Quantifying DNA Melting Transitions Using Single-Molecule Force Spectroscopy

Christopher P. Calderon[†,§,*], Wei-Hung Chen[¶,‡,*], Kuan-Jiuh Lin[‡], Nolan C. Harris [¶], and Ching-Hwa Kiang [¶]

[†] Department of Computational and Applied Mathematics, Rice University, Houston, TX, USA.

[§] Department of Statistics, Rice University, Houston, TX, USA.

[¶] Department of Physics and Astronomy, Rice University, Houston, TX, USA.

[‡] Department of Chemistry, National Chung Hsing University, Taichung, Taiwan, ROC.

[*] These two authors contributed equally

E-mail: chkiang@rice.edu

Abstract.

We stretched a DNA molecule using atomic force microscope and quantified the mechanical properties associated with B and S forms of double-stranded DNA (dsDNA), molten DNA, and single-stranded DNA (ssDNA). We also fit overdamped diffusion models to the AFM time series and used these models to extract additional kinetic information about the system. Our analysis provides additional evidence supporting the view that S-DNA is a stable intermediate encountered during dsDNA melting by mechanical force. In addition, we demonstrated that the estimated diffusion models can detect dynamical signatures of conformational degrees of freedom not directly observed in experiments.

1. Introduction

There has been renewed interest in understanding the thermodynamics and kinetics of DNA melting due to recent advances in both single-molecule experiment techniques [1–4] and theoretical modeling for such systems [5–7]. DNA's mechanical properties influence a variety of its biological functions such as how it wraps around histones, packs into phage heads, and interacts with proteins [1]. It is believed that many biological machines depend on the mechanical properties of dsDNA [4, 5]. These mechanical properties can be exploited by novel therapeutics whose design is guided by information extractable from single-molecule force spectroscopy [8, 9].

Here we report experiments on repeated stretching/relaxation cycles of dsDNA using atomic force microscope (AFM) and quantification of the dynamical response of these experiments using time series modeling techniques. An individual λ-DNA molecule was stretched and relaxed at constant velocity. A force induced transition between B form DNA (B-DNA) and S form DNA (S-DNA) occurs prior to dsDNA melting by mechanical force [2, 5]. Diffusion models were used to extract additional information from the AFM time series. A collection of stochastic dynamical models were fit to each observed time series using maximum likelihood techniques [10–14]. Using these models, we were able to extract approximate dynamical system information, *e.g.* effective internal system force and effective diffusion coefficients, from single-molecule force spectroscopy data.

It has been suggested that the region of phase space where S-DNA is stable contains many large kinetic barriers in the underlying energy landscape [5]. We provided evidence that supports this view and also demonstrated that the magnitude of the effective diffusion coefficient of dsDNA depends on the molecular extension, suggesting that the dynamics in this portion of phase space are indeed quite complex [2, 3, 10–12, 15, 16]. We also demonstrated that pathwise estimation techniques can capture signatures of large-scale conformational fluctuations using batches of AFM time series. This type of information was inferred by comparing the diffusion models obtained using different AFM time series for estimation. The modeling techniques presented are aimed at more fully utilizing the wealth of information contained in single-molecule time series. In the DNA molecule studied, we showed that the measured diffusion coefficient is strongly correlated with DNA reannealing dynamics.

2. Materials and Methods

2.1. AFM Experiments

DNA molecules of λ-BstE II digest (Sigma-Aldrich), with length distribution 117–8454 bp, in Tris-EDTA buffer solution (150mM NaCl, 10mM Tris, 1mM EDTA, pH 8), was allowed to absorb on a gold substrate. A Si_3N_4 AFM tip (Veeco), with a cantilever spring constant of 0.03 N/m and a resonance frequency in buffer of 2.5 kHz, was used to pick up individual DNA molecules. The spring constant was determined using the equipartition theorem [17,18]. After a single DNA molecule was attached between the tip and the substrate, repeated stretching

and relaxation cycles at 0.25 μm/s were performed at a sampling rate of 10 kHz. The force as a function of time was recorded.

