

## ABSTRACT

Title of dissertation:

METHODS OF SINGLE-MOLECULE  
ENERGY LANDSCAPE RECONSTRUCTION  
WITH OPTICAL TRAPS

Michel de Messières, Ph.D., 2012

Directed by:

Assistant Professor Arthur La Porta  
Department of Physics

Optical traps facilitate measurement of force and position as single molecules of DNA, RNA, or protein are unfolded and refolded. The effective energy landscape of a biomolecule can be reconstructed from the force and position data, providing insight into its structure and regulatory functions. We have developed new experimental and analytical methods to reconstruct energy landscapes by taking advantage of the harmonic constraint of an optical trap. We demonstrate the effectiveness of these methods using a model DNA hairpin and then apply these methods to study problems of practical biophysical interest. CCR5 mRNA has been demonstrated to stimulate -1 programmed ribosomal frameshifting and we measure its structural properties. We measure the binding energy of a GA/AG tandem mismatch, one of many mismatches with unusual properties. We use our single-molecule methods to reproduce bulk measurements of the nearest-neighbor DNA base-pair free energy parameters and we consider possible refinements to the model. We also study an alternative method of measuring energy landscapes, Dynamic Force Spectroscopy (DFS), and conduct experiments on DNA quadruplexes to demonstrate the effectiveness of DFS with optical traps. Finally, we develop theory to elucidate the role of noise in optical trap measurements of energy landscapes.

METHODS OF SINGLE-MOLECULE ENERGY LANDSCAPE  
RECONSTRUCTION WITH OPTICAL TRAPS

By

Michel de Messières

Dissertation submitted to the Faculty of the Graduate School of the  
University of Maryland, College Park, in partial fulfillment  
of the requirements for the degree of  
Doctor of Philosophy  
2012

Advisory Committee:

Assistant Professor Arthur La Porta, Chair

Assistant Professor Michelle Girvan

Professor Christopher Jarzynski

Associate Professor Jason Kahn

Assistant Professor Arpita Upadhyaya

© Copyright by  
Michel de Messières  
2012

## Acknowledgements

First I would like to thank my advisor Arthur La Porta, who was the inspiration and guidance for all of the work presented here and showed me how to properly think about scientific research.

The friends I made here supported me and contributed to making these years productive and pleasant. I thank all the professors and mentors who educated me and provided essential support and advice. I would especially like to thank my colleagues Barbara and Jen-Chien, who were my friends and collaborators.

I thank my father who valued education so highly and my mother who provides endless support in all things. My siblings inspire me through their own successes and encouragement. Finally, I thank my wife Candace who was at my side every step of the way.

# Table of Contents

Acknowledgements.....	ii
Table of Contents.....	iii
List of Tables.....	v
List of Figures.....	vi
Chapter 1: Introduction to Energy Landscapes and Optical Traps.....	1
Introduction.....	1
Energy Landscapes and Applied Force.....	1
Principles of Optical Trapping.....	3
Mechanical Design.....	5
Principles of Calibration.....	7
Chapter 2: Energy Landscape of a Harmonically Constrained Biopolymer.....	11
Introduction.....	11
Background.....	11
Harmonic Constraint Technique.....	15
Implications.....	22
Materials and Methods.....	24
Sample Preparation.....	24
Instrumentation.....	25
Calibration.....	26
Experimental Setup.....	27
Theoretical Model.....	27
Results & Discussion.....	28
Harmonic Constraint Data Analysis.....	28
Force Clamp Data Analysis.....	32
Combined Samples.....	33
Discussion.....	35
Conclusion.....	36
Chapter 3: CCR5 mRNA Unfolding Pathways.....	38
Introduction.....	38
Background.....	38
Pseudoknot Structures and Slippery Sites.....	40
Theoretical mRNA Structures.....	41
Disruption Pathways.....	43
Substeps in the Disruption Pathways.....	49
Blocking First and Last 20 Bases of mRNA with DNA Oligos.....	52
Identification of the Pseudoknot Conformation.....	54
Folding Times and Intermediate States.....	55
Role of miRNA-1224.....	56
Methods.....	58
Details of Worm-Like Chain Theory.....	61
Conclusion.....	62
Chapter 4: Theoretical Modeling of Energy Landscapes.....	63
Introduction.....	63

Methods.....	63
Hydrogen Bond Correction.....	67
Loop Initiation .....	69
Conclusion .....	70
Chapter 5: Tandem GA Mismatch Energy .....	71
Introduction.....	71
Background.....	71
Methods.....	72
Analysis.....	74
Conclusion .....	74
Chapter 6: Nearest-Neighbor Free Energy Parameters for Duplex DNA.....	76
Introduction.....	76
Background.....	76
Methods.....	77
Reversibility.....	78
Fitting Method and Results.....	81
Conclusion .....	83
Chapter 7: G-Quadruplex Dynamic Force Spectroscopy .....	84
Introduction.....	84
Background.....	84
Theoretical Models .....	85
Methods.....	87
Analysis.....	91
Interpretation.....	95
Conclusion .....	98
Chapter 8: Noise Associated With Nonconservative Forces in Optical Traps .....	99
Introduction.....	99
Background.....	99
Simplified Model .....	102
Calculation of Excess Fluctuations.....	104
Application to a Realistic Force Field .....	109
Additional Fluctuations When Stretching a DNA Tether.....	116
Conclusion .....	121
Bibliography .....	123

## List of Tables

Table 7.1	Energy landscape parameters from maximum-likelihood fits	93
-----------	--	----

# List of Figures

## **Chapter 1**

Figure 1.1	Energy landscapes under applied force	2
Figure 1.2	Experimental setup	4
Figure 1.3	Conservation of momentum in optical traps	5
Figure 1.4	Schematic of optical trap	6
Figure 1.5	Photo of optical trap	7
Figure 1.6	Calibration scans relating voltage to position	8

## **Chapter 2**

Figure 2.1	Simulated force clamp and harmonic constraint	16
Figure 2.2	Simulated energy landscape reconstruction	19
Figure 2.3	Structure of the long DNA hairpin	30
Figure 2.4	Experimental harmonic constraint	31
Figure 2.5	Experimental force clamp	33
Figure 2.6	Combined samples for harmonic constraint	34
Figure 2.7	Reconstructed energy landscape with new theoretical model	37

## **Chapter 3**

Figure 3.1	Pseudoknots and slippery sites	41
Figure 3.2	Theoretically predicted pseudoknot structures	42
Figure 3.3	Disruption pathways of CCR5 mRNA	45
Figure 3.4	Classification of disruption pathway types	47
Figure 3.5	Forward and reverse scans of CCR5 mRNA	48
Figure 3.6	Dependency of structure on $Mg^{2+}$	49
Figure 3.7	Substeps of disruption pathways	51
Figure 3.8	Blue conformation substep	52



Figure 3.9	mRNA with first and last 20 bases blocked by DNA oligos	53
Figure 3.10	Comparison of substeps to theoretical predictions	54
Figure 3.11	Time dynamics of intermediate states	56
Figure 3.12	Influence of miRNA on mRNA structure	57

#### **Chapter 4**

Figure 4.1	Structure of the example long DNA hairpin	64
Figure 4.2	Mfold energies for a long DNA hairpin	64
Figure 4.3	Ensemble of energy states per extension	65
Figure 4.4	Energy as a function of end-to-end distance	66
Figure 4.5	Theory and data with hydrogen bond correction	69

#### **Chapter 5**

Figure 5.1	GA/AG mismatch energy and controls	73
------------	------------------------------------	----

#### **Chapter 6**

Figure 6.1	Structure of sequence A and sequence B	78
Figure 6.2	Example folding and unfolding data	78
Figure 6.3	Reconstructed energy landscapes	80
Figure 6.4	Fits for 10 nearest-neighbor free energy parameters	82

## **Chapter 7**

Figure 7.1	Dynamic force spectroscopy energy landscapes	87
Figure 7.2	Experimental setup	89
Figure 7.3	Example force-extension curves	90
Figure 7.4	Opening distance for quadruplex structure	91
Figure 7.5	Binned representation of maximum-likelihood fits	94
Figure 7.6	Possible model of quadruplex disruption	96

## **Chapter 8**

Figure 8.1	Scattering profiles and circulation for small and large particles	101
Figure 8.2	Large particle force and position spectral density	108
Figure 8.3	Small particle force and position spectral density	109
Figure 8.4	Force spectral density in a 3D Gaussian optical mode	113
Figure 8.5	Force spectral density for a particle tethered by a DNA molecule	118

# Chapter 1: Introduction to Energy Landscapes and Optical Traps

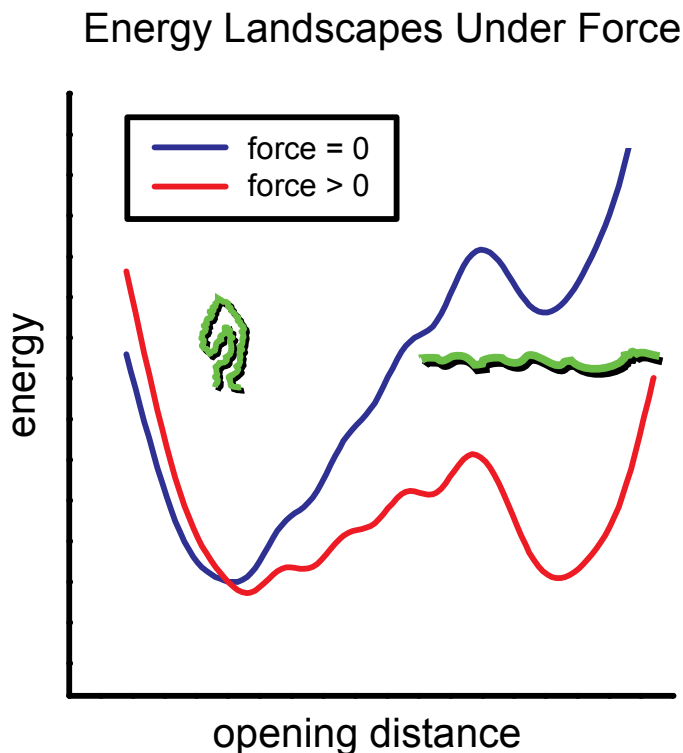
## *Introduction*

This thesis details single-molecule experiments conducted using optical traps to reconstruct the effective free energy landscapes of biomolecules. The measurement of energy landscapes at the molecular scale has broad implications, providing insight into kinetics and regulatory function. In this chapter, the basic principles of energy landscapes and optical traps are reviewed. Chapter 2 details a measurement and analysis technique which is the foundation of this thesis. The following chapters summarize research projects which used an optical trap to measure and interpret energy landscapes of various biomolecular systems. These research projects build upon each other and are presented in an order which highlights this progressive development and cohesiveness.

## *Energy Landscapes and Applied Force*

The effective free energy landscape of a biomolecule can be represented as a function of energy ( $y$ -axis) relative to a reaction coordinate or end-to-end opening distance ( $x$ -axis). Figure 1.1 shows an example energy landscape (blue) where the opening distance ( $x$ ) as a function of time can be modeled as a particle which thermally diffuses along this energy landscape, occupying each position with probability given by the Boltzmann distribution<sup>1</sup>. In this example, the molecule (Fig. 1.1 green) usually remains in the lower-energy folded conformation. However, the application of external force tilts the energy landscape by an amount  $-F x$  (Fig. 1.1 red), where  $F$  is the applied force and  $x$  is the distance along the reaction coordinate.

This tilted energy landscape lowers the energy barrier and increases the probability that the molecule visits the open conformation.



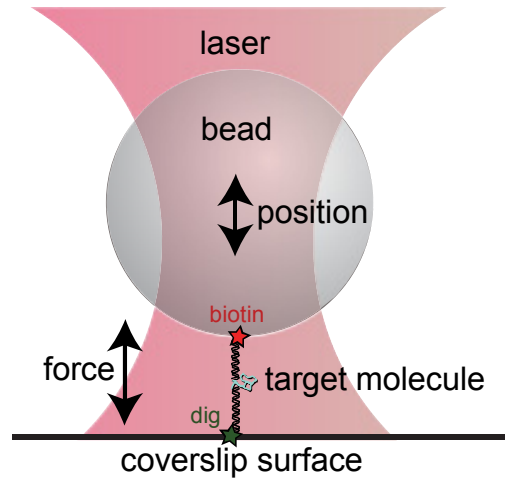
**Figure 1.1** Effect of applied force on an energy landscape. A molecule (green) prefers the folded conformation (small opening distance). Applied force lowers the energy barrier from blue to red, which increases the probability that a transition to the unfolded conformation will occur.

The assumption of a one-dimensional reaction coordinate is not true in general but will be demonstrated to be a highly effective approximation for the systems we study here. Some bonds and biomolecules can be approximated as two-state systems, adopting either the closed or open conformation. For much of this research, we will be considering systems which are more complicated than simple two-state systems. Therefore, we will be interested in accurately measuring not only the initial and final conformation, but also all intermediate conformations the biomolecule may visit. To

access all of the available conformations, we can apply force to a biomolecule using a device called an optical trap.

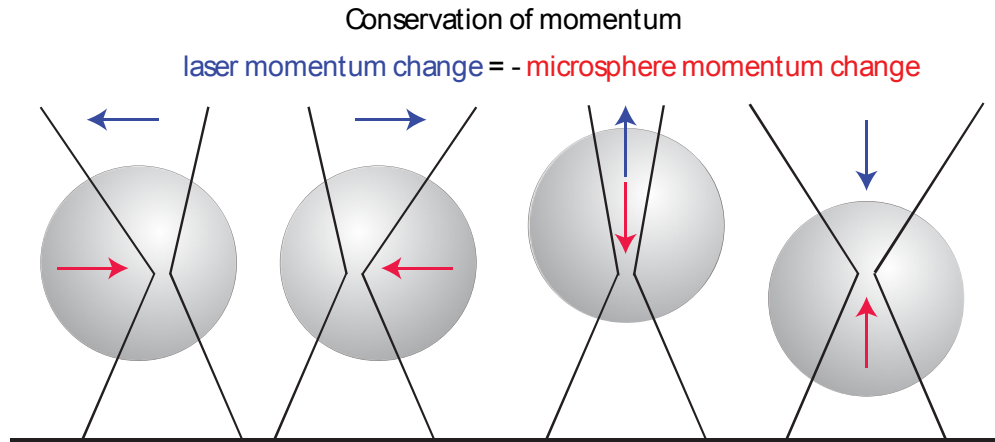
### ***Principles of Optical Trapping***

The optical trap, first developed by Ashkin<sup>2-4</sup>, allows for the precise measurement of force and opening distance on a target molecule. Typical units for these measurements are piconewtons (pN) and nanometers (nm). An optical trap is created with a high numerical aperture lens and a laser. If the laser is sufficiently focused, dielectric microspheres (and many other particles with dielectric properties such as bacteria and viruses<sup>5</sup>) can be stably trapped at the trap center which is approximately located at the focal point of the laser. Figure 1.2 shows a typical experimental setup for this research. A target molecule of interest is attached to the surface and the microsphere by double-stranded DNA handles with biotin or digoxigenin labels. Other setups are possible<sup>6,7</sup> including double-beam traps which have advantages of stability<sup>8</sup>, though here we focus solely on single-beam optical trap experiments. Force on the molecule may be controlled either by moving the stage on which the glass coverslip is attached, or by changing the power of the laser.



