\left(1-\frac{\nu F x^\ddagger}{\Delta G^\ddagger}\right)^{1/\nu}\right]\right\}. \quad (1)$$

Values $\nu = 2/3$ and $\nu = 1/2$ of the scaling parameter ν correspond to the linear-cubic potential $[U_0(x) = (3/2)\Delta G^\ddagger x/x^\ddagger - 2\Delta G^\ddagger(x/x^\ddagger)^3]$ and cusp potential $[U_0(x) = \Delta G^\ddagger(x/x^\ddagger)^2$ for $(x < x^\ddagger)$ and $-\infty$ for $(x \geq x^\ddagger)]$, respectively. For $\nu = 1$, or for $\Delta G^\ddagger \to \infty$ independent of ν, the phenomenological expression of Bell (9) is recovered from Eq. 1. For $\nu \neq 1$, this expression becomes invalid when F approaches the critical force $F_c = \Delta G^\ddagger/(\nu x^\ddagger)$ at which the barrier to rupture vanishes. This problem is caused by the use of the Kramers high-barrier approximation, and can be circumvented by using the full mean-first-passage times (MFPT) formula (15,23,24) to determine the force-dependent rate of escape from the well at x_{min} to a point x_M beyond the barrier, $k_{MFPT}(F) = D[\int_{x_{min}}^{x_M} dx e^{\beta[U_0(x)-Fx]}\int_{-\infty}^{x} dy e^{-\beta[U_0(y)-Fy]}]^{-1}$. In the high-barrier regime probed by low-to-intermediate forces F, the full MFPT expression for the force-dependent rate of

rupture is practically identical to the approximate Eq. 1. However, the use of $k_{\mathrm{MFPT}}(F)$ leads to substantial improvements at the highest ramp speeds where forces near F_c are probed.

When the force is ramped up linearly with time, $F(t) = K\upsilon t$ (Fig. 1 *c*), where υ is the pulling speed and $K\upsilon$ is the force loading rate, the distribution of forces at rupture is (19)

$$p(F|\upsilon) = (K\upsilon)^{-1}k(F)e^{\frac{k_0}{\beta x^{\ddagger}K\upsilon}}e^{-\frac{k(F)}{\beta x^{\ddagger}K\upsilon}\left(1-\frac{\nu F x^{\ddagger}}{\Delta G^{\ddagger}}\right)^{1-1/\nu}}, \qquad (2)$$

where $k(F)$ is the force-dependent rate of Eq. 1. For intermediate values of the speed υ, the approximate analytical expressions for the mean rupture force $\langle F\rangle = \int Fp(F|V)dF$ and the variance $\sigma_F^2 = \langle F^2\rangle - \langle F\rangle^2$ are

$$\langle F\rangle \cong \frac{\Delta G^{\ddagger}}{\nu x^{\ddagger}}\left\{1 - \left[\frac{1}{\beta\Delta G^{\ddagger}}\ln\frac{k_0 e^{\beta\Delta G^{\ddagger}+\gamma}}{\beta x^{\ddagger}K\upsilon}\right]^{\nu}\right\} \qquad (3)$$

$$\sigma_F^2 \cong \frac{\pi^2}{6(\beta x^{\ddagger})^2}\left[\frac{1}{\beta\Delta G^{\ddagger}}\ln\frac{k_0 e^{\beta\Delta G^{\ddagger}+\tilde{\gamma}}}{\beta x^{\ddagger}K\upsilon}\right]^{2\nu-2}. \qquad (4)$$

Here $\tilde{\gamma} = \gamma^2 - 3/\pi^2\psi''(1) \approx 1.064$, $\gamma \approx 0.577$, and $\psi''(1) \approx -2.404$ (25). When γ is formally set to zero, Eq. 3 is a good approximation for the maximum (mode) of the rupture force distribution. For $\nu = 2/3$, these expressions have the same υ dependence as those of Dudko et al. (18) which, however, involved the critical force at rupture and the diffusion coefficient rather than k_0 and $x.^{\ddagger}$ The theory in Eqs. 1–4 can be analytically continued to all ν, and thus ν can be used as an additional fitting parameter.

Implicit in the derivation of both phenomenological and microscopic models is the adiabatic assumption that the pulling speed is slow enough so that by the time the barrier is so low that Kramers theory is invalid, the survival probability is effectively zero (17). For a single-well potential, the survival probability then satisfies a first-order differential equation with a time-dependent rate, $dS(t)/dt = -k[F(t)]S(t)$. If this adiabatic approximation is indeed valid and $F(t) = K\upsilon t$, then the product $\upsilon\ln S[t(F)]$ as a function of F is independent of υ (26), where $t(F) = F/K\upsilon$. In this case, the following relation between the constant-force experiments (measuring $k(F)$) and constant speed experiments (measuring $p(F|\upsilon)$) has been established (19):

$$k(F) = \frac{K\upsilon p(F|\upsilon)}{1 - \int_0^F p(F'|\upsilon)dF'} \;\forall\upsilon. \qquad (5)$$

This equation relates the two kinds of experiments in a model-free way. It predicts that the rupture force distributions $p(F|\upsilon)$ at different ramp speeds can be collapsed onto a single curve for the force-dependent rate of molecular rupture, $k(F)$.