2.2. Local Maximum Likelihood Estimation

Stochastic differential equations (SDEs) were fit to time series of the AFM experimental data. Each time series resulted in the estimation of a new SDE. The global model representing the dynamics of one path was assumed to be a nonlinear diffusion of the form [11,12]:

$$d\xi_t = \mu(t,\xi_t)dt + \sqrt{2}\sigma(\xi_t)dW_t \tag{1}$$

$$y_t = \xi_t + \epsilon_t, \tag{2}$$

where ξ_t is the molecular extension at time t, W_t is the standard Brownian motion, $\mu(\cdot,\cdot)$ is the time dependent drift function, and $\sigma^2(\cdot)$ is the diffusion coefficient. The drift term is time dependent because the AFM stage was moving at constant velocity. The measurement noise process ϵ_t does not allow us to directly observe ξ_t, instead we observe $y_t = \xi_t + \epsilon_t$.

In single-molecule systems, the complexity of the atomistic system often cannot be ignored and can significantly complicate developing physically based, accurate stochastic parametric SDE models [10,15,19,20] from *a priori* considerations. Therefore, we assumed that the global dynamics were unknown *a priori*, so appealing to a standard parametric estimation scheme was problematic. However, in the case when a physics based model exists, many of the pathwise inference tools presented here are still applicable.

To overcome the difficulty of unknown global drift and diffusion functions, we used local models [10–12] to fit the coefficients of polynomial SDEs whose functional form was motivated by the overdamped Langevin equation [12]. The relevant equations are:

$$\sigma^{\mathrm{loc}}(\xi) := \quad (C + D(\xi - \xi^o)) \tag{3}$$

$$F^{\mathrm{Ext}}(t,\xi) := k_{\mathrm{pull}}(\lambda(t) - \xi) \tag{4}$$

$$F^{\mathrm{Int}}(\xi) := \quad (A + B(\xi - \xi^o)) \tag{5}$$

$$\mu^{\mathrm{loc}}(t,\xi) := \frac{\left(\sigma^{\mathrm{loc}}(\xi)\right)^2}{kT}\Big(F^{\mathrm{Int}}(\xi) - F^{\mathrm{Ext}}(t,\xi)\Big), \tag{6}$$

where kT is the Boltzmann constant times the system temperature, F^{Ext} is the external force added to the system, $\lambda(\cdot)$ is the AFM stage position we control, k_{pull} is the spring constant associated with the cantilever which applies force to the molecule, F^{Int} is the force due to internal molecular forces, and $\theta \equiv (A,B,C,D)$ is the local parameter vector estimated by approximate maximum likelihood estimation of hidden Markov models [14]. ξ^o is a free parameter; since a constant velocity protocol is used for $\lambda(\cdot)$, we set ξ^o to the average value in local time series windows. The windows were formed by dividing a single global time series into $M = 80$ windows, each representing an equal temporal fraction of the total time series. A can be interpreted as the internal system force associated with the value ξ^o and B the associated linear sensitivity (similarly for C and D). The time series used for estimation contained uniformly spaced observations, consecutive observations were separated by 0.20 ms. With $M = 80$, each local window consisted of ~ 400 temporal observations.

The modeling ideas behind the two-scale realized volatility estimator (TSRV) [13] were used to approximate the variance of the noise process in each local window ($\mathbb{E}[\epsilon_t^2] \approx \frac{1}{2n}\sum_{i=1}^{n}(y_i - y_{i-1})^2$). A white measurement noise was assumed [21]. This information was used in conjunction with likelihood based methods to estimate quantities describing the ξ

dynamics. The fitting criterion we used was motivated by the local linear maximum likelihood type method outlined in [14]. The stage location $\lambda(\cdot)$ was denoised using the Daubechies (5 vanishing moments) wavelet family and all measurement noise was assumed to be contained in the measured molecular extension, ξ. The validity of the various assumptions, like diffusive noise, local linearity, white measurement noise, *etc.*, were tested in an *a posteriori* fashion using the probability integral transform based Q-test developed in Ref. [22].

To obtain a global model, a second degree penalized spline was used to stitch the piecewise polynomial models together [23]. The generalized cross-validation procedure was used to pick the smoothing parameter and knot selection was automated using an algorithm found in Appendix B of [23]. Note that for each observed time series separate μ and σ functions are constructed from information contained in the corresponding estimated μ^{loc} and σ^{loc} functions. The spline fitting procedure and additional methodological details are outlined in Ref. [21].