**Figure 1.2 Typical experimental setup for this research where relatively short double-stranded DNA handles (black) are used to attach the target molecule (blue) between a microsphere (bead) and the glass coverslip surface. The optical trap allows us to track the applied force and the position of the microsphere.**

In principle, an optical trap is a simple device requiring only an objective lens and a laser. However, accurately measuring the state of the microsphere (position and force) is more complex. The focused laser beam applies a force to the microsphere, drawing it towards the trap center. Momentum conservation demands that the laser beam be deflected in the opposite direction of the applied momentum and this deflection can be measured on a photodiode detector which collects the laser light emerging from the sample chamber. The vertical direction along the laser-beam axis can also be used since conservation of momentum causes the beam to be focused or broadened if the microsphere is respectively pulled toward or away from the laser source. Figure 1.3 illustrates lateral and axial conservation of momentum, where the momentum imparted to the microsphere (red) is always equal to the change in momentum of the laser beam (blue).



**Figure 1.3 Conservation of momentum demands that momentum imparted to the microsphere (red) is matched by equal and opposite momentum change of the laser beam (blue). Lateral and axial momentum changes can both be monitored using a photodiode which collects light emerging from the sample chamber.**

### ***Mechanical Design***

Construction of the optical trap (schematic in Fig. 1.4 and photo in Fig. 1.5) was a major component of this research. The trap is constructed on a gas-suspended table for stability. A polarized 1064 nm infrared Nd:YAG laser is sent through an acousto-optic modulator (AOM) which splits the beam and allows us to control the power of the primary beam through rapid feedback. The beam is then coupled into a single-mode polarization preserving optical fiber. The fiber purifies the beam polarization and minimizes the effects from laser-pointing drift. The beam is collimated and 10% is diverged to a photodiode detector which allows us to monitor the absolute power of the laser for feedback control. The primary beam is aligned on the back of a numerical aperture (NA) 1.49 oil immersion 100x objective which focuses it on the sample. Sample chambers were mounted on a nanometer-resolution digital piezo stage. Light exiting the sample chamber was collected by a condenser

and then clipped by an iris<sup>9</sup>. The clipping gives us sensitivity to axial motion of the microsphere, which broadens or narrows the beam as shown in Fig. 1.3. The beam is then projected onto a dual-axis position-sensitive photodiode which monitors total intensity ( $z$  calibration) and deflection of the beam in the lateral directions ( $x$  and  $y$  calibration). A camera provides an image by which we can visually monitor the position of the microspheres.

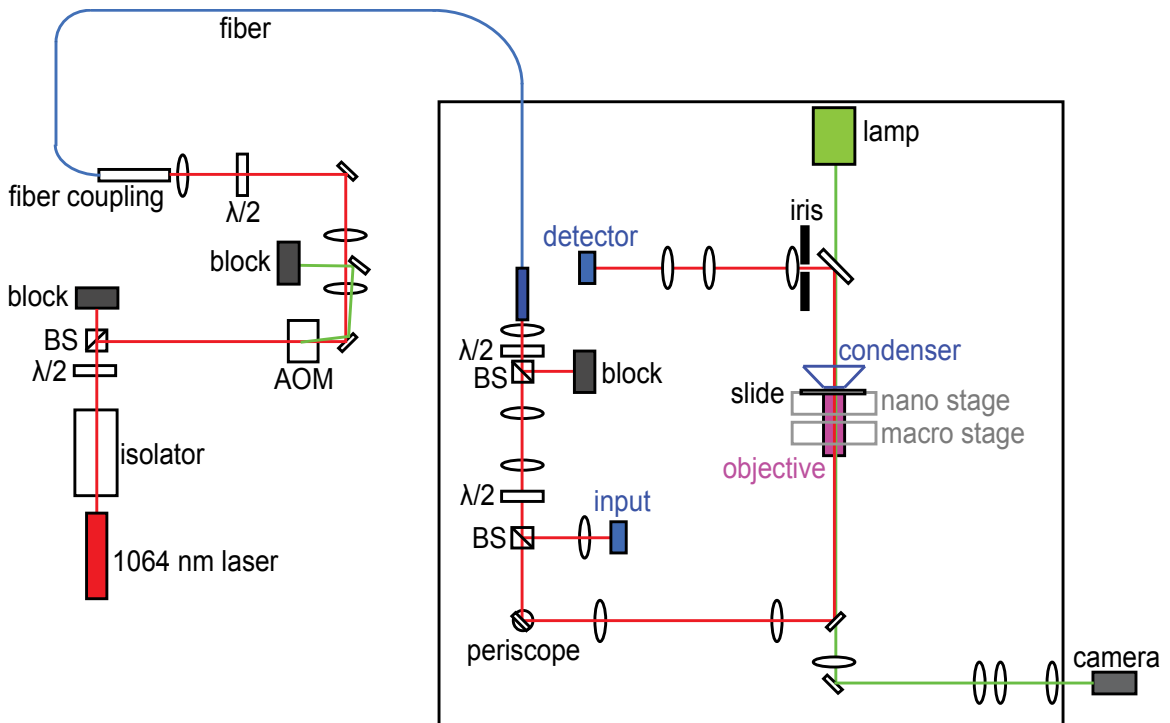
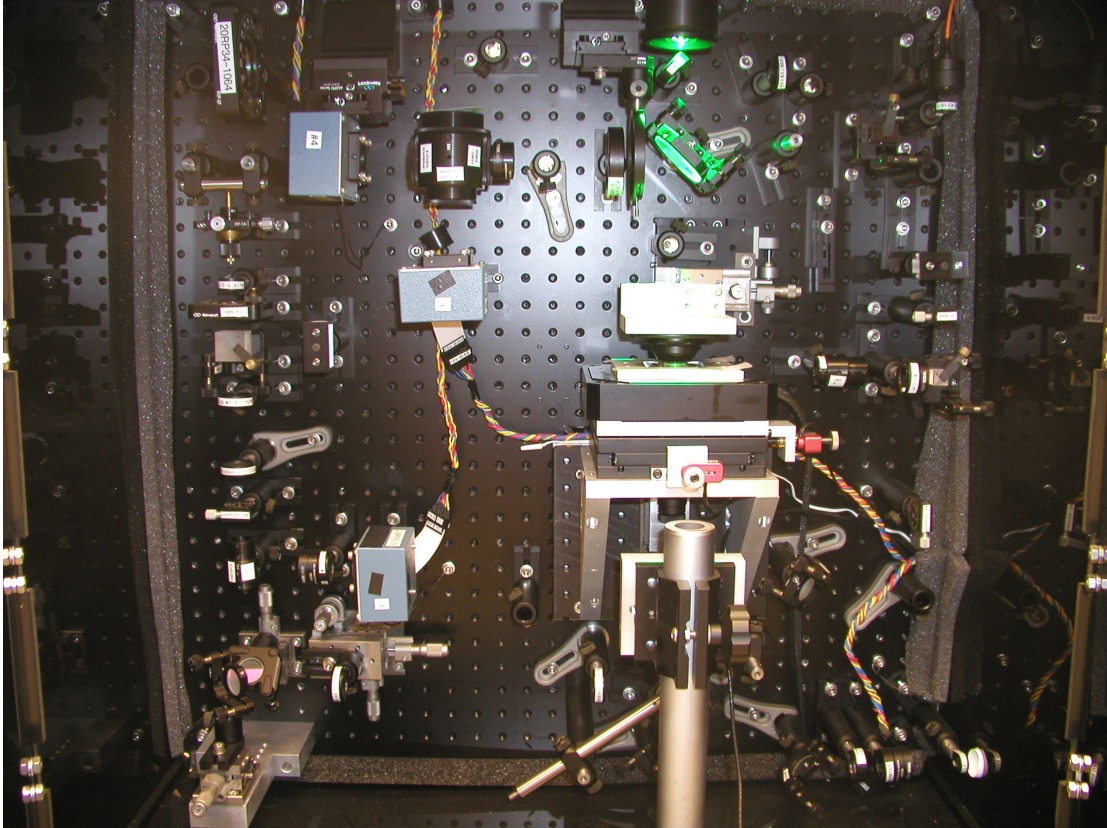


Figure 1.4 Schematic of the optical trap used for this research. A 1064 nm infrared laser passes through an acousto-optic modulator (AOM) which can be used to control the laser power. The laser is then coupled into an optical fiber before being sent to the primary setup, which is mounted on a single backboard for stability. After passing through the sample chamber, the emerging laser is collected on photodiode detectors allowing determination of force and position on the microsphere.  $\lambda/2$  indicates half-wave plates which rotate the linear polarization of the laser. Polarizing beam splitters (BS) split the beam into two linearly polarized components. The beam dumps are indicated by *block*.





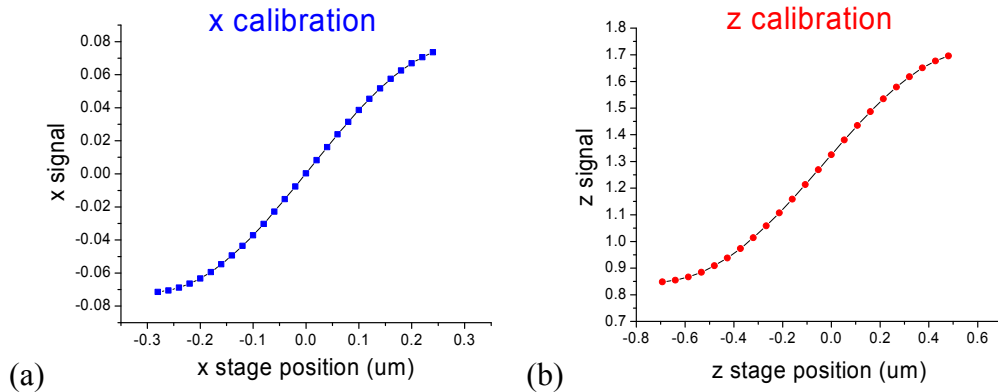
**Figure 1.5** Image of the optical trap used for all the experiments in this thesis. The boxed region of Fig. 1.4 corresponds to the layout shown in this photo. The optics are mounted on a vertical backboard.

### ***Principles of Calibration***

Calibration is the method by which we convert voltage changes on the photodiode due to momentum conservation (Fig. 1.3) into position of the microsphere relative to the trap center and force applied to the microsphere. The voltages measured are the total intensity of the beam before entering the objective ( $V_{\text{input}}$ ), the  $x$  and  $y$  signals ( $V_x$  and  $V_y$ ) on the final detector which are proportional to the deflected momentum of the beam in the  $x$  and  $y$  directions, and the total laser intensity on the

final detector ( $V_{\text{sum}}$ ). We define the normalized  $x$ ,  $y$ , and  $z$  signals as  $x_s \equiv V_x/V_{\text{sum}}$ ,  $y_s \equiv V_y/V_{\text{sum}}$  and  $z_s \equiv V_{\text{sum}}/V_{\text{input}}$ .

Calibrating the position may be done by several methods<sup>10</sup>. For this research, position was calibrated by creating slides with microspheres in 100 mM KCl, 10 mM sodium acetate (pH~4.5), which causes the microspheres to be firmly stuck to the surface. A single microsphere is moved relative to the trap by controlling the nanometer-accurate digital piezo stage. Position of the stage and voltages on the photodiodes are recorded providing a mapping from voltage to position. Several beads may be scanned and averaged to reduce error from bead variance. Figure 1.6 shows one example of the  $x$  and  $z$  calibrations though these were frequently updated. The data points are fit to a 5<sup>th</sup>-order polynomial which defines the position calibration. Note that the position calibration is approximately linear for small displacements from the trap center.



**Figure 1.6 Calibration scans for the  $x$  direction (a) and  $z$  direction (b). The relationship between position and signal defines the position calibration.**

Calibrating force is more complex and this standard method illustrates important physics related to the optical trap so we will outline it here<sup>10</sup>. We consider

the  $x$  direction with identical results following for the  $y$  direction. Subtleties for the scattering force, which acts primarily in the  $z$  direction, will be discussed in great detail in chapter 8. To first order, the trap behaves as a harmonic spring and the force may be written  $F_x = -\alpha x$ . Since the energy of a harmonic spring has the quadratic form  $E_x = \frac{1}{2}\alpha x^2$ , we may relate RMS fluctuations in  $x$  to the trap stiffness with the equipartition theorem, giving:

$$\frac{1}{2}\alpha\langle x^2 \rangle = \frac{1}{2}k_B T \quad 1.1$$

Equation 1.1 illustrates that the stiffness can be determined independently of the Stokes drag, which would have significant error due to hydrodynamic effects given the microsphere's proximity to the surface for these experiments. Implementation of Eq. 1.1 is achieved by fitting the  $x$  fluctuations in frequency space. We begin by defining the Langevin equation for position of the bead in the  $x$  direction  $x(t)$ , where  $F_T(t)$  represents the Brownian thermal noise,  $\alpha$  is the trap stiffness to be determined, and  $\beta$  is the drag on the bead.

$$F_T(t) - \alpha x(t) - \beta \dot{x}(t) = 0 \quad 1.2$$

Applying a Fourier transform gives:

$$F_T(\omega) - \alpha x(\omega) - i\omega\beta x(\omega) = 0 \quad 1.3$$

Now we can calculate the power spectrum of  $x$  fluctuations as:

$$|x(\omega)|^2 = \frac{|F_T(\omega)|^2}{\alpha^2 + \omega^2 \beta^2} \quad 1.4$$

It is more natural to contemplate frequency space so we convert using  $\omega = 2\pi f$ . The thermal noise is approximately constant across frequency space (white noise) so we define a constant parameter  $A$  giving:

$$|x(f)|^2 = \frac{A^2}{f_o^2 + f^2} \quad f_o = \frac{\alpha}{2\pi\beta} \quad 1.5$$

We now integrate to determine the RMS of fluctuations in  $x$  and relate this to the equipartition theorem using Eq. 1.1.

$$\langle x^2 \rangle = \int_0^\infty |x(f)|^2 df = \frac{A^2 \pi}{2f_o} = \frac{k_B T}{\alpha} \quad 1.6$$

We now have expressions for both  $A$  and  $f_o$ :

$$f_o = \frac{\alpha}{2\pi\beta} \quad A^2 = \frac{k_B T}{\pi^2 \beta} \quad 1.7$$

Beads are captured in the trap and the  $x$ ,  $y$ , and  $z$  fluctuations are measured, Fourier transformed, and then fit to determine  $A$  and  $f_o$  for each axis. Using Eq. 1.7 we can determine the drag  $\beta$  and stiffness  $\alpha$ . Since the force is proportional to change in voltage  $\Delta V$ , we may write  $F = -\alpha s \Delta V$ , where the sensitivity  $s$ , in units of  $\mu\text{m}/\text{Volts}$ , is determined by the relationship between position and voltage shown in Fig. 1.6. During a calibration scan where we collect the RMS fluctuation data, the microsphere has relatively small displacements ( $\sim 10$  nm) and the linear approximation is valid. We may subsequently apply the force calibration for larger displacements under the assumption that force will be proportional to change in voltage.