We note that the theory, Eqs. 1–5, developed here for molecular unfolding, or unbinding can be applied equally to the reverse process, folding or binding, in which the force is reduced with time. Both forward and reverse transitions have been observed, for instance, in mechanical unfolding experiments of RNA hairpins (2,27), and studies of the resulting hysteresis effects (28) should prove insightful.

In the nanopore unzipping experiments, the applied voltage V is analogous to the applied mechanical force F in pulling experiments (e.g., those using atomic force microscopes or optical tweezers). The voltage drop across the membrane-spanning nanopore results in an electric field that generates a mechanical force on the charged DNA strand threaded into the nanopore (Fig. 1 *a*). To adapt the above formalism to the nanopore-unzipping experiments, we follow Mathé et al. (7) and define the voltage $V^{\ddagger} = k_BT/Q_{\mathrm{eff}}$ as the characteristic of the transition state (which is V_β in their notation), where Q_{eff} is the effective charge of the DNA inside the pore. We can then use Eqs. 1–5 by replacing $\beta Fx^{\ddagger} \rightarrow V/V^{\ddagger}$, and $K\upsilon \rightarrow \dot{V}$, where $\dot{V}$ is the voltage ramp speed.

Maximum likelihood analysis of experimental data

Consider a series of constant voltage-ramp speed experiments at ramp speeds $\dot{V}_j (j = 1, \ldots, N)$. DNA unzipping is observed at different voltages V_{ij} ($i = 1, \ldots, M_j$). Our objective is to find the parameter set $\{V^{\ddagger}, \Delta G^{\ddagger}, k_0\}$ that maximizes the likelihood L of the $M_1 \times M_2 \times \ldots \times M_N$ measurements, and to explore whether the theories unified by Eqs. 1–5 with the ML estimates for $V^{\ddagger}, \Delta G^{\ddagger}$, and k_0 can accurately reproduce the data both for unzipping under constant voltage and unzipping under constant ramp speed.

For model parameters $\{V^{\ddagger}, \Delta G^{\ddagger}, k_0\}$ and ramp speeds $\dot{V}_j$, the likelihood L of a single observation of DNA unzipping at a voltage V_{ij} is given by Eq. 2, $L \equiv p(V_{ij}|V^{\ddagger}, \Delta G^{\ddagger}, k_0; \dot{V}_j)$. The likelihood of a series of such observations, assuming their statistical independence, is

$$L = \prod_{j=1}^{N}\left\{\prod_{i=1}^{M_j} p(V_{ij}|V^{\ddagger}, \Delta G^{\ddagger}, k_0; \dot{V}_j)\right\}, \qquad (6)$$

where the two products are over all values of the ramp speed $\dot{V}_j$, and the corresponding unzipping events V_{ij}.

The likelihood L assesses how well a given model describes the distributions of observables. By maximizing L with respect to model parameters (here, $V^{\ddagger}, \Delta G^{\ddagger}, k_0$), we seek to determine the optimal parameters of the models for different values of the scaling parameter ν. The corresponding maximum of L (or, equivalently, of the log-likelihood $\ln(L)$) then allows us to compare their relative performance. However, one has to take into account that models of higher complexity will tend to have larger scores just due to having more free parameters. Based on the Schwarz-Bayesian information criterion (29), one would expect an increase in the log-likelihood of $(\ln n)/2$ for every parameter added, with n being the number of measurements. Here, we use data from ~17,000 measurements for the ML fit, so that adding one parameter should increase the log-likelihood by ~5. A larger increase indicates a substantially better model.

Using the ML formalism, we globally fitted Eq. 2 to the voltage ramp data binned into histograms for loading rates of 12 V/s or less. Experimental data collected at higher ramp speeds were set aside for subsequent validation. Note that in some of the experiments at those highest ramp speeds, the DNA hairpin was still intact when the maximum voltage 0.2 V had been reached.

The likelihood function was maximized numerically with respect to the model parameters using simplex search and Monte Carlo methods. We also applied Bayesian inference to extract model parameters, assuming uniform prior distributions. All methods gave essentially the same results for the fitted model parameters.