3. Results and Discussion

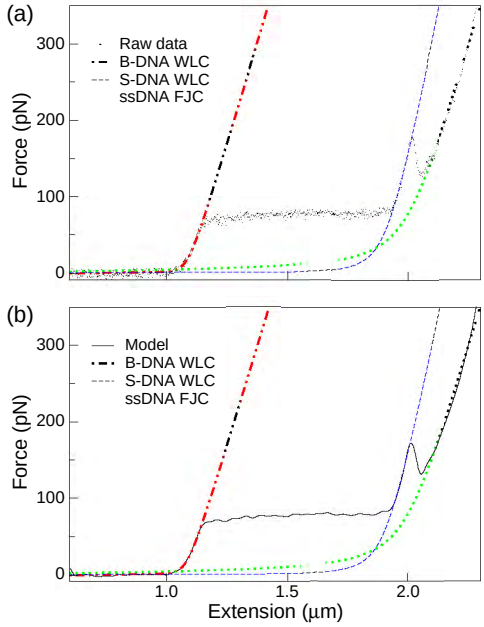

Figure 1. (a) Experimental force-extension data for the stretching of double-stranded λ-DNA. (b) Force-extension for the same λ-DNA molecule generated using overdamped diffusion models (see text). Fitting using the extensible WLC and FJC models suggests a B to S transition followed by DNA melting.

Figure 1 shows the force-extension curve of segment of digested, double-stranded λ-DNA. There is a plateau at low extension and a force peak at 150 pN, followed by ssDNA stretching. The plateau indicates a B to S transition, and the force peak is a product of dsDNA melting. The general feature is consistent with previous observations [2, 3], however, the detailed shapes of the transition are different. The melting transition was usually observed as a sloped plateau, but in our data a clear force peak was always present. This may be caused by a number of factors such as the pulling velocity,

cantilever stiffness, and sample preparation. In Refs. [2, 3], the slope of the second plateau depends on the pulling velocity, with slower pulling velocity (< 0.5 μm/s) showing better defined plateaus. The cantilever spring stiffness may also play a role. The spring used in our experiment (0.03 N/m) is stiffer than those used in previous experiments (0.01 N/m) [2,3], and the stiffer cantilever is supposed to track the rugged energy landscape better [24]. Note that a few curves in Ref. [2] do show a force peak like feature. Other parameters such as sample incubation time may also affect the shape of the curve. It will be interesting to know exactly how the experimental parameters influence the results of these nonequilibrium measurements.

As shown in Fig. 1a, the mechanical stretching of dsDNA can be described by an extensible wormlike chain (WLC) model [25, 26],

$$x = b_{ds}\left[1 - \frac{1}{\sqrt{4\beta P_{ds} F}} + \frac{F}{K_{ds}}\right], \tag{7}$$

where P_{ds}, b_{ds}, and K_{ds} are the persistence length, contour length, and elastic stretch modulus for dsDNA. For the portion of the curve before the plateau, we obtained $P_{ds} =$ 50 nm, $b_{ds} = 1100$ nm, and $K_{ds} = 1200$ pN, consistent with previous estimates of B-DNA [5, 8]. For the portion of the curve immediately after the plateau, we obtained