## Chapter 2: Energy Landscape of a Harmonically Constrained Biopolymer<sup>a</sup>

### ***Introduction***

In this chapter we will introduce a new method of measuring energy landscapes. We first show that the technique works using simulations. Then we demonstrate the technique experimentally, using a model DNA hairpin system with well-understood properties. We will demonstrate that under many circumstances, conducting experiments using our method is highly advantageous and provides better resolution of the energy landscape. This method forms the foundation for many of the experiments that will be discussed in subsequent chapters of this thesis.

### ***Background***

Pioneering studies have shown that the probability distribution of opening length for a DNA hairpin, recorded under constant force using an optical trap, can be used to reconstruct the energy landscape of the transition<sup>11,12</sup>. However, measurements made under constant force are subject to some limitations. Under constant force a system with a sufficiently high energy barrier spends most of its time in the closed or open conformation, with relatively few statistics collected in the transition state region. We describe a measurement scheme in which the system is driven progressively through the transition by an optical trap and an algorithm is used to extract the energy landscape of the transition from the fluctuations recorded during this process. We illustrate this technique in simulations and demonstrate its

---

<sup>a</sup>Adapted from de Messieres, M., Brawn-Cinani, B. & La Porta, A. Measuring the folding landscape of a harmonically constrained biopolymer. *Biophys. J.* **100**, 2736-2744 (2011).

effectiveness in experiments on a DNA hairpin. We find that the combination of this technique with the use of short DNA handles facilitates a high resolution measurement of the folding landscape of the hairpin with a very short measurement time.

The concept of a transition state, or activation complex, which limits the rate at which chemical reactions proceed, dates to the late 19th and early 20th century<sup>13-15</sup>. Later work by Kramers<sup>1</sup> provided an explicit model for chemical reaction kinetics in terms of thermally driven diffusion on an effective energy landscape. The one-dimensional domain in which the diffusion is assumed to occur is the reaction coordinate, a variable with dimension of length that parameterizes the physical rearrangement necessary for the reaction to occur. The height of the landscape is the effective free energy of the system as a function of the reaction coordinate. In this formalism, transition states are peaks in the energy landscape that impede diffusion between the initial and final states along the reaction coordinate.

Kramers was motivated by simple chemical reactions that could be assumed to occur by a well-defined pathway. However, this model has been extremely successful in describing the chemical kinetics of biological macromolecules, which often undergo complex conformational dynamics in the process of folding into their active conformation or catalyzing a chemical reaction. Although the system typically explores a high-dimensional phase space in the course of the transition, it is often possible to characterize it in terms of a well-defined effective path that connects the initial and final conformation<sup>16-19</sup>.

The development of single-molecule manipulation techniques (based on optical trapping, magnetic force, atomic force microscopy, and fluorescence) has made it possible to monitor progression of a folding/unfolding transition of a biomolecule, opening a new window into Kramer's model<sup>10,20</sup>. One widely used technique employs optical tweezers. The molecule in question is localized and tagged with a microsphere that is held in an optical trap. If the conformational change required to complete the reaction pulls the microsphere away from the center of the trap, the molecule will have to perform an additional portion of work to complete the reaction. Furthermore, the progress of the reaction can be monitored by observing the motion of the microsphere with respect to the trap center. If the optical trap is configured to produce constant force  $F_o$ , opposing the motion, and a displacement  $\Delta x$  occurs in the course of the reaction, a term  $-F_o\Delta x$  is added to the free energy of the system. The effect is to uniformly tilt the energy landscape and modify the transition state and final state energy by a well-defined amount<sup>21</sup>. Large forces can also result in an alteration of the precise location of the transition state along the reaction coordinate, depending on the shape of the barrier<sup>22</sup>.

There are two methods by which measurements made under an external force have typically been used to characterize kinetics of a folding/unfolding transition. The first is referred to as dynamic force spectroscopy (DFS). In DFS, the height of the barrier and its effective distance from the initial state are determined from the dependence of reaction rate on the applied force, or the dependence of disruption force on the rate at which force is increased (the loading rate). A particularly powerful implementation of this technique is to fit the distribution of disruption

forces as a function of loading rate to the predicted functional form<sup>21-25</sup>. The results of this analysis, however, are difficult to interpret unless it can be assumed that the reaction rate is limited by a single barrier. Furthermore, the parameters obtained will depend on the assumed shape of the barrier.

Another technique, which is especially useful in studying the folding/unfolding dynamics of biopolymers, is to apply a constant external force  $F_{1/2}$  that is sufficient to bring the initial and final conformation of the system to a state of equal occupancy. In this case, the system makes repeated transitions between the two conformations<sup>26</sup> and a histogram of the reaction coordinate — which is defined as the end-to-end extension of the polymer itself — is compiled and used to calculate the probability density function (PDF) defined as  $p(x)$ <sup>11,12</sup>. The energy as a function of the reaction coordinate  $x$  is then obtained using

$$E(x) = -k_B T \ln(p(x)) \tag{2.1}$$

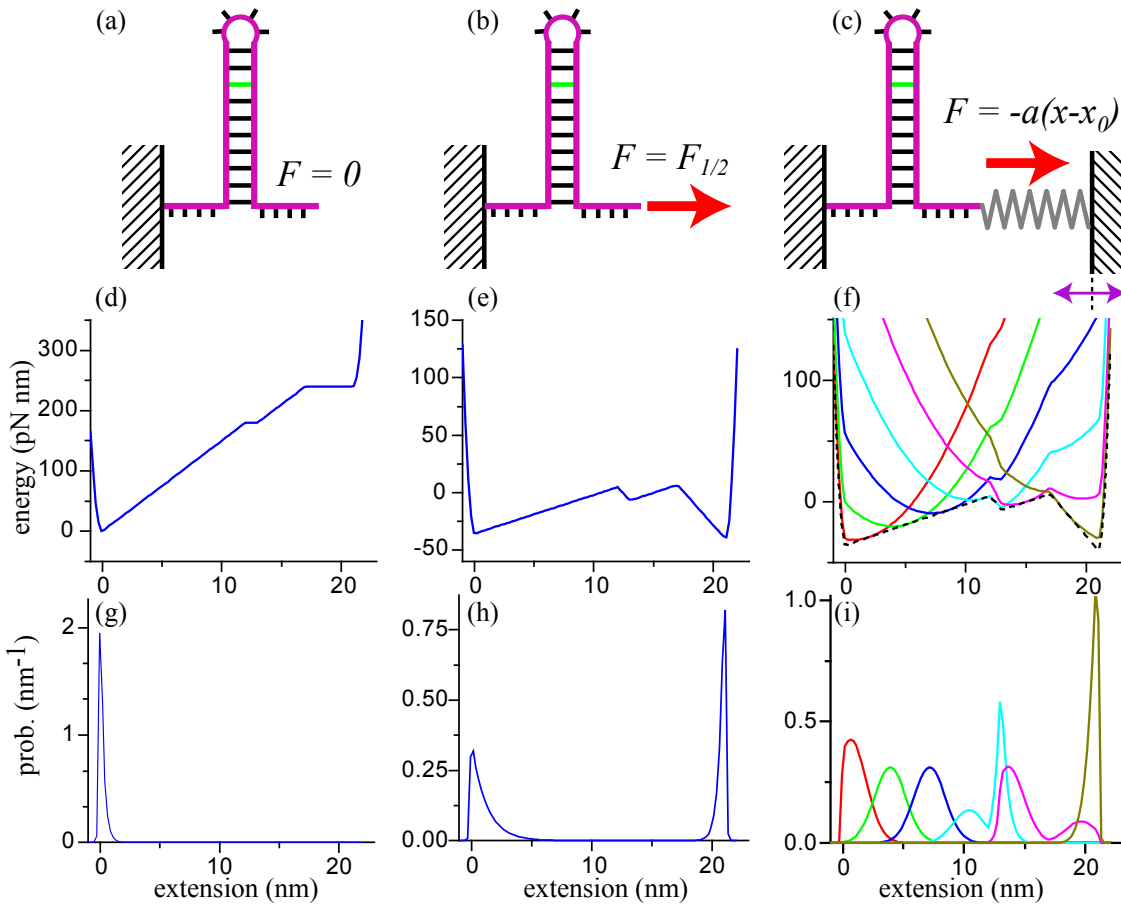
Although the transition rate at zero force is often the quantity most relevant to the biological function of a molecule, it has been argued that the detailed energy landscape is more useful in estimating this rate than a measurement of the transition rate itself as a function of force<sup>27</sup>. This is because the transition rate depends not only on the characteristics of the energy landscape, but also on the effective diffusion constant for evolution of the reaction coordinate. The attachment of handles to the structure, normally necessary for single-molecule manipulation studies, can impede movement of the structure in solution and reduce this diffusion constant. As a result, an extrapolation of measured transition rate to zero force can dramatically underestimate the native zero-force transition rate<sup>27</sup>. An experimental measurement



of the energy landscape combined with an estimate of the diffusion constant obtained through numerical simulations has been suggested as a method of obtaining an accurate estimate of the native zero-force transition rate<sup>27</sup>.

### ***Harmonic Constraint Technique***

The question remains as to the most effective method of measuring the energy landscape of a folding transition. Measurement at constant force is illustrated in Figs. 2.1a and 2.1b. In Figs. 2.1d and 2.1e the energy landscape for folding of a nucleic acid hairpin structure is represented at zero force, and for the force  $F_{1/2}$ . The expected PDF of the reaction coordinate is shown in Figs. 2.1g and 2.1h. For illustrative purposes, it has been assumed that disruption of the structure requires constant energy per base-pair (bp) except for a weak base-pair in the stem close to the loop (corresponding to a mismatch) as illustrated in Figs. 2.1a and 2.1b, creating a recognizable feature in the energy landscape near the transition state. For sufficiently high transition state energy the hairpin at  $F_{1/2}$  is expected to exhibit infrequent transitions between the open and closed conformation. As a result the time required for the system to reach equilibrium between the open and closed conformation may become impractically long, and the hairpin will spend a very small fraction of its time in the vicinity of the transition state.



**Figure 2.1** (a, b, c) Schematics for experimental configurations in which the hairpin evolves with no external force, for a constant external force  $F_{1/2}$  such that the system has equal probability of occupying the open and closed conformations, and subject to a harmonic constraint characterized by stiffness  $\alpha$  and origin  $x_0$  which tends to constrain the system along the reaction coordinate  $x$ . (d, e, f) Model energy landscapes constructed under the assumption that the energy of the hairpin decreases a constant amount for each base-pair hybridized, except for one base-pair (depicted in green) which hybridizes with no change in energy. Landscapes are shown (d) for zero force, (e) for force  $F_{1/2}$  and (f) under harmonic constraint for six different choices of the constraint origin. (g, h, i) Probability density function of position for the corresponding energy landscapes.

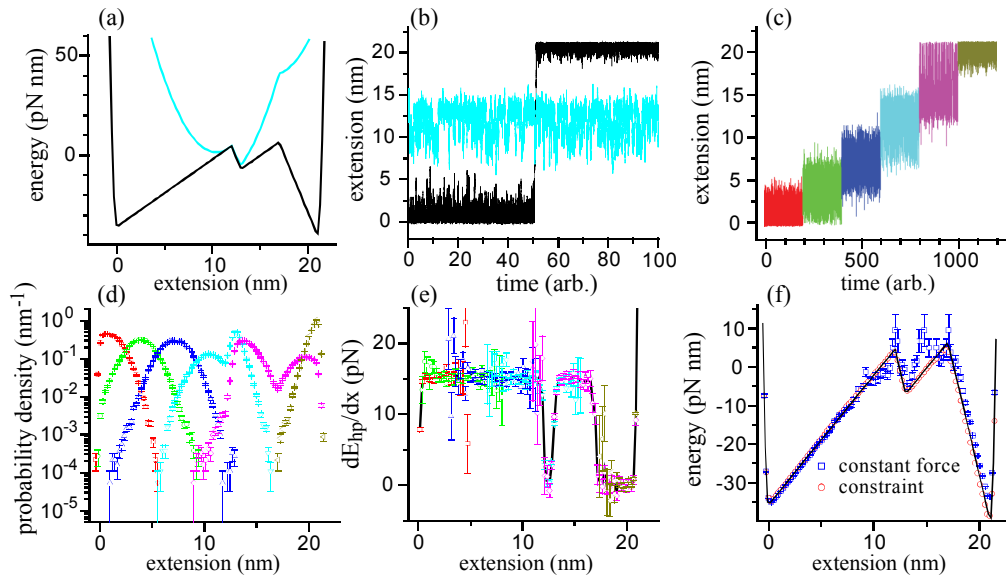
However, the energy landscape for a folding transition can also be accurately constructed by taking data with the system subject to force which depends linearly on the end-to-end extension of the molecule, so that when the extension increases the tension decreases, and vice versa. Such a harmonic restoring force can be created by taking advantage of the linear-force regime of the optical trap itself, and this experimental configuration has been employed in previous experiments to measure transition rates of hairpins at different constraint positions<sup>28,29</sup>. The harmonic constraint manifests itself as a parabolic term superimposed on the native energy landscape of the polymer structure and tends to confine the system to a well-defined region along the reaction coordinate. The native energy landscape in this region can be obtained from the constrained data by correcting for the effect of the imposed constraint. The effect is similar to the umbrella sampling technique which is frequently used in molecular dynamics simulations to enhance statistical sampling of rarely visited conformations<sup>30</sup>.