RESULTS

ML estimates for the intrinsic rate coefficient, k_0, and the characteristic of the transition state, $V^\ddagger$, are listed in Table 1 for the linear-cubic theory ($\nu = 2/3$), cusp theory ($\nu = 1/2$), and phenomenological model ($\nu = 1$), as well as for the case when the exponent ν is used as an additional fitting parameter. Also listed are estimates for the free energy of activation, $\Delta G^\ddagger$, for the microscopic theories, and the maximum values of $\ln(L)$ (including, in parenthesis, the number of log-likelihood units gained or lost by a given model relative to the linear-cubic model). Parameters of the traditional phenomenological fit to the most probable unzipping voltage (7) are included for comparison.

As seen from Table 1, the two microscopic theories ($\nu = 2/3$ and $\nu = 1/2$) produce consistent estimates for the parameters $V^\ddagger, \Delta G^\ddagger, k_0$ and have similar likelihoods (with the cusp model having some advantage). Fitting the data with ν as an additional free parameter produces an optimal $\nu = 0.554$ intermediate between the linear-cubic and cusp values, but the corresponding gain of ~4 units in log-likelihood compared to the cusp model does not justify adding a fourth parameter, as assessed by the Schwarz-Bayesian information criterion. In contrast, as a result of the global ML fit, the phenomenological theory ($\nu = 1$) produces a dramatically lower (i.e., unfavorable) likelihood score that cannot be attributed to having only two instead of three free parameters. The value for the rate k_0 for the phenomenological theory is overestimated by an order of magnitude and the characteristic of the transition state, $V^\ddagger$, is double the typical number for the microscopic theories. Traditionally, the phenomenological model is fitted to the most probable unzipping voltage V_m as a function of ramp speed. Such a fit results in even larger estimates for k_0 and $V^\ddagger$, producing a rate k_0 that is more than an order of magnitude higher than those from the microscopic fit.

TABLE 1 Maximum-likelihood estimates of the kinetic parameters for nanopore unzipping of DNA

ν	k_0 (s^{-1})	$V^\ddagger$ (mV)	$\Delta G^\ddagger(k_BT)$	max{ln L}
2/3	0.12	12.7	10.5	39,989
1/2	0.05	9.9	11.9	40,037 (+48)
1	0.61	21.7	–	39,075 (−914)
1* (fit to V_m)	1.39	24.7	–	–
0.554 (floating)	0.06	10.5	11.5	40,041 (+52)

Estimates calculated from data at ramp speeds ≤12 V/s for the microscopic theories ($\nu = 2/3$, 1/2, and floating) and the phenomenological model ($\nu = 1$). *Traditional least-squares fit of the phenomenological model to the most probable unzipping voltage V_m for ramp speeds >1.6 V/s (7).

For the linear-cubic model, the rate k_0 and free energy barrier $\Delta G^\ddagger$ define a characteristic relaxation time in the folded well, $\tau_{rel} \approx \exp(-\beta\Delta G^\ddagger)/(2\pi k_0)$. It would be interesting to relate the resulting $\tau_{rel} \approx 40\,\mu$s to molecular processes such as basepair opening, with characteristic times of 1–50 ms in free solution (30). However, k_0 may contain contributions from processes other than unfolding, as discussed in the ''Concluding remarks''. We therefore caution against such detailed microscopic interpretations.

In Fig. 2, we compare the measured distributions of unzipping voltages to those obtained for the fitted theoretical models. We find that the microscopic theories reproduce the measured distributions of unzipping voltages very well in the regime used for the fit and, remarkably, even outside that regime.

The phenomenological model with parameter estimates obtained from the traditional fit to V_m turns out to be accurate in reproducing the experimental distributions at moderate ramp speeds, the regime used for fitting, whereas at low and high speeds the deviations are substantial (Fig. 2). We found that the phenomenological model with the estimates obtained from the global ML fit performs significantly better at low ramp speeds; however, the agreement at high speeds deteriorates further.

Fig. 3 shows that the most probable unzipping voltage V_m (i.e., the mode of the voltage distribution) from experiment depends nonlinearly on the logarithm of the voltage ramp speed. In contrast, the phenomenological theory (Eq. 3 with γ set to 0 and $\nu = 1$) predicts a linear dependence of the mode on the logarithm of speed. Within the framework of the phenomenological approach, one may attribute the observed curvature to additional molecular processes, such as hairpin rezipping or switching between multiple states. However, as shown in Fig. 3 the microscopic theories are fully consistent with this nonlinear dependence of the mode (and average) of the unzipping voltage distribution on the ramp speed, even in the regime not used for the fit. A single-well energy landscape thus appears to be adequate to explain the experimental data.

The inset in Fig. 3 shows that the variance σ_V^2 of the unzipping voltage distribution sharply increases with increasing ramp speed, in agreement with the microscopic theories, Eq. 4, fitted to ramp speeds ≤12 V/s. In contrast, the phenomenological theory predicts an essentially constant σ_V^2. The observed increase in the variance of the unzipping voltage (or force) at higher ramp speeds, thus, serves as a simple indicator that the phenomenological model cannot explain the data.