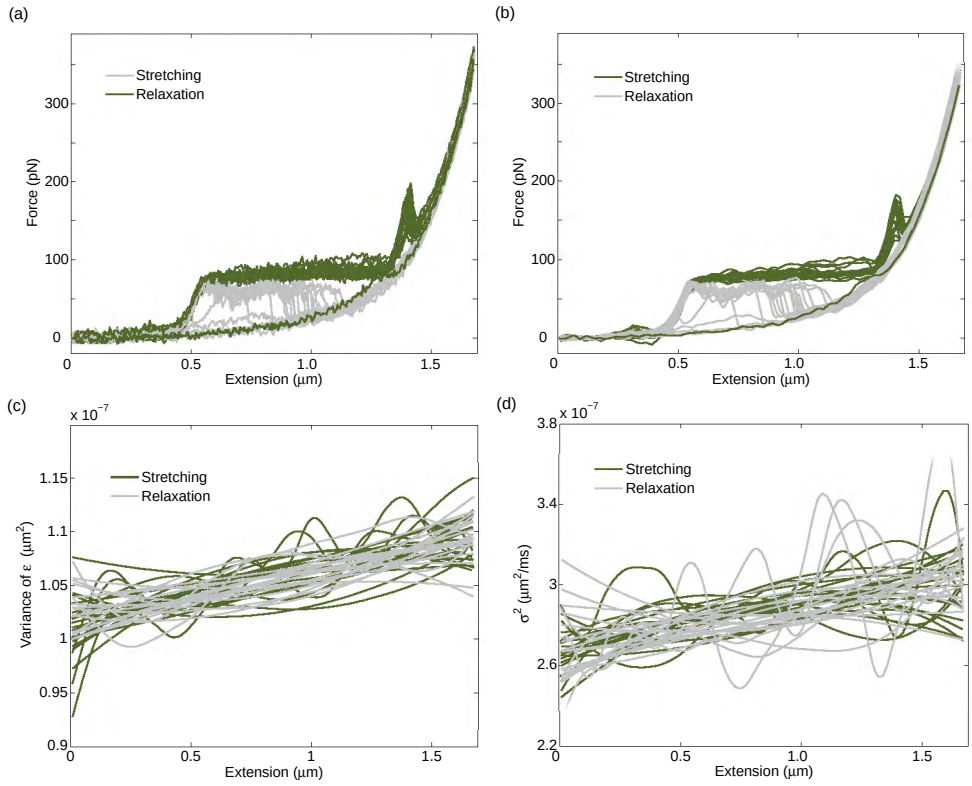

Figure 2. Stretching and relaxation of a segment of λ-DNA. (a) AFM experimental data. (b) Estimated internal force from SPA functions fit using 20 extension and relaxation cycles shown in (a). (c) The estimated measurement noise variance. (d) The estimated diffusion coefficient. Each function was measured on a pathwise basis.

$P_{ds} = 10$ nm, $b_{ds} = 2000$ nm, and $K_{ds} = 3200$ pN. These values agree with theoretical models for S-DNA, where the persistence length ranges between 7–12 nm and the stretch modulus is higher than that of B-DNA [27]. Following the S-DNA melting, which displays a force peak like feature, ssDNA stretching was approximated using the extensible freely jointed chain model [8],

$$x = b_{ss}\left[\coth(2\beta P_{ss}F) - \frac{1}{2\beta P_{ss}F}\right]\left[1 + \frac{F}{K_{ss}}\right], \tag{8}$$

where P_{ss}, b_{ss}, and K_{ss} are the persistence length, contour length, and stretch modulus for ssDNA. Fitting to our raw data results in $P_{ss} = 0.75$ nm, $b_{ss} = 2000$ nm, and $K_{ss} = 2200$ pN. The force-extension generated by our surrogate diffusion models (see Fig. 1b) returns the same fitted values for the parameters as our experimental data for B-DNA, S-DNA, and ssDNA.

Figure 2 displays the experimental data, the estimated internal force, the estimated measurement noise, and the diffusion coefficient as a function of the molecular extension for both extension and relaxation portions of the cycle. Significant hysteresis exists between the extension and relaxation curves, indicating only one strand was attached to the substrate and the tip [3], thus allowing the other strand to melt and reanneal while still partially connected to the stretched strand. A total of 20 cycles were analyzed. Note that in both extension and relaxation, the effective diffusion coefficient tends to increase with molecular extension. The same trend was observed in the estimated variance of the measurement noise.

Figure 3 shows the region associated with the melting of dsDNA. It has been suggested that S-DNA is an intermediate encountered when DNA is mechanically melted starting from B-DNA [1, 2, 5, 7, 28] and that S-DNA has higher stretch modulus than B-DNA [1, 3, 5]. Our data suggest the existence of a state with distinct mechanical properties, though this state may only be metastable. Figure 3b displays the measured AFM time series along with the estimated internal force made by appealing to the diffusion models.

Recall that we were conducting an experiment in which the AFM stage was set to move at a constant velocity and we were using overdamped diffusion models to describe the dynamics of molecular extension and approximate the effective internal force [29] experienced by the molecule. Therefore it is not surprising that, on average, a slight offset in the measured external and estimated internal force exists. This feature may facilitate the understanding of complex transitions experienced in some other single-molecule experiments [30–33]. This type of analysis could potentially help in more accurately determining the rupture times and forces in situations where single-molecule experiments encounter noise problems [34]. For example, in this system the estimated internal force aids in making it apparent that in each experiment the local force maxima related to S-DNA occurs at roughly the same extension. However, there is still a discernible distribution of forces at this local maximum and the diffusion models facilitate in extracting this type of information because the estimated diffusion models help increase the signal to noise ratio. The signal to noise ratio is believed to decrease as

the the pulling velocity decreases in AFM and optical tweezers experiments [35]. Slow single-molecule pulling experiments can provide quantitative information describing energetic and kinetic barriers [5, 16, 31] associated with biomolecules, so statistical analysis tools that can help in deciphering noisy information from such experiments may shed light on the governing physics associated with various physical phenomena such as DNA melting and annealing.