Although harmonic constraint tends to restrict the system to a limited range of the reaction coordinate, the full landscape can be constructed by taking a series of runs with different origins, confining the system to a set of overlapping regions along the reaction pathway. The effective energy landscapes for six such constraint origins are shown in Fig. 2.1f. A key feature of this configuration is that if the constraint is sufficiently strong, the large energy barrier which separates the open and closed conformations is no longer the dominant feature of the effective landscape. The predicted PDFs are shown in Fig. 2.1i, and reflect the parabolic shape of the constraining potential superimposed on the hairpin's native energy landscape. In

contrast to the constant-force measurement, every region of the energy landscape is visited with high probability for at least one value of the constraint center, resulting in a much more uniform convergence of statistics. This is in contrast to the PDF obtained from the constant-force landscape, Fig. 2.1h.

The advantage of harmonic constraint relative to constant force is apparent in Langevin dynamics simulations (Fig. 2.2) based on the energy landscapes in Fig. 2.1e and 2.1f. A single simulation was run at constant force for a time equal to the combined time of six individual harmonic constraint simulations. In Fig. 2.2a the effective energy landscape at constant force  $F_{1/2}$  is contrasted with the landscape for one of the constrained runs restricting the system to remain near the center of the transition. Fig. 2.2b shows simulated dynamics governed by the two potentials in Fig. 2.2a. The trajectory at constant force, which is dominated by the large barrier between the open and closed conformations, shows only one transition across the barrier. This time interval deliberately selects the one observed transition from the full run of 1200 time units. The trajectory taken from the constrained system shows rapid fluctuations within the region favored by the constraint potential. The energy barrier that dominates the dynamics at constant force has two distinct disadvantages compared with the constrained landscape. The first is that the position PDF calculated from the simulation at  $F_{1/2}$  (Fig. 2.2b black curve) will show poor convergence in the transition region because the probability of occupying this region is extremely low. The second difficulty is that a system starting in the closed conformation will take a relatively long time to traverse the barrier and sample the open conformation (and vice versa), despite the fact that thermodynamically, the open

and closed conformations should be equally populated. Aside from non-uniform sampling of the reaction coordinate, the system at constant force requires a long time to manifest thermal equilibrium. Even if the system can be sampled at an extremely high rate, it is necessary to wait for the system to equilibrate. The constrained system, in contrast, equilibrates more quickly, facilitating measurement in a short time interval.



**Figure 2.2** (a) Comparison of the native energy under constant force (black) and for a harmonically-constrained force (cyan) that constrains the system to remain between the open and closed conformations. (b) Simulated trajectory of the hairpin held with constant external force and under the constraining force, corresponding to the two potentials in (a). The constant-force trace is selected from a run of 1200 time units and the constrained trace is selected from a run of 200 time units. (c) Simulated trajectories of the hairpin subject to the harmonic constraint for six positions of the constraint origin. (d) Probability density functions calculated from each trajectory in (c) and displayed in matching color. (e)  $dE_{hp}/dx$  constructed from the data in (d). The black curve is the combined  $dE_{hp}/dx$  curve calculated from Eq 2.1. (f) Energy landscape calculated from constant-force data (blue squares) and harmonic constraint (red circles) compared with the effective energy (black line), evaluated at  $F_{1/2}$ .

Although the individual runs taken with the harmonic constraint converge quickly, no single run samples the full transition. Figure 2.2c-f demonstrates the construction of the full energy landscape from a series of runs whose combined domains of convergence span the entire transition. The array of origins used to create the effective landscapes in Fig. 2.1f is chosen so that the corresponding PDFs show good convergence in overlapping domains of the reaction coordinate, such that every position on the reaction coordinate is well-converged in at least one of the PDFs. If we were to apply Eq. 2.1 to each of these curves we would obtain the total energy, including the contribution of the harmonic constraint. However, if the strength of the constraint  $\alpha_c$  and the origin  $x_c$  are sufficiently well known, the effect of the constraint can be subtracted to obtain the underlying energy landscape,

$$E_{\text{hp}}(x) = -k_{\text{B}}T \ln(p(x)) - \frac{\alpha_c}{2}(x - x_c)^2 + A_c \quad 2.2$$

where  $E_{\text{hp}}$  is the energy of the hairpin we intend to measure and  $A_c$  is an arbitrary constant. However, for each value of the constraint position  $x_c$  this equation will give  $E_{\text{hp}}(x)$  which is well converged only within the narrow domain visited by the system. Combining these constrained energy landscapes to form the global energy landscape would require us to choose the constants  $A_c$  to form a continuous energy landscape spanning the individual runs. Although this can be done using standard techniques<sup>31</sup> the difficulty of determining the constants  $A_c$  is circumvented by calculating the rate of change of  $E_{\text{hp}}$  with respect to the reaction coordinate,

$$\frac{dE_{\text{hp}}}{dx}(x) = -k_{\text{B}}T \frac{d}{dx} \ln(p(x)) - \alpha_c(x - x_c) \quad 2.3$$

Since the constants  $A_c$  have been eliminated, the global  $dE_{\text{hp}}/dx$  can be obtained by averaging the individual curves, giving each term statistical weight inversely proportional to the square of the standard deviation of the term<sup>32</sup>. This is accomplished by evaluating

$$\frac{dE_{\text{hp}}}{dx}(x_i) = \frac{\sum_j \left[ -k_B T \frac{d}{dx} \ln(p_{x_j}(x_i)) - \alpha_c(x_i - x_j) \right] p_{x_j}(x_i)}{\sum_j p_{x_j}(x_i)} \quad 2.4$$

where  $x_i$  is the point at which  $dE_{\text{hp}}/dx$  is being calculated,  $x_j$  is the origin of the harmonic constraint for run  $j$ , and  $p_{x_j}(x_i)$  is the probability the system is observed at  $x_i$  when the constraint origin is  $x_j$ . Fig. 2.2d shows the probability distributions  $p_{x_j}(x_i)$  and Fig. 2.2e shows  $dE_{\text{hp}}/dx$  calculated for individual constraint origins (symbols with error bars) based on the trajectories in Fig. 2.2c. as well as the combined  $dE_{\text{hp}}/dx$  (black curve). Note that the individual  $dE_{\text{hp}}/dx$  curves have large error bars at the margins of their range, but the combined curve is constructed with high resolution throughout, because every part of the reaction coordinate is well converged in at least one of the runs. The function  $dE_{\text{hp}}/dx$  is a useful way to characterize the transition, and can be integrated along  $x$  to obtain the energy itself.

Figure 2.2f shows the energy landscape (plotted at  $F_{1/2}$ ) obtained from the simulated constant-force measurements and from integration of  $dE_{\text{hp}}/dx$  for the constrained measurements, compared with the landscape used as the input to the two simulations. The same total amount of measurement time is simulated for both experiments, but the constrained measurement is well converged over the entire landscape, while the constant-force data shows poor convergence in the critical

transition region. Equally good results can, in principle, be obtained from the constant-force configuration, although a  $\sim 100$  fold longer measurement time is required.

### ***Implications***

One might be tempted to conclude from the comparison of the two measurement methods that the constant-force measurement can be made to equal the constrained version as long as the measurement time is extended sufficiently. Although this is true in the idealized conditions of these simulations, in real experiments the constrained method has advantages that cannot generally be duplicated by extending the measurement period of a constant-force data set. One example would be a system which undergoes an irreversible transition. In such a system the constrained measurement could be used to probe the energy landscape in the neighborhood of the irreversible transition, adjusting the constraint to prevent the system from committing to the transition.

More generally, the ability to make measurements in a shorter time interval allows a wider latitude in choosing the measurement technology. The long measurement times necessary to reach equilibrium in constant-force measurements have motivated researchers to shun single-beam optical trapping assays in favor of a dual-beam assay<sup>8,11,33</sup>. This configuration, in which the construct is stretched between two particles held in two optical traps originating from the same laser, offers unequalled long-range stability, making it possible to record trajectories for many minutes without suffering significant drift. However, the trap centers must be maintained at a distance comparable to the optical wavelength to prevent the two



traps from effectively merging into a single trap. The double-stranded DNA handles used to attach the structure are typically chosen to have a total length of ~1700 base-pairs or more to accommodate this geometry<sup>11</sup>. Thermally driven fluctuations in the extension of these handles is a source of uncertainty in the measurement of the conformation of the system.

Taking advantage of the fact that the constrained measurements can be completed in a short time interval, we have adopted a single-beam surface-based axial-pulling configuration in which the construct is stretched between a surface attachment and the bead. The vertical-pulling geometry allows us to use short handles to couple the nucleic acid construct to the surface and bead, with total length of 428 base-pairs. This results in a tighter coupling between the structure being studied and the bead, and less uncertainty in the measurement of the system conformation. This is expected to result in higher resolution in the determination of the energy landscape, compared with measurements made with longer handles. Ultimately the choice of handle length will have to take into account other effects, such as the increase in viscous drag that occurs when a bead closely approaches a fixed surface or another bead, or interactions between the bead and the structure being studied. Recently, optical trapping experiments have made use of handles as short as 29 base-pairs, offering the possibility that even higher resolution could be achieved<sup>34,35</sup>.

## **Materials and Methods**

### **Sample Preparation**

Two double-stranded DNA handles were generated by PCR using biotin and digoxigenin labeled primers (Invitrogen, Gaithersburg, MD) for attachment to the surface and bead respectively. Handles were digested with BstXI or BtgI (New England Biolabs, Ipswich, MA) and gel extracted, with final lengths of 234 base-pairs and 194 base-pairs, excluding the 4 base-pair overhang resulting from the digestion.

Two different hairpins were synthesized (Invitrogen, Gaithersburg, MD) with a flanking sequence of TTTT on each end for improved ligation efficiency, and 4 base overhangs of cgtg and cgat corresponding to the BtgI and BstXI enzymes respectively. We will refer to the hairpins as Short (cgtgttttgagtcacagctctggatcctgttttcaggatccagacgttgactcttttcgat) and Long (cgtgttttccgcgcatctgagtcgaggcagtagccgctcgtctcgcatttttttcgagacgacggctactgcctcgactcagatgcgcggttttcgat). (The short hairpin was previously studied<sup>11</sup>.) The handles were ligated to the hairpins, gel extracted, and diluted to approximately 50 pM in the final experimental buffer (50 mM sodium phosphate buffer pH 7.0, 50 mM NaCl, 10 mM EDTA, and 0.02% Tween 20<sup>36</sup>).

The sample chamber cover slips were scrubbed with Windex, rinsed, and dried. Polyclonal anti-digoxigenin (Roche Molecular Biochemicals, Indianapolis, IN) was diluted to 20 µg/ml in a phosphate buffer solution (137 mM NaCl, 8.1 mM Na<sub>2</sub>HPO<sub>4</sub>, 1.8 mM KH<sub>2</sub>PO<sub>4</sub>, 2.7 mM KCl) and incubated in the chamber for 20 minutes. To prevent bead and DNA interactions with the slide, a blotting buffer was incubated in the slide for 3 cycles of 20 minutes each, containing 1 mg/ml blotting-

grade blocker (Bio-Rad, Hercules, CA) in popping buffer. The ligated hairpin constructs were then incubated in the sample chamber for 45 minutes for attachment of the DNA dig label to the anti-digoxigenin on the surface. Carboxyl microspheres (Bangs Laboratories, Inc., Fishers, IN) with radius 410 nm were streptavidin coated with a PolyLink Protein Coupling Kit (Polysciences, Inc), diluted to 0.1 g/ml in popping buffer, and incubated for 20 minutes for attachment to the biotin labels on the DNA. A final flow through left the sample in the final experimental buffer (50 mM sodium phosphate buffer pH 7.0, 50 mM NaCl, 10 mM EDTA, and 0.02% Tween 20<sup>36</sup>) and oxygen-scavenging solution (721 ug/ml glucose oxidase (Sigma, St. Louis, MO), 144 ug/ml catalase (Sigma), and 3.9 mg/ml glucose)<sup>37</sup>. We found that the oxygen scavenger was essential for working at the high laser powers used.

## **Instrumentation**

Data was collected on a single-beam optical trap using the experimental configuration shown in the Fig. 2.4a insert. A 1064 nm laser (BL-106C, Spectra-Physics, Santa Clara, CA) was coupled to a single-mode, polarization-preserving optical fiber (Thorlabs, Newton, NJ). The trap was constructed with a NA 1.49 oil immersion objective (CFI Apo TIRF 100x, Nikon Instruments, Lewisville, TX). Light was collected with a NA 1.4 oil immersion condenser (Nikon Instruments, Lewisville, TX) on a dual-axis position-sensitive diode (DL100-7PCBA3, Pacific Silicon Sensor, Westlake Village, CA). A fraction of the beam incident on the objective lens was also measured to monitor optical power. An acousto-optic modulator (Isomet, Springfield, VA) was used to either stabilize the power of the beam or maintain constant force using software feedback. Data was collected at 60 kHz on a 16-bit digital acquisition

board (National Instruments, Austin, TX). The analog signal was filtered using an 8-pole Bessel filter at 30 kHz (Krohn-Hite, Brockton, MA), and filtered offline at 20 kHz. An iris after the condenser clipped the beam to provide signal dependent on the bead height<sup>9</sup>.

## Calibration

The double-stranded DNA handles were modeled as a WLC (worm-like chain)<sup>38</sup> with total contour length 160 nm determined by assuming 0.338 nm/bp<sup>39</sup> with an additional 15 nm added to estimate the size of the biotin and dig labels, stretch modulus of 1205 pN, and persistence length 43.1 nm<sup>40</sup>. Single-stranded DNA was modeled as a WLC using average values from Woodside et al. of persistence length 1.25 nm, contour length per base of 0.625 nm, and neglecting enthalpic contributions<sup>11</sup>. The measured force-extension curves are fit to the worm-like chain theory so our results are relatively insensitive to the precise values, except for the contour length per base of single-stranded DNA, which determines the conversion of opening distance (nanometers) to number of bases released. This is accounted for by fitting a scaling correction to the opening distance (~2%) which is discussed in the results.

The optical trap was calibrated using standard techniques based on the Fourier spectrum of Brownian fluctuations and scans using a bead rigidly attached to the surface of the sample chamber<sup>41</sup>. The position calibrations were adjusted to reproduce the previously reported opening distance of the short hairpin<sup>11</sup>. The vertical trap stiffness was measured to be  $249 \pm 3$  pN/ $\mu$ m (standard deviation of the mean) with 860 mW of power on the back of the overfilled objective.