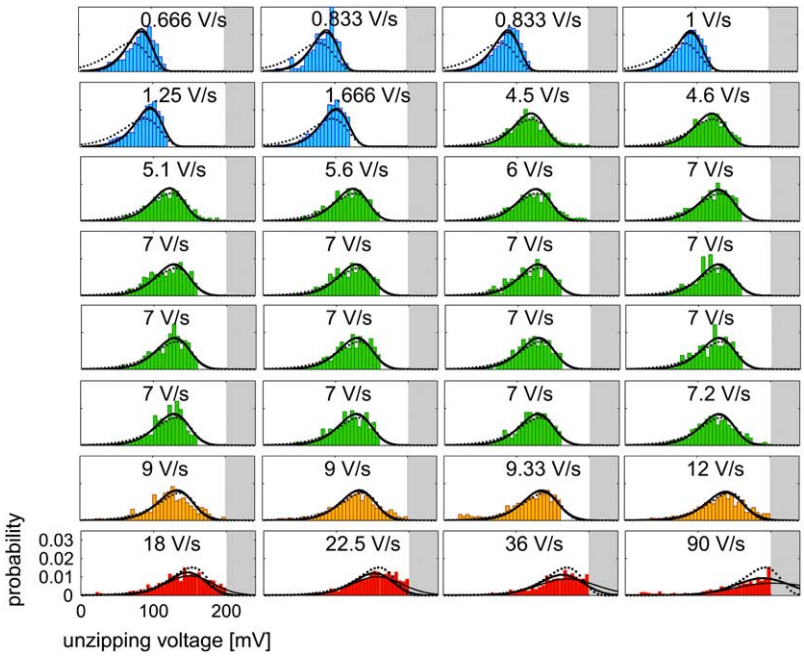

FIGURE 2 Distributions of the unzipping voltage from experiment (*histograms*) and theory, Eq. 2 (*lines*). (*Bold solid* and *fine solid lines*) Linear-cubic theory and cusp theory, respectively, with the parameters extracted from the global ML fit of the theories to experimental histograms obtained at speeds of 12 V/s or less. (*Dotted line*) Phenomenological theory with the parameters obtained from a traditional least-squares fit to the maximum of the unzipping voltage distributions for ramp speeds >1.6 V/s. Histograms at 22.5, 36, and 90 V/s were normalized assuming that 5, 20, and 40% of measurements, respectively, exceeded the 200 mV experimental cutoff. Colors reflect different ranges of the ramp speed and correspond to those in Figs. 3 and 4. Shaded areas indicate the range of experimentally inaccessible voltages.

To test whether the adiabatic approximation is valid for the experimental data set, and whether the constant voltage experiments are consistent with the voltage-ramp experiments, we transform the histograms from constant ramp-speed measurements according to Eq. 5. As confirmed in Fig. 4, the histograms collected over more than two orders of magnitude in the voltage ramp speed indeed collapse onto a single curve. That curve, in turn, matches the rates $k(V)$ obtained using constant voltage experiments (see Fig. 4). This collapse indicates that the two types of measurements (at constant voltage and at constant ramp speed) are fully consistent with each other. Data collected at different ramp speeds (colors correspond to those in Fig. 2) probe different ranges of the DNA unzipping rate $k(V)$. Taken together, they cover four orders of magnitude of the unzipping rate. We further observe in Fig. 4 that, although constant voltage data were not used in the fit, the data are actually predicted by Eq. 1, with the parameters obtained above from fitting to unzipping-voltage distributions in voltage ramp experiments.

Fig. 4 also shows that the experimental data from both types of measurements exhibit curvature in the voltage dependence of the logarithmic unzipping rate, $\ln[k(V)]$. This curvature is reproduced by the linear-cubic theory (Eq. 1 with $\nu = 2/3$) and cusp theory (Eq. 1 with $\nu = 1/2$). The curvature highlights the limitations of the Bell formula, which postulates a linear relation between $\ln(k)$ and the applied voltage, which in turn leads to a systematic overestimation of the intrinsic rates when used to fit data in the intermediate-to-high forces regime.

CONCLUDING REMARKS

We have illustrated and tested a recently proposed systematic procedure (19) for extracting kinetic information from SM force experiments, by analyzing an extensive set of experimental data for unfolding of DNA hairpins. Single DNA hairpin molecules were unzipped in a nanopore in two ways: under constant voltage and under voltage ramped up at a constant rate (7). We explored how well different theories reproduce the data, by performing ML global fits to the low-to-intermediate ramp speed data, as well as by examining the traditional phenomenological fit to the maxima of unzipping-voltage distributions. The fitted models were then validated by comparing their predictions to data collected at high ramp speeds and at constant voltage.