Figure 4 shows the relaxation portion of the cycle. The region associated with the molten DNA (fit with ssDNA model) of the force-extension curves appear similar to those of the stretch cycle. Hysteresis in the force is only observed in the region associated with the plateau in the stretch portion of the cycle [1–3]. The nature of the hysteresis has been shown to depend on the pulling velocity. However, here we analyzed only one pulling velocity and still observed a significant variation in dynamical responses. In figure 4, we focus on curves that either failed to return to the plateau, characteristic of B to S transition, or returned to this plateau much later than the typical curve. It appears that the typical return to the plateau occurs around 1 μm for this particular λ-DNA molecule. The common feature in these solid highlighted curves is that their corresponding effective diffusion coefficient functions show pronounced differences

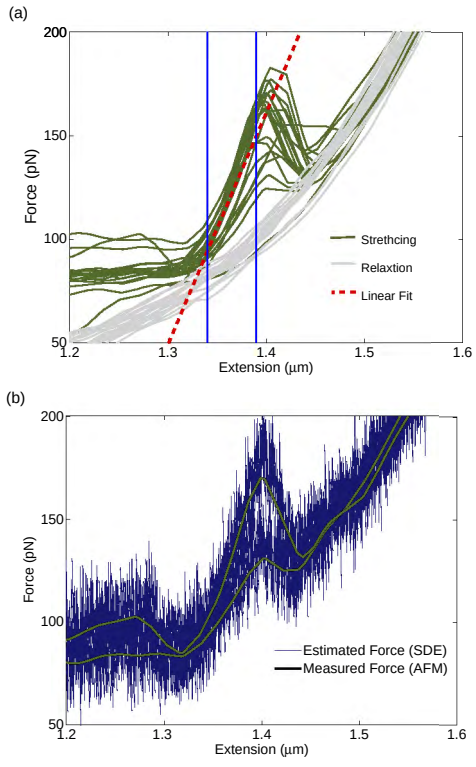

Figure 3. (a) The force-extension relationship around the force peak calculated using the internal force estimated by the overdamped diffusion models. The dashed line represents a linear least squares fit using the data between the vertical lines. (b) The estimated internal force is plotted with experimental data. The sign of external AFM force is changed from Eq. 6 to facilitate comparing the forces.

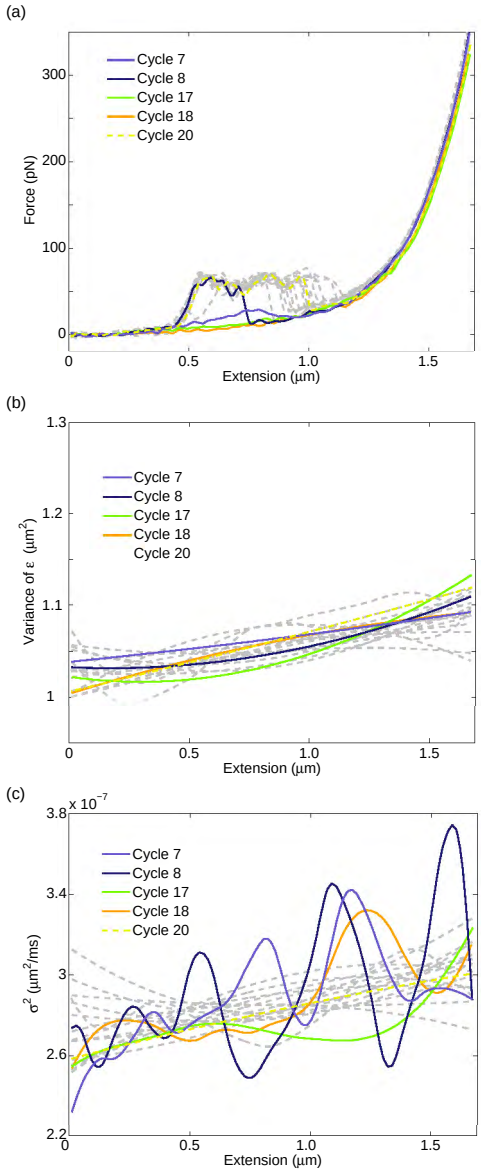

Figure 4. Estimated SPA functions obtained using 20 relaxation cycles from AFM time series. Those experiments where the plateau was either not observed or occurred late are shown as solid lines. (a) Estimated internal force, (b) variance of measurement noise, and (c) diffusion coefficient. The diffusion coefficient of the solid curves displays unusually oscillatory behavior compared to the rest of the curve population. The oscillation starts to become apparent in the region of phase space where S-DNA is stable [5]. Note that the force-extension curves show no apparent hysteresis until much later in the cycle.