## Experimental Setup

We studied hairpins in either force-clamp configuration or harmonically-constrained configuration. For the force clamp, a software feedback loop controlled the power of the trap through the acousto-optic modulator to maintain a constant force on the hairpin as it fluctuated between the open and closed conformations. For the constrained configuration, we lowered the stage at a constant rate to sweep across the hairpin from the closed conformation to the opened conformation. During constrained scans, the feedback loop controlling the acousto-optic modulator held the trap power constant. A normalized  $z$  signal was calculated as the intensity measured at the detector divided by the intensity incident on the objective<sup>9</sup>. We recorded the normalized  $z$  signal, trap power, and stage position for the analysis.

## Theoretical Model

We compared our results to a simple theoretical model for the free energy of the hairpin, based on Mfold v.3.2, (using a monovalent salt correction<sup>42</sup> of 140 mM at 24°C) which gives the energy released by sequential hybridization of the base-pairs<sup>43,44</sup>. For a given extension of the structure, the number of base-pairs hybridized is determined, taking into account the extension of the single-stranded DNA released at the average observed force. The energy is the combination of the hybridization energy specified by Mfold and the elastic energy of the single-stranded DNA released. There is some uncertainty in the increase in extension when the hairpin loop opens. We adjusted this parameter to reproduce the observed well depth for the fully open hairpin. To approximate the blurring effect due to fluctuations of the single-stranded DNA, we followed previous research<sup>12</sup> and applied a pseudo-convolution, in

which the width of the Gaussian kernel varies across the landscape by  $\sigma = \sqrt{kT/\alpha}$ . The varying stiffness  $\alpha$  was calculated using the worm-like chain model for the single-stranded DNA released from the hairpin at the average measured force for each bin. This specifies the intrinsic energy landscape of the hairpin. This procedure is an approximation and refinements to this method based on first principles are discussed in chapter 4.

To account for fluctuations in the apparent extension of the structure arising from thermal fluctuations of the double-stranded DNA handles and trapped bead, we convolved the probability distribution derived from the energy landscape model with a Gaussian of uncertainty 1.1 nm, determined by fitting to the closed well of the hairpin. The model is similar to that employed by Woodside et al, except we account for different parts of the landscape being measured at different average force for constrained measurement. An example of the original model with single-stranded elasticity applied, the convolved model, and experimental data are shown in Fig. 2.4e and 2.4f.

## ***Results & Discussion***

### **Harmonic Constraint Data Analysis**

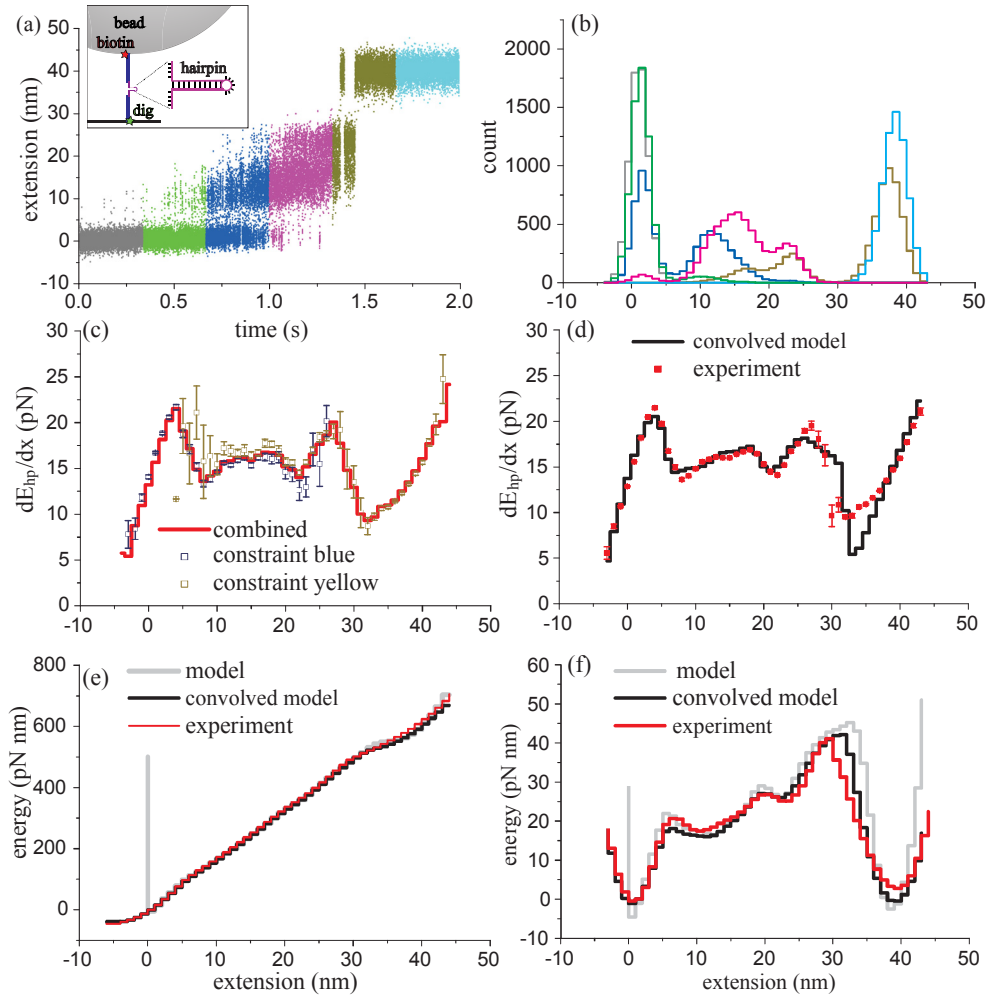
To implement the harmonic-constraint method, we lowered the stage at a constant rate averaging 20 nm/s. The resulting records of opening distance vs. time were divided into six equally spaced time intervals where each interval is approximated to represent a constant position. Using a greater number of intervals makes this approximation more exact but does not significantly change our final results. Fig. 2.4a shows the resulting data for one scan of the long hairpin (structure

shown in Fig. 2.3). The position of the trap (defined as the point where force on the bead is zero) and the location of the surface define the relationship between the harmonic constraint and the hairpin. For each interval, the double-stranded DNA handles were approximated as a linear spring using the WLC stiffness at the average force for that interval. The stiffness of the trap was assumed to be constant for all measurements. The trap and the DNA were treated as two linear springs in series with a combined stiffness  $\alpha_c = (1/\alpha_{DNA} + 1/\alpha_{Trap})^{-1}$ , where  $\alpha_c$  represents the effective stiffness of the harmonic constraint. The energy of the system including the constraint follows Eq. 2.3, where  $E_{hp}$  includes the energy for stretching the single-stranded DNA after it is liberated from the hairpin. A scaling correction (which averaged 2%) was applied to each data set so that the total base-pairs released matched the expected value. A set of probability distributions of the measured opening distance was accumulated in equally spaced 1.0 nm bins (Fig. 2.4b).

A discrete version of Eq. 4 gives us a measurement of  $dE_{hp}/dx$  for each interval. As in the simulation, each interval only provides accurate results for one region of the landscape. In Fig. 2.4c we have shown two of these intervals (blue and yellow) for clarity. Blue provides good statistics in the middle region of the landscape while yellow provides good statistics for the regime when the hairpin is mostly open. The energy landscape for this hairpin is more irregular than that assumed in the simulation, so the regions of good statistical convergence are more widely distributed in the experimental data. The combined graph of  $dE_{hp}/dx$  was calculated using Eq. 2.4 and in Fig. 2.4c red we show the combined result of the 6 intervals from the first scan. In Fig. 2.4d we show the combined results (red) for 3 scans on the same sample



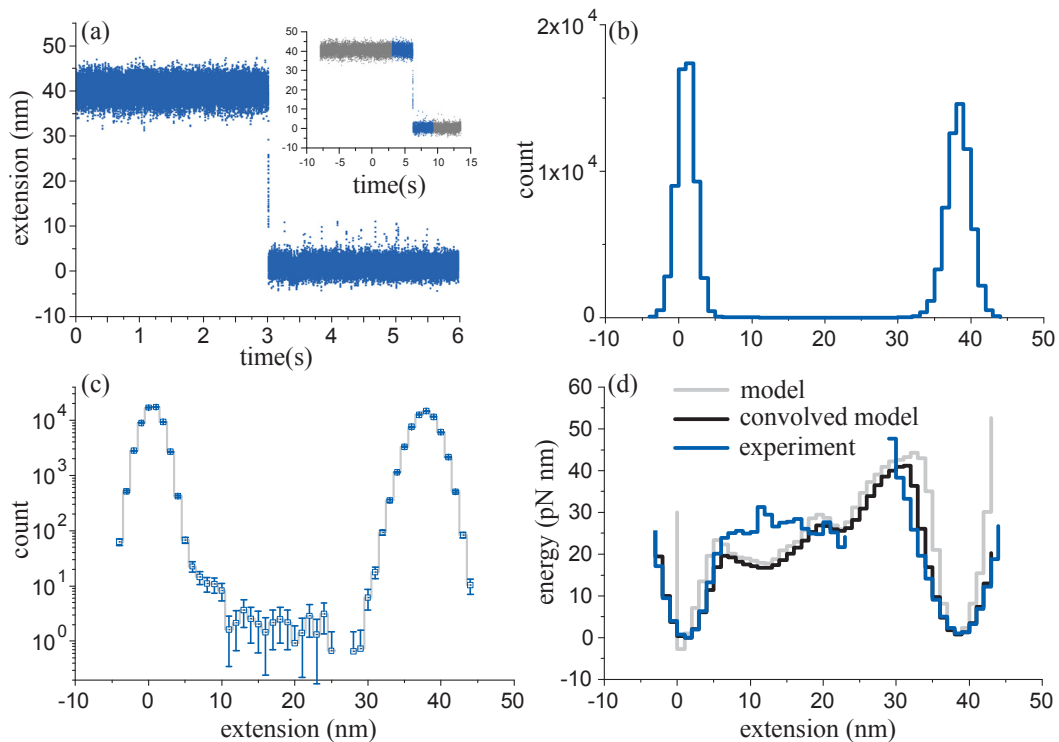




**Figure 2.4** Harmonic constraint applied to a single long DNA hairpin. The trap position was moved for 3 scans of duration 2 s. (a) Opening distance for the first scan divided into 6 intervals. Inset shows the experimental configuration, with force applied in the vertical direction. Double-stranded DNA handles with total length 428 base-pairs were used to attach the hairpin between the surface and the trapped bead using digoxigenin and biotin labels respectively. (b) Probability distribution with 1 nm bins for the 6 intervals of the scan shown in (a). (c) Change in energy per nm shown for two of the intervals (blue and yellow) along with the combined result for the 6 intervals shown in a (red). (d) Change in energy per nm for the combined result of the 3 scans (red) compared to the predictions of the convolved model (black). (e) The model after accounting for single-stranded elasticity (gray), the convolved model (black), and experimental results (red). (f) The same data shown in (e), except plotted at  $F_{1/2}$ .

## Force Clamp Data Analysis

To compare with the harmonically-constrained data analysis, we held the same hairpin sample in force-clamp mode and measured data for a time period equal to the total time of the constrained data sweep. The opening distance is plotted over 6 s (Fig. 2.5a) giving the probability distribution (Fig. 2.5b). At  $F_{1/2}$  ( $\sim 15$  pN) the mean time between transitions is much larger than the proscribed measurement time. In order to obtain at least one transition during a measurement period we began with the system in the open conformation and reduced the force to 14.2 pN, which increases the rate of transitions to the closed conformation. The system remained in the open conformation for 14 seconds before making a transition to the closed conformation and we selected a 6 s interval centered on the transition for analysis. The probability distribution can be compared with the harmonic-constraint result (Fig. 2.4b). We also show the log of the probability with uncertainties (Fig. 2.5c). The final landscape was reconstructed directly from Eq. 2.5, correcting for the energy for stretching the single-stranded DNA. (Fig. 2.5d) Since the system made only one transition, the fraction of time spent in the open and closed conformations is arbitrary and the statistical convergence of the transition state is extremely poor. The shape of the energy landscape in the neighborhood of the open and closed conformations is statistically converged and well defined, but the relative energy of the open and closed conformations and the shape of the transition state landscape are not determined by this data.



**Figure 2.5** Force-clamp data for the same hairpin sample measured in Fig. 2.4. (a) Opening distance measured as the hairpin makes one transition during 6 s (Compare to Fig. 2.4a). Inset: Shows the same data over a longer time frame and the portion selected for analysis. (b) Probability distribution of the opening distance with 1 nm bins shown in a (Compare to Fig. 2.4b). (c) Log of the probability distribution with uncertainties. (d) Reconstructed energy landscape (blue), the model (gray), and convolved model (black) (Compare to Fig. 2.4f).

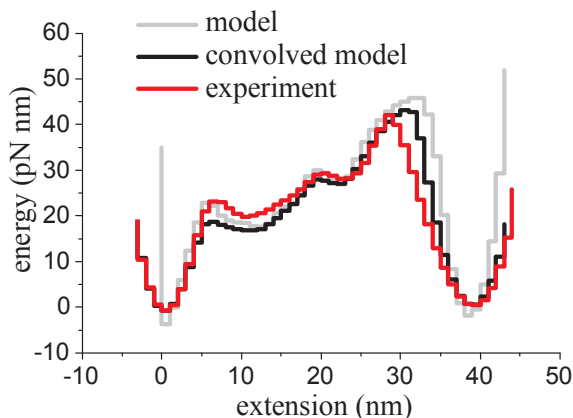
## Combined Samples

In Fig. 2.6 we show the combined results for 5 unique hairpin molecules, all measured on the same slide. Each hairpin contributed one scan of 6 intervals with a bin spacing of 1 nm. These intervals were combined with appropriate weights by the same method we used to combine intervals for the single hairpin sample in Fig. 2.4.

Some of our long hairpin samples deviated from the model with an initial opening force significantly below forces required to open the remainder of the

hairpin. These features were reproducible on subsequent scans and we attributed them to mutations in the original synthesis of the 99 base-pair hairpin, contamination or deterioration of the sample, or optical damage. We assumed that any defects would lower rather than increase the hybridization energy and therefore implemented a simple scoring system to delete hairpins which had anomalously low total opening energy. We scored each sample as the highest force measured before it opened beyond 5 base-pairs. The top 40% of samples were found to be self-consistent and we combined them to produce the average landscape shown in Fig. 2.6.

In Fig. 2.6 we find good agreement between the measured landscape and the model which includes the expected blurring resulting from of the double-stranded DNA handles. The loss of resolution compared to the intrinsic landscape of the hairpin is 1.1 nm, which corresponds to approximately one base-pair opening of the hairpin.