In these simplest models, the kinetics of force-induced rupture is described as an irreversible escape from a single well over a barrier in the potential of mean force. In the presence of a constant force F, the distribution of rupture times is $k(F)e^{-k(F)t}$. When $k(F)$ is given by Bell's formula, $\ln[k(F)]$ depends linearly on F. In a force-ramp experiment with constant ramp speed, the mean (and mode) of the rupture-force distribution will depend linearly on the logarithm of the force loading rate. More sophisticated microscopic theories of the force-dependent rupture rate predict that $\ln[k(F)]$ is a nonlinear function of F. This nonlinear dependence is a simple consequence of the fact that the distance to the transition state decreases as the force increases, eventually vanishing when the barrier disappears. Hence, in a constant-speed pulling experiment, the mean rupture force will depend nonlinearly on the logarithm of the loading rate.

A nonlinear dependence of the rupture force on the logarithmic ramp speed has been observed experimentally for the voltage-dependent unzipping of DNA hairpins (7). Here, we show that such data can be successfully described using a class of microscopic models of force-induced crossing of a single barrier. These models lead to simple analytic

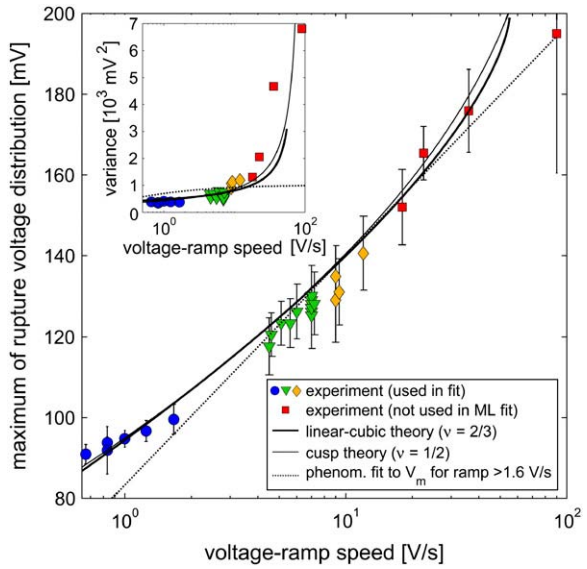

FIGURE 3 Dependence of the maximum (mode) of the unzipping voltage distribution on the voltage-ramp speed from experiment and theory (Eq. 3 with $\gamma = 0$). Model parameters for the linear-cubic and cusp theories were obtained from global ML fits. Parameters of the phenomenological model were obtained from a traditional least-squares fit (7) to the maximum V_m of unzipping voltage distributions for ramp speeds >1.6 V/s. The inset shows the dependence of the variance of the unzipping voltage on the voltage-ramp speed from experiment (*symbols*) and theory (*lines*). The microscopic theories (Eq. 4 with $\nu = 1/2$, *thin line*; and $\nu = 2/3$, *bold line*) show a sharp increase in the variance with increasing ramp speed. The variance of the phenomenological model (*dotted line*), evaluated directly from the distribution Eq. 2, shows a weak dependence on the ramp speed, and approaches the constant value at higher speeds in agreement with the approximate expression, Eq. 4, for $\nu = 1$. Color coding as in Fig. 2.

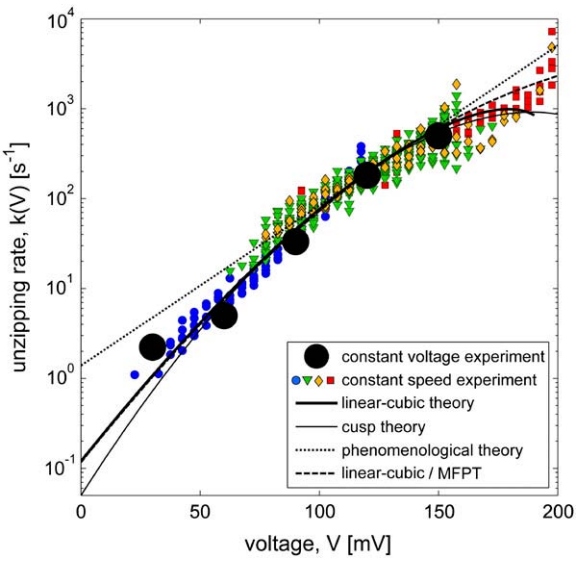

FIGURE 4 Quantitative relation between experiments at constant voltage and constant voltage ramp speed. Histograms from constant speed measurements are collapsed according to Eq. 5. The symbol colors indicate different ramp-speed ranges as in Fig. 2 (*blue spheres*, <2 V/s; *green triangles*, 2–8 V/s; *orange diamonds*, 8–15 V/s; *red squares*, >15 V/s). Constant voltage data are shown as large solid circles. The $k(V)$ curves of the theoretical models are shown as lines. Only histogram bins with high statistical significance (>10 counts) were used.

expressions for the distribution of rupture forces, which greatly facilitate the analysis of experimental rupture force distributions by enabling direct applications of Bayesian or ML methods. We note that the dominance of a single barrier in nanopore unzipping of DNA is consistent with the interpretation of pulling experiments on short nucleic-acid hairpins (1,2), despite the inherent richness of the calculated free energy landscape (27).