from the rest of the curves well before they reach the extension of 1 μm. We noticed that a force plateau was not observed in the extension portion of cycle 18.In Fig. 2a only one curve (cycle 18) does not exhibit the typical force plateau during extension. Apparently a high degree of conformational "frustration" was experienced in this part of the experiment. This frustration was enough to prevent the molecule from properly

reannealing at zero force. Note that with the slow pulling speeds used and length of the DNA molecule retained, the molecule had several seconds to recover in the low force regime. The frustration apparently was not severe enough to completely prevent reannealing in future experiments, *e.g.* a strand did not melt off completely, because on cycle 20 we again observed the force plateau (this dashed curve is denoted in the legend as well).

The information in the surrogate models allowed us to identify these dynamical responses as atypical, as indicated by the difference in diffusion coefficient at extensions around 1.3 μm. This region coincide with the region where S-DNA is stable. It is believed that there are many local minima possessing large kinetic barriers on the energy landscape describing dsDNA [5]. The addition of external force further would complicate such a surface. The rapid change of diffusion coefficient magnitude is consistent with the idea that system hops between different regions as the landscape evolves, but in addition each local minima appears to be associated with effective diffusion coefficients possessing measurably different magnitudes [16]. Exploring this portion of phase space is by no means rare given that 4 of 20 curves appear to deviate from the typical annealing path. By using the overdamped models, we indirectly quantified information about conformational information not directly measured in our experiments [5]. One plausible explanation for the observed oscillatory diffusion coefficient would be that overdamped diffusion model is not valid in this regime. Pathwise statistical hypothesis tests [22] suitable for nonstationary signals were applied to test the goodness-of-fit of the overdamped diffusion proxy. Our model presented in Eq. 4 performed well under this test even for the curves with an oscillatory diffusion coefficient indicating that the phenomenon was not an artifact of a poor dynamical model.

Previous works [30] have demonstrated how a variance analysis can provide a researcher with physical system information that a single force extension curve alone can not easily provide. That is, when one analyzes the extension and relaxation curves, hysteresis does not seem apparent. Parameter estimates of a simple polymer model like the extended FJC may even yield nearly indistinguishable parameters for extension and relaxation on repeated cycles. One may assume that this indicates that the process is quasi-static and effectively at equilibrium. However, hidden conformational degrees of freedom, which may differ substantially from path to path, can modulate the observed dynamics and provide evidence of nonequilibrium dynamics [30]. Such information is also contained in the diffusion coefficient of the models we estimated from the AFM time series. Note that the measurement noise dominates the raw variance signal in the system analyzed here. Thermal noise was only a small fraction (10-30 %) of the signal that would be determined by typical variance analysis methods [30, 36]. The highlighted curves in the estimated measurement noise variance (Fig. 4c) show no substantial deviations from the other curves. Simultaneously observing typical measurement noise and atypical effective diffusion coefficients in frustrated experimental cycles gave further evidence that the observed oscillations in the diffusion coefficient correspond to a physically relevant phenomenon. Extracting such a signal requires us to wrap a loose structure around the

data set.

4. Conclusions

Stable *S*-DNA was observed during mechanical stretching of dsDNA. Its persistence length and elastic stretch modulus were obtained by fitting the data to the extensible WLC model. The AFM time series were analyzed with diffusion models. Using these models, we were able to detect dynamical signatures that standard single-molecule analysis techniques for analyzing force extension curves could not discern. An unusually oscillatory diffusion coefficient in the estimated models was demonstrated to correlate with failed reannealing of melted dsDNA. Stochastic dynamical modeling coupled with functional data analysis ideas [12,37,38] shows promise in helping to understand complex single-molecule dynamics. The use of pathwise inference procedures may also be combined with first principles physical models to help in understanding such systems [16, 30, 39]. Single-molecule force spectroscopy will likely continue to help in providing a better understanding of complex biological processes like dsDNA melting [16, 40–42].