**Figure 2.6 Energy landscape (red) produced by combining scans of 5 different data samples, the model (gray), and convolved model (black). Data analyzed with bin spacing 1.0 nm and 6 intervals for each scan.**

## ***Discussion***

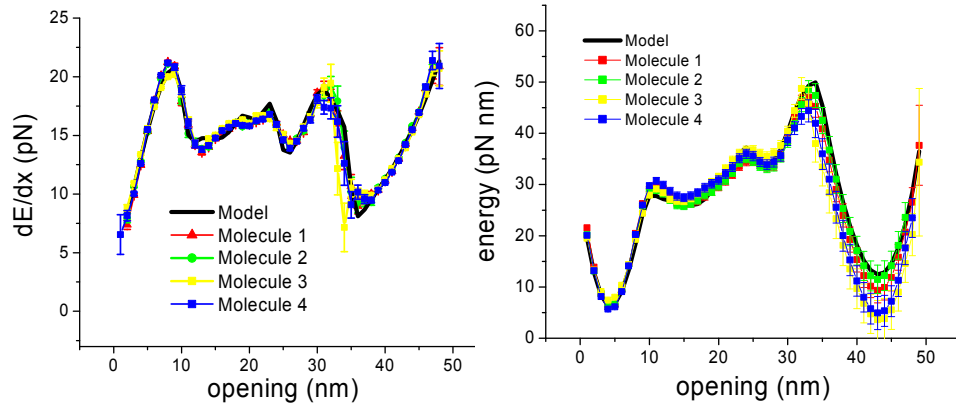
The principal benefit of collecting data under a harmonic constraint rather than constant force occurs in cases where the system has a large energy barrier between the open and closed conformation. Provided the constraint can be made strong enough, the effective barrier will be smaller for the constrained system and it will equilibrate much more quickly under the harmonic constraint than under constant force. Since the occupancy of the transition state is proportional to  $\exp(-E_t/k_B T)$ , where  $E_t$  is the transition state energy, a dramatic decrease in the transition time can be achieved from even a modest reduction in the transition state energy. In such cases the harmonic-constraint method allows the landscape to be reconstructed from an ensemble of constrained measurements with good statistical convergence and high resolution using a much smaller total measurement time than would be necessary for a constant-force measurement (Compare Figs. 2.4d and 2.5d or see Fig. 2.6).

The ability to complete the measurement in a short time reduces the sensitivity to drift, and allows us to use a simplified experimental configuration in which the structure is attached to the trapped bead using a relatively short pair of double-stranded DNA handles. Fluctuations in the length of the handle are a significant source of uncertainty in the instantaneous state of the system and cause blurring of the measured energy landscape. Comparison of our measured landscape with the Mfold-based model suggests that fluctuations of the handle introduce blurring equivalent to convolution of the true landscape with a Gaussian function of standard deviation 1.1 nm. Neglecting end effects, thermal fluctuations in the length of the handles are expected to scale as the square root of the length. Since our handles are a factor of 4

shorter than those used in double-beam experiments by Woodside et. al<sup>11</sup>, this implies that our resolution should be better by a factor of 2. Woodside employed deconvolution to compensate for handle effects. However, the landscapes reported by Woodside in the absence of deconvolution show a level of blurring consistent with this estimate. We expect that further reduction in the handle length, either in the vertical-pulling optical trap assay or in other single-molecule assays, will further increase the resolution of the measurement of the energy landscape.

## ***Conclusion***

The method we have presented demonstrates that quantitatively accurate energy landscapes can be obtained from data taken with strong harmonic constraint by removing the effect of the constraint and combining data taken at different constraint positions. We have demonstrated this technique in a single-beam optical trapping experiment but we anticipate that it can be applied to measurements made using other technologies, including magnetic tweezers, or atomic force microscopy, provided that the harmonic constraint can be implemented with sufficient stiffness and the measurement resolution is adequate. The harmonic constraint can facilitate more rapid measurement and may also be preferred in situations where the system is near an effectively irreversible transition. As a result, this technique may broaden the scope of energy landscape reconstruction to systems which currently can only be studied by DFS techniques.



**Figure 2.7** Data from multiple samples of the long hairpin analyzed relative to a new model which will be developed in Chapter 4.

In this chapter we used the calculation of  $dE/dx$  as an intermediate step to determine the energy landscape of the molecule. The results have been presented as in our original publication. However, in Chapter 3 we will develop an analysis of mRNA structure based solely on the  $dE/dx$  values and discuss advantages of this viewpoint. In Chapter 4 we will discuss an improved model which accounts for discrepancies between the data and the original model. As an example, Fig. 2.7 shows data from this chapter for four different molecules analyzed relative to the new model.

## Chapter 3: CCR5 mRNA Unfolding Pathways<sup>b</sup>

### ***Introduction***

In the previous chapter we developed a new method of reconstructing energy landscapes. As an intermediate step, we calculated  $dE/dx$ , or the slope of the energy landscape. We now consider experiments to determine RNA structure and find that looking at the  $dE/dx$  values as a final result rather than an intermediate step is an exceptionally useful way to classify the multiple conformations we observe. Therefore there are two important results we will describe in this chapter. The first is biological and illuminates the structure of CCR5 mRNA. The second is technical and demonstrates that calculating  $dE/dx$  is a useful general technique for interpretation of single-molecule results.

### ***Background***

The ribosome translates messenger RNA (mRNA) into protein by moving along the mRNA molecule and adding an amino acid to the nascent protein chain for every three-base codon it encounters<sup>45,46</sup>. Secondary or tertiary structure in the mRNA during translation can create a physical barrier, requiring the ribosome to generate additional force in order to translocate<sup>47-51</sup>. The presence of such a barrier can dramatically increase the probability that the ribosome will backslide and shift into an alternate reading frame, in which a different set of codons is recognized<sup>52</sup>. The detailed mechanism by which frameshifting is induced remains unknown. Here we employ optical trapping techniques to show that CCR5 mRNA, which encodes the

---

<sup>b</sup>Adapted from de Messieres, M., Chang, J.-C., Belew, A. T., Meskauskas, A., Dinman, J. D. & La Porta, A. Single-molecule measurements of the CCR5 mRNA unfolding pathways. *In preparation* (2012)

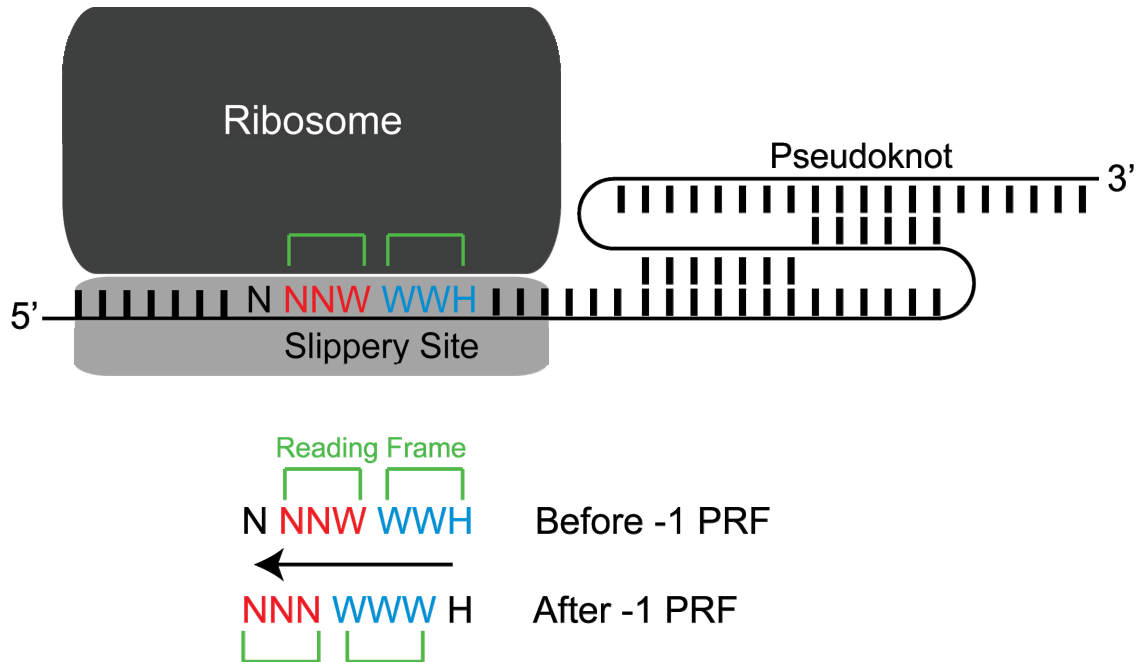


co-receptor for HIV-1<sup>53</sup>, manifests several distinct unfolding pathways when subject to end-to-end force, one of which is consistent with a proposed pseudoknot conformation<sup>54-57</sup>, and another of which we have identified as a folding intermediate. We find that the introduction of micro RNA (miRNA) molecules<sup>58</sup>, that have been found to increase CCR5 frameshifting *in vivo*<sup>57</sup>, changes the relative occupancy of these conformations, raising the possibility that miRNAs regulate frameshifting by binding to the mRNA molecule and driving it towards specific conformations. Our study illuminates the relationship between mRNA structure and frameshifting by measuring the amount of mechanical energy the ribosome must expend to disrupt the downstream mRNA. CCR5 mRNA is the first identified human programmed -1 ribosomal frameshift (-1 PRF) of non-viral origin<sup>57</sup>, but frameshift signals are also common in viral genomes<sup>59-63</sup>, so developments in this field may have broad impact in the treatment of human disease.

Extensive research has demonstrated that two important features of the mRNA, a slippery site and a downstream pseudoknot structure, can promote efficient -1 PRF<sup>55,64</sup>. First we define these concepts followed by discussion of computationally predicted structures for the mRNA. Then we discuss our optical trap studies where we repeatedly stretched individual molecules of CCR5 mRNA. The single-molecule experiments are an approximation of the unfolding kinetics the mRNA would exhibit while being pulled into the ribosome's active site. We compare our measurements to the computationally predicted structures and draw conclusions regarding a pseudoknot, an intermediate conformation, other low-force conformations, and the influence of miRNA-1224 on the occupancy of the multiple conformations observed.

## ***Pseudoknot Structures and Slippery Sites***

Pseudoknots occur when single-stranded nucleotides in loops of mRNA structure base pair with nucleotides outside that loop<sup>55</sup>. Figure 3.1 shows a pseudoknot which has formed downstream of a ribosome positioned at a slippery site. The slippery site is identified as a heptamer of form N NNW WWH which increases the probability that the ribosome will frameshift<sup>55</sup>. The notation is IUPAC (N = any base, W = any 3 identical weak bases (A or T), H = A, C, or U). Figure 3.1 depicts the ribosome's reading frame before and after a -1 PRF, where the -1 PRF event causes the ribosome to be translocated in the 5' direction of the mRNA. The reading frame (shown in green) indicates the alignment of the ribosome with each mRNA triplet (codon). Each codon specifies an amino acid and the -1 PRF changes the definition of all subsequent codons. The presence of a pseudoknot structure approximately 6-8 nucleotides downstream of the 3' end of the slippery site can significantly enhance the efficiency of -1 PRF<sup>55</sup>.



**Figure 3.1 A -1 Programmed Ribosomal Frameshift (-1 PRF) occurs when the ribosome is positioned over a slippery site heptamer of form N NNW WWH (IUPAC notation). The ribosome translocates one nucleotide in the 5' direction of the mRNA, resulting in a -1 PRF. The presence of a downstream pseudoknot can significantly increase the probability of a -1 PRF occurring.**

### ***Theoretical mRNA Structures***

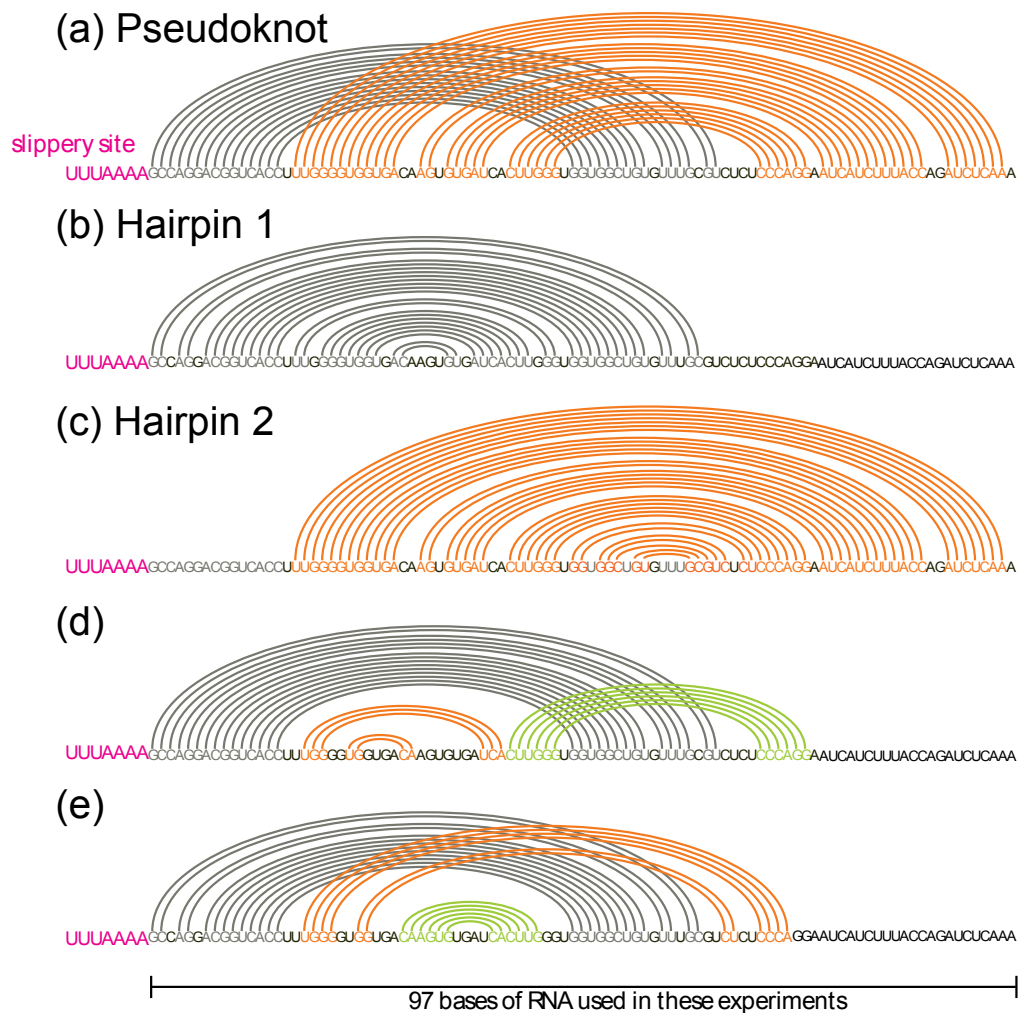
We now consider theoretically proposed structures for the CCR5 mRNA. Using the Predicted Ribosomal Frameshift Database<sup>c</sup> (PRFdb)<sup>57,65</sup> and Mfold<sup>44</sup> we obtained a set of possible conformations for the CCR5 mRNA. Fig. 3.2a represents a pseudoknot with two primary stems (black and orange) which correspond to two possible hairpins (Fig. 3.2b or Fig. 3.2c respectively). Two alternate pseudoknot conformations are also shown (Fig. 3.2d and Fig. 3.2e). We designed our optical trapping experiment to study the 97 bases of single-stranded RNA immediately

<sup>c</sup> [http://prfdb.umd.edu/detail.html?accession=NM\\_000579&slipstart=473](http://prfdb.umd.edu/detail.html?accession=NM_000579&slipstart=473)

following the UUUAAAA slippery site, as indicated in Fig. 3.2. The single-stranded RNA for our experiments is:

***GCCAGGACGGUCACCUUUGGGGUGGUGACAAGUGUGAUCACUUGGGUGG  
UGGCUGUGUUUGCGUCUCUCCCAGGAAUCAUCUUUACCAGAUCUCAAA***

Details of this construction are given in the methods section. This choice of sequence allows for all of the structures in Fig. 3.2 to potentially form in the single-stranded RNA regime.