The microscopic theories considered here are arguably the simplest alternatives to the phenomenological approach and contain only one additional parameter (the free energy of activation, $\Delta G^{\ddagger}$). Despite their remarkable success in predicting the high ramp speed and constant voltage data, these models are not necessarily unique. In particular, it is always possible that more elaborate models with additional parameters capture the experimental data equally well or even better.

In general, more complex models involving, for instance, multiple states will not necessarily lead to single-exponential rupture time distributions in constant-force experiments. In this case, the rupture force histograms cannot be collapsed onto a master curve using the relation between the constant-force and constant-speed experiments, Eq. 5. Conversely, if the experimental data do collapse, as in this article, then any model for which $k(F)$ is consistent with the constant-force data will reproduce the rupture-force distributions at different force-loading rates.

In fact, a multistate model has been proposed by Mathé et al. (20). The following kinetic scheme was considered:

$$Z \underset{k_f e^{-V/V_f}}{\overset{k_u e^{V/V_u}}{\rightleftarrows}} U \xrightarrow{k_e e^{V/V_e}} exit, \tag{7}$$

where Z and U are the zipped and unzipped hairpin states, respectively, and k_e describes the escape of the unzipped DNA from the pore. The rate constant k_e and the rate constants k_f and k_u for folding and unfolding, respectively, depend on force according to Bell's formula with ''transition states'' V_e, V_f, and V_u. Now, if it is assumed that U is in steady state (20), then the distribution of DNA unzipping times will be single exponential involving the overall force-dependent rate of unzipping:

$$k(V) = \frac{k_u k_e e^{V/V_u}}{k_e + k_f e^{-V/V_{fe}}}, \tag{8}$$

where $1/V_{fe} = 1/V_f + 1/V_e$. If the escape rate is fast, this $k(V)$ reduces to Bell's formula; if escape is slow, it again reduces to a Bell-like formula, but with a different effective ''transition-state'' location. In the high-voltage limit, $k(V \to \infty) = k_u e^{V/V_u}$, and in the low-voltage limit $k(0) = k_u k_e/(k_e + k_f) = k_e/(k_e/k_u + \exp[\beta \Delta G])$, where $\beta \Delta G = \ln(k_f/k_u)$ is the equilibrium stability. This model contains an extra (fourth) parameter and

predicts a nonlinear dependence of $k(V)$ on V, even though the elementary rates are assumed to satisfy Bell's formula. Such a model can indeed fit the collapsed histogram data from ramp experiments, as well as constant voltage data (Fig. 5). If one directly fits $k(V)$ of Eq. 8 to the collapsed data, the fitted curve almost perfectly superimposes on that obtained from the single-state linear-cubic theory ($\nu = 2/3$), and extrapolates to a rate $k(0)$ of $\sim 0.1\ s^{-1}$. However, in this case the extrapolated $k(0)$ is not equal to the rate k_u of unzipping, but rather is essentially equal to the product of the escape rate and the equilibrium constant for folding.

It is somewhat disconcerting that two rather different theories can account for the data about equally well. Clearly, one cannot easily distinguish between different models and mechanisms on the basis of the SM experimental data alone. By applying ''Occam's razor'', one might favor the simpler single-state models; on the other hand, based on the physics of the problem, the multistate models may seem more appealing. Additional information can help in choosing the appropriate model. For example, we can compare estimates of the equilibrium stability of the DNA hairpin obtained by using the *mfold* server (31) (16.4 kcal/mol) with the activation free energy of unfolding estimated by the microscopic models ($\sim$7 kcal/mol; Table 1). Conversely, from the zero-force limit of the multistate model, we can find a lower bound for the DNA-exit rate from the pore at zero voltage, $k_e \geq e^{\beta\Delta G} k(0)$. For the *mfold* value of ΔG, and $k(0) \approx 0.1\ s^{-1}$ (Figs. 4 and 5), that leads to unrealistically large exit rates k_e exceeding $\sim 10^{11}\ s^{-1}$. Interestingly, if the ΔG value was consistent with the $\Delta G^{\ddagger} \sim 7$ kcal/mol activation free energy of our single-state models, then the problem with an unrealistic exit rate in the multistate model would also disappear. Such lowering of the free energy could arise, for instance, from interactions between the unfolded DNA with the pore. In light of these results, one might argue that both the kinetics and thermodynamics of unfolding of the DNA in the pore could indeed be substantially different from that in solution. Clearly, further experiments and simulations (32–39) are needed to address this interesting point.