5. Acknowledgments

CPC thanks NSF DMS 0240058, NSF ACI-0325081, NIH T90DK070121 and the computer support provided by NSF CNS-0421109 and a partnership between Rice AMD and Cray. CHK thanks NSF DMR-0505814 and Welch Foundation C-1632, NCH thanks NIH T90DK70121, and KJL thanks NSC 962627M005001 for support.

[1] S. B. Smith, Y. Cui, and C. Bustamante. Overstretching B-DNA: The elastic response of individual double-stranded and single-stranded DNA molecules. *Science*, 271:795 – 799, 1996.

[2] M. Rief, H. Clausen-Schaumann, and H. E. Gaub. Sequence-dependent mechanics of single DNA molecules. *Nature Structural Biology*, 6:346–349, 1999.

[3] H. Clausen-Schaumann, M. Rief, C. Tolksdorf, and H. E. Gaub. Mechanical stability of single DNA molecules. *Biophys. J.*, 78:1997–2007, 2000.

[4] C.H. Albrecht, G. Neuert, R.A. Lugmaier, and H. E. Gaub. Molecular force balance measurements reveal that double-stranded DNA unbinds under force in rate-dependent pathways. *Biophys. J.*, 94:4766–4774, 2008.

[5] S. Cocco, J. Yan, J.-F. Leger, D. Chatenay, and J.F. Marko. Overstretching and force-driven separation of double-helix DNA. *Phys. Rev. E*, 70:011910, 2004.

[6] T. Ambjornsson, S. K. Banik, O. Krichevsky, and R. Metzler. Breathing dynamics in heteropolymer DNA. *Biophys. J.*, 92:2674–2684, 2007.

[7] S. Whitelam, S. Pronk, and P. L. Geissler. There and (slowly) back again: Entropy-driven hysteresis in a model of DNA overstretching. *Biophys. J.*, 94:2452–2469, 2008.

[8] M. J. McCauley and M.C. Williams. Mechanisms of DNA binding determined in optical tweezer experiments. *Biopolymers*, 85:154 – 168, 2004.

[9] C. Ke, M. Humeniuk, H. S-Gracz, and P. E. Marszalek. Direct measurements of base stacking interactions in DNA by single-molecule atomic-force spectroscopy. *Phys. Rev. Lett.*, 99:018302, 2007.

[10] C.P. Calderon. Fitting effective diffusion models to data associated with a "glassy potential":

Estimation, classical inference procedures and some heuristics. *Mutliscale Modeling and Simulation*, 6:656–687, 2007.

[11] C.P. Calderon. On the use of local diffusion for path ensemble averaging in potential of mean force computations. *J. Chem. Phys.*, 126:084106, 2007.

[12] C.P. Calderon and R. Chelli. Approximating nonequilibrium processes using a collection of surrogate diffusion models. *J. Chem. Phys.*, 128:145103, 2008.

[13] L. Zhang, P. A. Mykland, and Y. Ait-Sahalia. A tale of two time scales: Determining integrated volatility with noisy high-frequency data. *Journal of the American Statistical Association*, 100:1394–1411, 2005.

[14] J.C. Jimenez and T. Ozaki. An approximate innovation method for the estimation of diffusion processes from discrete data. *J. Time Series Analysis*, 27:77–97, 2005.

[15] G. Hummer. Position-dependent diffusion coefficients and free energies from Bayesian analysis of equilibrium and replica molecular dynamics simulations. *New J. Phys.*, 7:1–14, 2005.

[16] J. Chahine, R.J. Oliveira, V.B.P. Leite, and J. Wang. Configurational-dependent diffusion can shift the kinetic transition state and barrier height of protein folding. *Proc. Natl. Acad. Sci. USA*, 104:14646, 2007.

[17] J. L. Hutter and J. Bechhoefer. Calibration of atomic-force microscope tips. *Rev. Sci. Instrum.*, 64:1868–1873, 1993.

[18] H.-J. Butt and M. Jaschke. Calculation of thermal noise in atomic force microscopy. *Nanotechnology*, 6:1–7, 1995.

[19] S. Park and K. Schulten. Calculating potentials of mean force from steered molecular dynamics simulations. *J. Chem. Phys.*, 120:5946–5961, 2004.

[20] G. Hummer and I.G. Kevrekidis. Coarse molecular dynamics of a peptide fragment: Free energy, kinetics, and long-time dynamics computations. *J. Chem. Phys.*, 118:10762–10773, 2003.