**Figure 3.2 Structures from the Predicted Ribosomal Frameshift Database<sup>57,65</sup> and Mfold<sup>44</sup> for CCR5 mRNA. (a) Predicted pseudoknot structure. (b, c) Hairpin structures which correspond to the two stems for the pseudoknot shown in (a). (d, e) Two alternate pseudoknot structures.**

## ***Disruption Pathways***

Examples of force-extension plots are shown in Fig. 3.3 where four colors (blue, green, yellow, and red) are used to identify four distinct disruption pathways observed. In principle, these might represent different disruption pathways of a common structure, or disruption pathways for different initial structures. Black lines represent the worm-like chain model<sup>38</sup> for the handles without the CCR5 mRNA and with the full 97 bases of mRNA released<sup>26,29</sup>. Alongside each force-extension curve is the corresponding plot of rate of change in energy of the structure per extension ( $dE/dx$ ), determined by reconstructing the energy landscapes from the force-extension data using the method discussed in chapter 2. The  $dE/dx$  curve indicates the steepness of the effective energy landscape as the structure is progressively disrupted and allows us to distinguish pathways that generate similar disruption forces, but at different points along the disruption pathway. This additional information is potentially important because frameshifting has been demonstrated to be more efficient when the ribosome's active site is aligned with the slippery site<sup>55,66</sup>. The single-stranded regime of the mRNA measured in our optical trapping experiments begins directly after the slippery site so high  $dE/dx$  values which occur for small opening distance (such as in Fig 3.3a) represent high mechanical stability when the slippery site would be approximately aligned with the active site of the ribosome. Precisely correlating opening distance with the expected alignment of the active site is complicated by the fact that the mRNA has an inherent structural width which is not determined until the specific conformation is known.

We have overlaid the individual  $dE/dx$  curves corresponding to each trace (colors) on top of a log density plot (grayscale) showing all  $dE/dx$  data taken for the CCR5 mRNA under the same conditions. Darker regions on the grayscale plot represent disruption paths that occurred more frequently. By overlaying an individual  $dE/dx$  trace on the complete data set, we can put individual disruptions in a context reflecting all of the data (Fig. 3.3). Each trace is identified with a color type, determined by the regions of the  $dE/dx$  graph that the trace visits (Fig. 3.4).

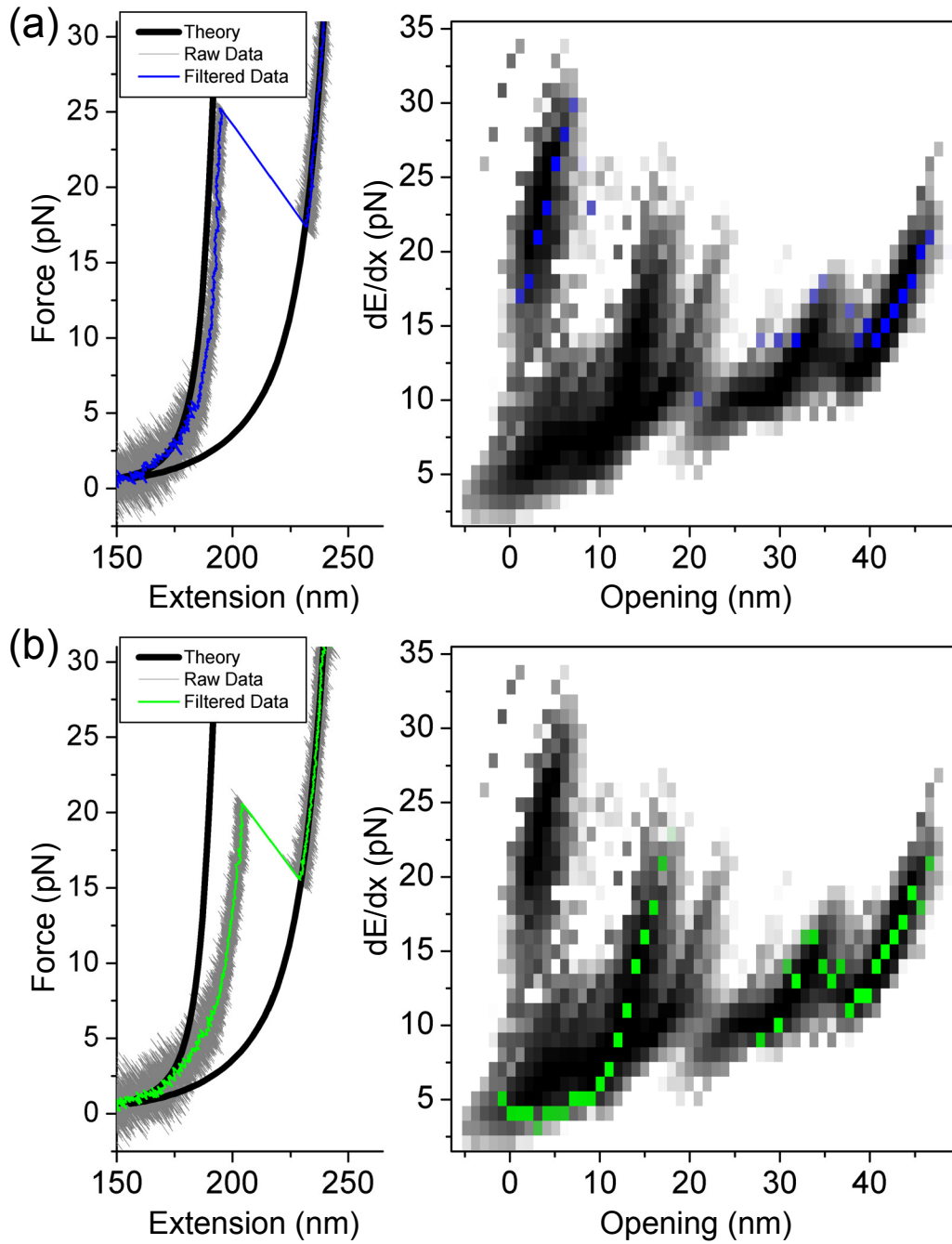


Fig 3.3 Continued next page

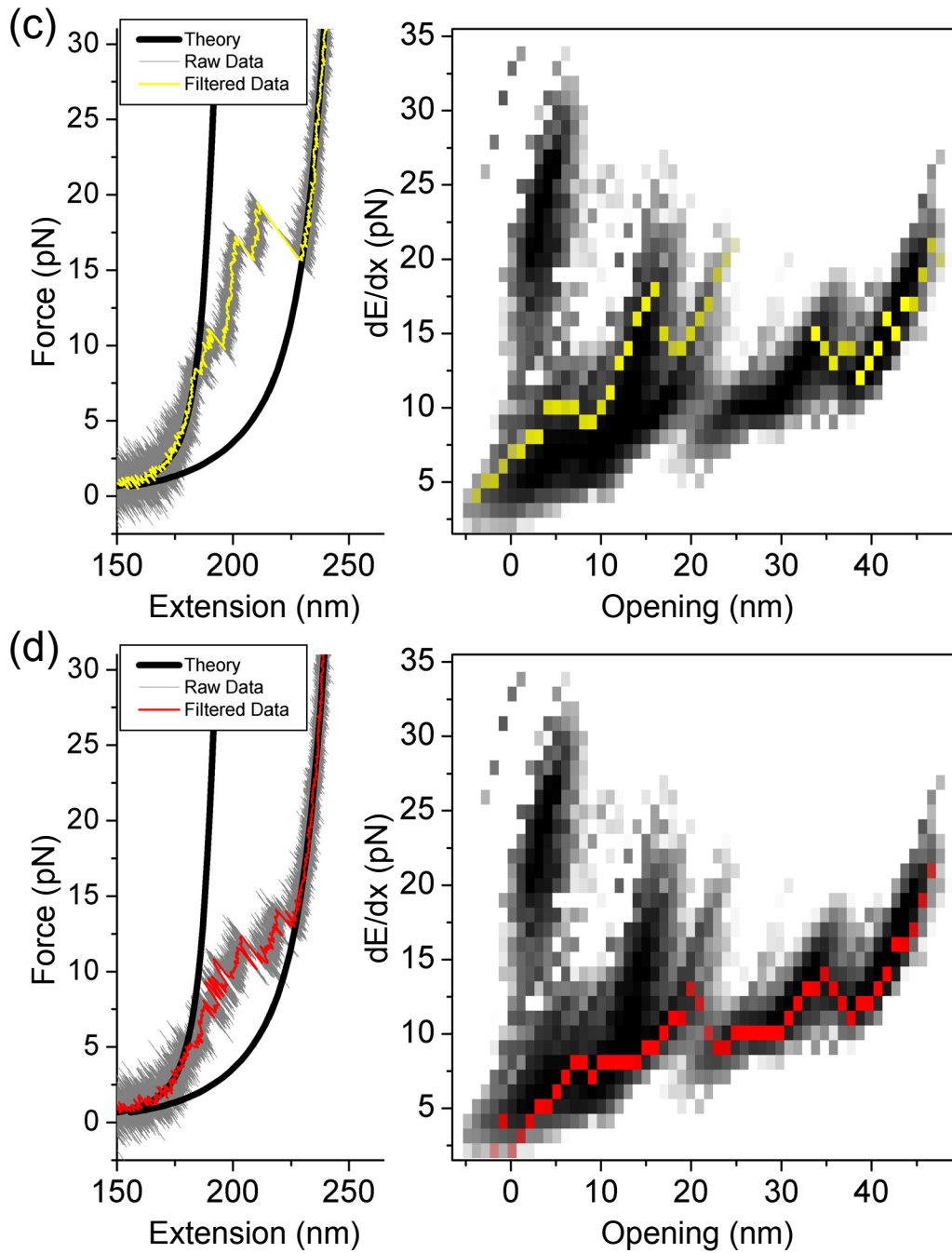
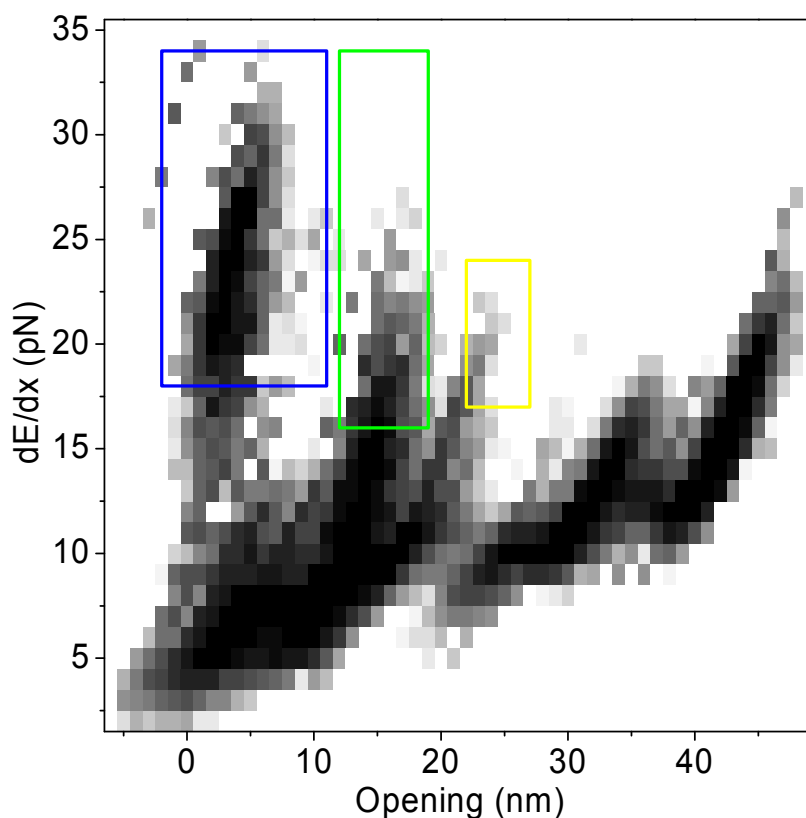


Figure 3.3 Black lines: Worm-like chain theory<sup>38</sup> for hybrid DNA/RNA handles with and without 97 additional bases of extended single-stranded mRNA. Change in energy per opening distance ( $dE/dx$ ) calculated as described previously in chapter 2. Gray scale density plot: The accumulated  $dE/dx$  probability density from all scans plotted on a log scale. Colored squares: The result from the single scan shown in the force-extension curve.