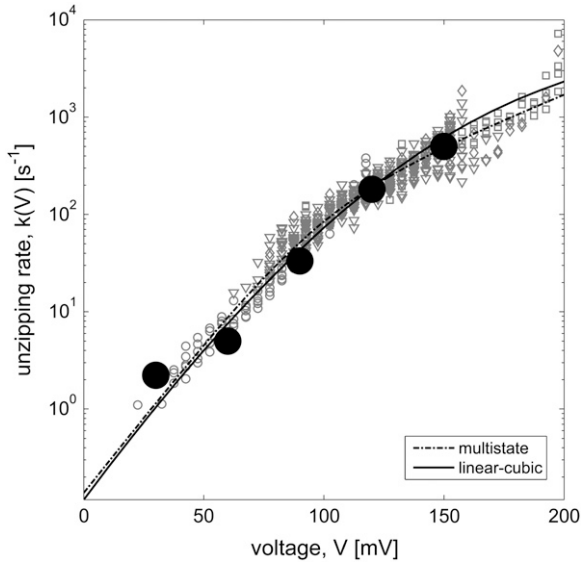

FIGURE 5 Multistate $k(V)$ of Eq. 8 fitted to collapsed unzipping-voltage histograms (as in Fig. 4). Also shown is the linear-cubic model. The fitted parameters of Eq. 8 are: $k_u = 15\ s^{-1}$, $k_f/k_e = 107$, $V_u = 42$ mV, and $V_{fe} = 21$ mV.

In summary, the excellent agreement between theory and experiment, even outside the regime used for the fit, demonstrates the significant progress that has been made in extracting information about the kinetics of molecular processes from SM data. Nevertheless, a complete understanding of the underlying molecular mechanisms will require a multifaceted approach and a critical analysis of all available data.

This research was supported by the Intramural Programs of the National Institute of Diabetes and Digestive and Kidney Diseases and Center for Information Technology, National Institutes of Health. A.M. acknowledges support from NIH awards HG003574 and GM072893.

REFERENCES

1. Liphardt, J., B. Onoa, S. B. Smith, I. Tinoco Jr., and C. Bustamante. 2001. Reversible unfolding of single RNA molecules using mechanical force. *Science*. 292:733–737.
2. Woodside, M. T., P. C. Anthony, W. M. Behnke-Parks, K. Larizadeh, D. Herschlag, and S. M. Block. 2006. Direct measurement of the full, sequence-dependent folding landscape of a nucleic acid. *Science*. 314: 1001–1004.
3. Kasianowicz, J. J., E. Brandin, D. Branton, and D. Deamer. 1996. Characterization of individual polynucleotide molecules using a membrane channel. *Proc. Natl. Acad. Sci. USA*. 93:13770–13773.
4. Akeson, M., D. Branton, J. J. Kasianowicz, E. Brandin, and D. Deamer. 1999. Microsecond time scale discrimination among polycytidylic acid, polyadenylic acid, and polyuridylic acid as homopolymers or as segments within single RNA molecules. *Biophys. J.* 77: 3227–3233.
5. Henrickson, S. E., M. Misakian, B. Robertson, and J. J. Kasianowicz. 2000. Driven DNA Transport into an asymmetric nanometer-scale pore. *Phys. Rev. Lett.* 85:3057–3060.
6. Meller, A., L. Nivon, E. Brandin, J. Golovchenko, and D. Branton. 2000. Rapid nanopore discrimination between single polynucleotide molecules. *Proc. Natl. Acad. Sci. USA*. 97:1079–1084.
7. Mathé, J., H. Visram, V. Viasnoff, Y. Rabin, and A. Meller. 2004. Nanopore unzipping of individual DNA hairpin molecules. *Biophys. J.* 87:3205–3212.
8. Evans, E., D. Berk, and A. Leung. 1991. Detachment of agglutinin-bonded red blood cells. I. Forces to rupture molecular-point attachments. *Biophys. J.* 59:838–848.
9. Bell, G. I. 1978. Models of the specific adhesion of cells to cells. *Science*. 200:618–627.
10. Dembo, M., D. C. Torney, K. Saxman, and D. Hammer. 1988. The reaction-limited kinetics of membrane-to-surface adhesion and detachment. *Proc. R. Soc. Lond. B Biol. Sci.* 234:55–83.
11. Hyeon, C. B., and D. Thirumalai. 2006. Forced-unfolding and force-quench refolding of RNA hairpins. *Biophys. J.* 90:3410–3427.
12. Koculi, E., D. Thirumalai, and S. A. Woodson. 2006. Counterion charge density determines the position and plasticity of RNA folding transition states. *J. Mol. Biol.* 359:446–454.
13. Kramers, H. A. 1940. Brownian motion in a field of force and the diffusion model of chemical reactions. *Phys.* 7:284–304.
14. Evans, E., and K. Ritchie. 1997. Dynamic strength of molecular adhesion bonds. *Biophys. J.* 72:1541–1555.