[21] C.P. Calderon, N. Harris, C.-H. Kiang, and D.D. Cox. Quantifying multiscale noise sources in single-molecule time series via pathwise statistical inference procedures. *submitted*, 2008.

[22] Y. Hong and H. Li. Nonparametric specification testing for continuous-time models with applications to term structure of interest rates. *The Review of Financial Studies*, 18:37–84, 2005.

[23] D. Ruppert, M.P. Wand, and R.J. Carroll. *Semiparametric Regression*. Cambridge University Press, New York, 2003.

[24] U. Gerland, R. Bundschuh, and T. Hwa. Mechanically probing the folding pathway of single RNA molecules. *Biophys. J.*, 84:2831–2840, 2003.

[25] T. Odijk. Stiff chains and filaments under tension. *Macromolecules*, 28:7016–7018, 1995.

[26] C. G. Baumann, S. B. Smith, V. A. Bloomfield, and C. Bustamante. Ionic effects on the elasticity of single DNA molecules. *Proc. Natl. Acad. Sci. USA*, 94:6185–6190, 1997.

[27] C. Storm and P. C. Nelson. Theory of high-force DNA stretching and overstretching. *Phys. Rev. E*, 67:051906, 2003.

[28] C. Danilowicz, V.W. Colijee, C. Bouzigues, D.K. Lubensky, D.R. Nelson, and M. Prentiss. DNA unzippped under a constant force exhibits multiple metastable intermediates. *Proc. Natl. Acad. Sci. USA*, 100:1694–1699, 2003.

[29] R. Kupferman and A.M. Stuart. Fitting SDE models to nonlinear Kac-Zwanzig heat bath. *Physica D*, 199:279–316, 2004.

[30] K.A Walther, J. Brujic, H. Li, and J. M. Fernandez. Sub-Angstrom conformational changes of a single molecule captured by AFM variance analysis. *Biophys. J.*, 90:3806–3812, 2006.

[31] N.C. Harris, Y. Song, and C.-H. Kiang. Experimental free energy surface reconstruction from single-molecule force spectroscopy using Jarzynski's equality. *Phys. Rev. Lett.*, 99:068101, 2007.

[32] M. Kawakami, K. Byrne, D. J. Brockwell, S.E. Radford, and D.A. Smith. Viscoelastic study of the mechanical unfolding of a protein by afm. *Biophys. J.*, 91:L16–L18, 2006.

[33] M. J. Higgins, J.E. Sader, and S.P. Jarvis. Frequency modulation atomic force microscopy reveals individual intermediates associated with each unfolded i27 titin domain. *Biophys. J.*, 90:640–

647, 2006.

[34] J.R. Moffitt, Y.R. Chemla, S.B. Smith, and C. Bustamante. Recent advances in optical tweezers. *Annual Review of Biochemistry*, 77:19.1–19.4, 2008.

[35] P. Maragakis, F. Ritort, C. Bustamante, M. Karplus, and G.E. Crooks. Bayesian estimates of free energies from nonequilibrium work data in the presence of instrument noise. *J. Chem. Phys.*, 129:024102, Jul 2008.

[36] M. J. Tysak, D. E. Dupuis, W. H. Guilford, J. B. Patlak, G. S. Waller, K. M. Trybus, D. M. Warshaw, and S. Lowey. Two heads of myosin are better than one for generating force and motion. *Proc. Natl. Acad. Sci. USA*, 96:44024407, 1999.

[37] J. Ramsay and B.W. Silverman. *Functional Data Analysis*. Springer-Verlag, New York, 2005.

[38] F. Ferraty and P. Vieu. *Nonparametric Functional Data Analysis*. Springer, New York, 2006.

[39] D.E. Makarov and P.K. Hansma and H. Metiu. Kinetic Monte Carlo simulation of titin unfolding. *J. Chem. Phys.*, 114:9663, 2001.

[40] S. Dixit, M. Singh-Zocchi, J. Hanne, and G. Zocchi. Mechanics of binding of a single integration-host-factor protein to DNA. *Phy. Rev. Lett.*, 94:118101, 2005.

[41] M. Vendruscolo and C. M. Dobson. Dynamic visions of enzymatic reactions. *Science*, 313:1586, 2006.

[42] S. Liu, G. Bokinsky, N. G. Walter, and X. Zhuang. Dissecting the multi-step reaction pathway of an RNA enzyme by single-molecule kinetic fingerprinting. *Proc. Natl. Acad. Sci. USA*, 104:12634–12639, 2007.