15. Izrailev, S., S. Stepaniants, M. Balsera, Y. Oono, and K. Schulten. 1997. Molecular dynamics study of unbinding of the avidin-biotin complex. *Biophys. J.* 72:1568–1581.

16. Shapiro, B. E., and H. Qian. 1997. A quantitative analysis of single protein-ligand complex separation with the atomic force microscope. *Biophys. Chem.* 67:211–219.

17. Hummer, G., and A. Szabo. 2003. Kinetics from nonequilibrium single-molecule pulling experiments. *Biophys. J.* 85:5–15.

18. Dudko, O. K., A. E. Filippov, J. Klafter, and M. Urbakh. 2003. Beyond the conventional description of dynamic force spectroscopy of adhesion bonds. *Proc. Natl. Acad. Sci. USA.* 100:11378–11381.

19. Dudko, O. K., G. Hummer, and A. Szabo. 2006. Intrinsic rates and activation free energies from single-molecule pulling experiments. *Phys. Rev. Lett.* 96:108101.

20. Mathé, J., A. Arinstein, Y. Rabin, and A. Meller. 2006. Equilibrium and irreversible unzipping of DNA in a nanopore. *Europhys. Lett.* 73:128–134.

21. Meller, A., L. Nivon, and D. Branton. 2001. Voltage-driven DNA translocations through a nanopore. *Phys. Rev. Lett.* 86:3435–3438.

22. Bates, M., M. Burns, and A. Meller. 2003. Dynamics of DNA molecules in a membrane channel probed by active control techniques. *Biophys. J.* 84:2366–2372.

23. Szabo, A., K. Schulten, and Z. Schulten. 1980. First passage time approach to diffusion controlled reactions. *J. Chem. Phys.* 72:4350–4357.

24. Shillcock, J., and U. Seifert. 1998. Escape from a metastable well under a time-ramped force. *Phys. Rev. E.* 57:7301–7304.

25. Abramowitz, M., and I. A. Stegun. 1972. Handbook of Mathematical Functions. Dover, New York.

26. Raible, M., M. Evstigneev, P. Reimann, F. W. Bartels, and P. Ros. 2004. Theoretical analysis of dynamic force spectroscopy experiments on ligand-receptor complexes. *J. Biotechnol.* 112:13–23.

27. Manosas, M., D. Collin, and F. Ritort. 2006. Force-dependent fragility in RNA hairpins. *Phys. Rev. Lett.* 96:218301.

28. Shapiro, B. E., and H. Qian. 1998. Hysteresis in force probe measurements: a dynamical systems perspective. *J. Theor. Biol.* 194:551–559.

29. Schwarz, G. 1978. Estimating the dimension of a model. *Ann. Statist.* 6:461–464.

30. Dornberger, U., M. Leijon, and H. Fritzsche. 1999. High base pair opening rates in tracts of GC base pairs. *J. Biol. Chem.* 274:6957–6962.

31. Zuker, M. 2003. Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Res.* 31:3406–3415.

32. Mathé, J., A. Aksimentiev, D. R. Nelson, K. Schulten, and A. Meller. 2005. Orientation discrimination of single-stranded DNA inside the alpha-hemolysin membrane channel. *Proc. Natl. Acad. Sci. USA.* 102:12377–12382.

33. Kirmizialtin, S., V. Ganesan, and D. E. Makarov. 2004. Translocation of a beta-hairpin-forming peptide through a cylindrical tunnel. *J. Chem. Phys.* 121:10268–10277.

34. Rief, M., and H. Grubmüller. 2002. Force spectroscopy of single biomolecules. *ChemPhysChem.* 3:255–261.

35. Flomenbom, O., and J. Klafter. 2003. Single stranded DNA translocation through a nanopore: a master equation approach. *Phys. Rev. E.* 68:041910.

36. Muthukumar, M., and C. Y. Kong. 2006. Simulation of polymer translocation through protein channels. *Proc. Natl. Acad. Sci. USA.* 103:5273–5278.

37. Lubensky, D. K., and D. R. Nelson. 1999. Driven polymer translocation through a narrow pore. *Biophys. J.* 77:1824–1838.

38. Berezhkovskii, A. M., and I. Gopich. 2003. Translocation of rodlike polymers through membrane channels. *Biophys. J.* 84:787–793.

39. Yeh, I.-C., and G. Hummer. 2004. Nucleic acid transport through carbon nanotube membranes. *Proc. Natl. Acad. Sci. USA.* 101:12177–12182.