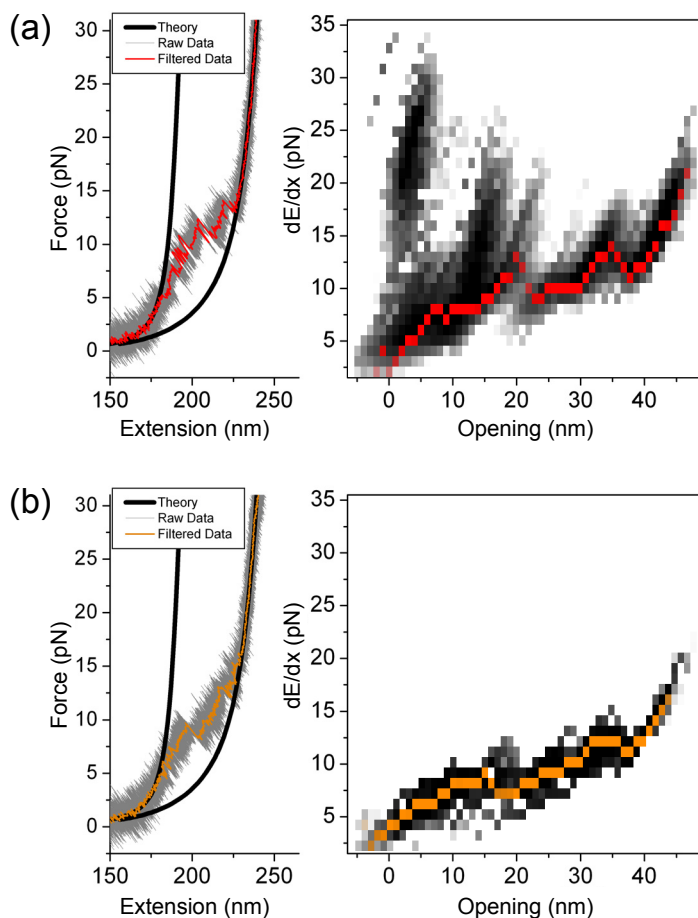




**Figure 3.4** Conformation type for individual scans was determined by testing whether  $dE/dx$  values visited one of the three specified boxes. The remaining set was assigned to be the red (low force) conformation. The choice of box location was optimized by inspecting the distribution of substeps, discussed later for Fig. 3.7a. Samples which visit more than one box were assigned in the following priority: Blue, Yellow, then Green. In particular, yellow conformations often visited the green region. Any sample which did not visit one of the boxes was defined as Red.

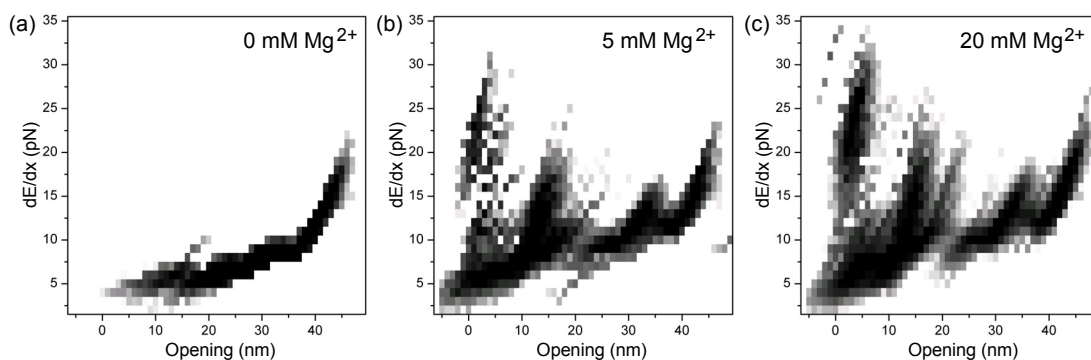
Before proceeding with quantitative analysis, we consider some general observations for the individual disruption pathways observed. Duplex RNA or DNA hairpins are generally observed to disrupt around 13-14 pN<sup>12,26</sup>. The red conformation has the lowest measured force and  $dE/dx$  levels, generally well below 13 pN and consistent with a conformation in which the mRNA has formed a hairpin

with some mismatches, such as the hairpins shown in Fig. 3.2b or Fig. 3.2c. A somewhat similar  $dE/dx$  signature is observed under reverse scans when the molecule is allowed to fold slowly, consistent with the fact that all of the substeps for the red pathway were observed to be semi-reversible (Fig. 3.5). In contrast to the red conformation, the blue pathway has significantly higher  $dE/dx$  values than would be expected for duplex RNA, suggesting the presence of tertiary structure.



**Figure 3.5 (a) Increasing force scan (red) reproduced from Fig. 3.3d. (b) Decreasing force scan (orange). The  $dE/dx$  log density plot for (b) shows the accumulation of several measured folding paths which are similar though not identical to the unfolding pathway for red shown in (a). Folding pathways (b) never exhibit the high  $dE/dx$  values measured during unfolding (a).**

All conformations were dependent on the presence of  $Mg^{2+}$  (Fig. 3.6), consistent with the association of these conformations with higher-order tertiary structure. Pseudoknot dependency on  $Mg^{2+}$  has been previously demonstrated for a different RNA sequence<sup>67</sup>.



**Figure 3.6** The relative distribution of the conformations is dependent on  $Mg^{2+}$  levels. (a) 0 mM  $Mg^{2+}$  (b) 5 mM  $Mg^{2+}$  (c) 20 mM  $Mg^{2+}$ . All data taken with 20 mM Tris pH 8.0 and 50 mM NaCl.

### ***Substeps in the Disruption Pathways***

To connect the disruption pathways with possible mRNA structures, we have plotted the distribution of disruption substeps in Fig. 3.7. Each disruption of an mRNA molecule is typically observed as a series of sub-disruptions which release portions of the mRNA (see Fig. 3.3c for example). Each sub-disruption is represented as a point on the scatter plot, where the  $x$ -axis is the amount of mRNA already released when the disruption starts and the  $y$ -axis is the amount of additional mRNA released. Blue disruptions, for example, show a single step of  $\sim 90$  bases resulting in a single point, or show a brief substep resulting in 2 points (example shown in Fig. 3.8). An alternative representation in Fig. 3.7b illustrates the average disruption pathway observed for the four conformation types, where double-ended

arrows indicate reversible steps and thinner arrows represent transitions that occurred less frequently. It is hypothetically possible for a single structure to manifest multiple pathways in Fig. 3.7b if different parts of the structure disrupted in a different order. In such a case arrows appearing in Fig. 3.7b would be of similar length but appear in a different sequence. This does not seem to be the case in our data, leading us to conclude that the distinct disruption pathways are associated with distinct folding conformations of the mRNA.

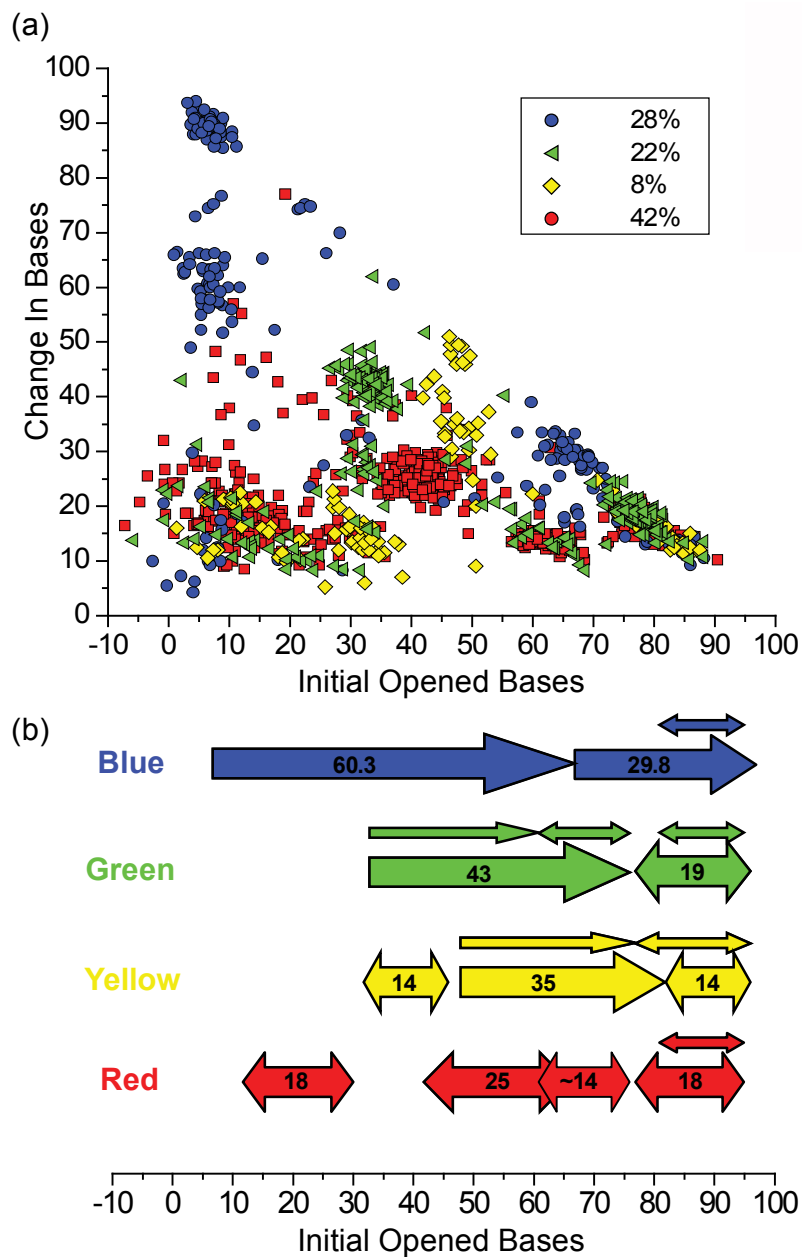


Figure 3.7 (a) An alternate view of the  $dE/dx$  data in Fig. 3.3. Each disruption path is composed of one or more substeps, plotted where  $x$  is the initial opened bases and  $y$  is the change in bases for that substep. Legend shows the percentage of 429 total scans for each conformation type. (b) Wide arrows indicate the mean behavior of the disruptions plotted in (a). Numbers indicate how many bases were released for the given substep on average. Double-ended arrows indicate substeps which were observed to be reversible. Thinner arrows indicate substeps observed with lower frequency.

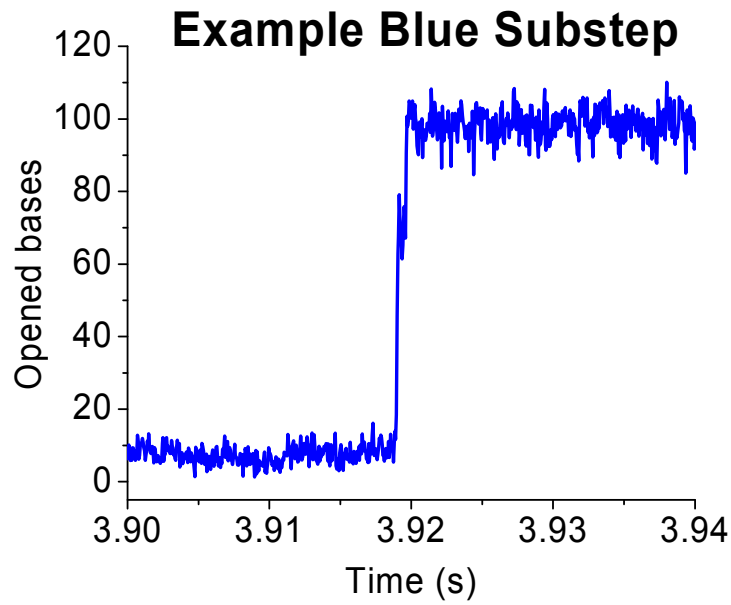
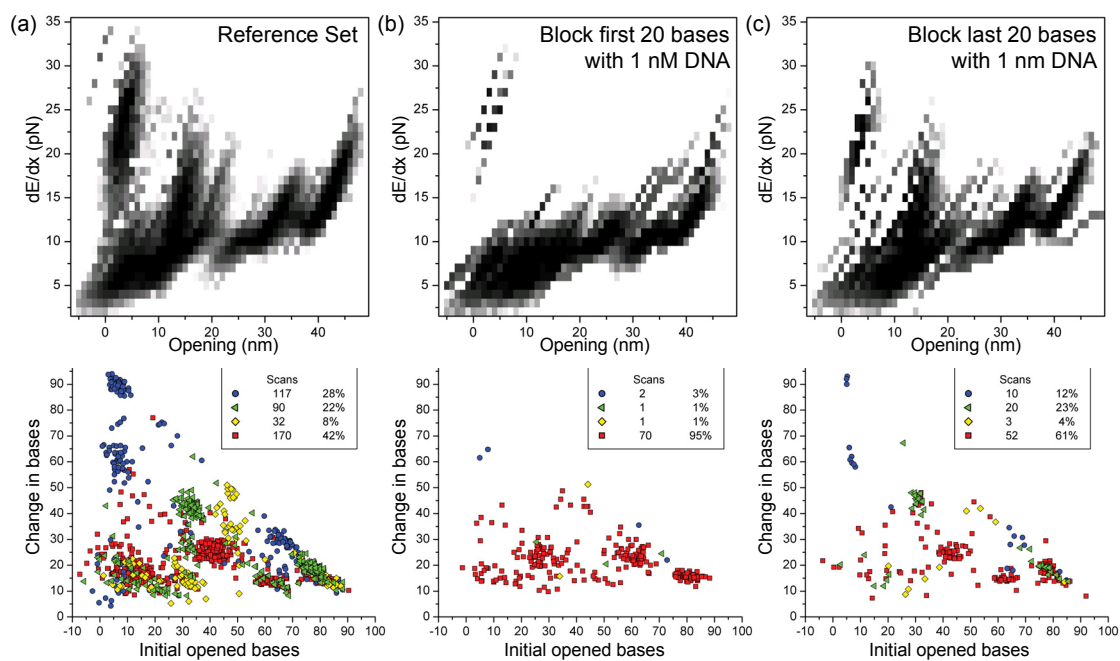


Figure 3.8 Opened bases plotted over time at 20 kHz for the blue conformation shown in Fig. 3.3a. In some cases this substep could not be resolved or was measured as a single data point. The accumulated average of these disruptions was used to generate the location of the blue substep shown in Fig. 3.7

### ***Blocking First and Last 20 Bases of mRNA with DNA Oligos***

The largest  $dE/dx$  values are observed for the blue conformation and occur before a substantial amount of mRNA has been released (Fig. 3.3a), indicating that most of the 97 bases of RNA participate in the structure. The apparent initial opening distance does not necessarily reflect released mRNA because the structure will have an inherent structural distance between the 5' and 3' ends. To test whether the blue conformation depends on the full mRNA sequence, we introduced 20 base DNA oligos complementary to either end of the mRNA sequence. The results (Fig. 3.9) demonstrate that blocking either end of the mRNA depletes the blue pathway, consistent with a computationally predicted pseudoknot structure (Fig. 3.2a) which

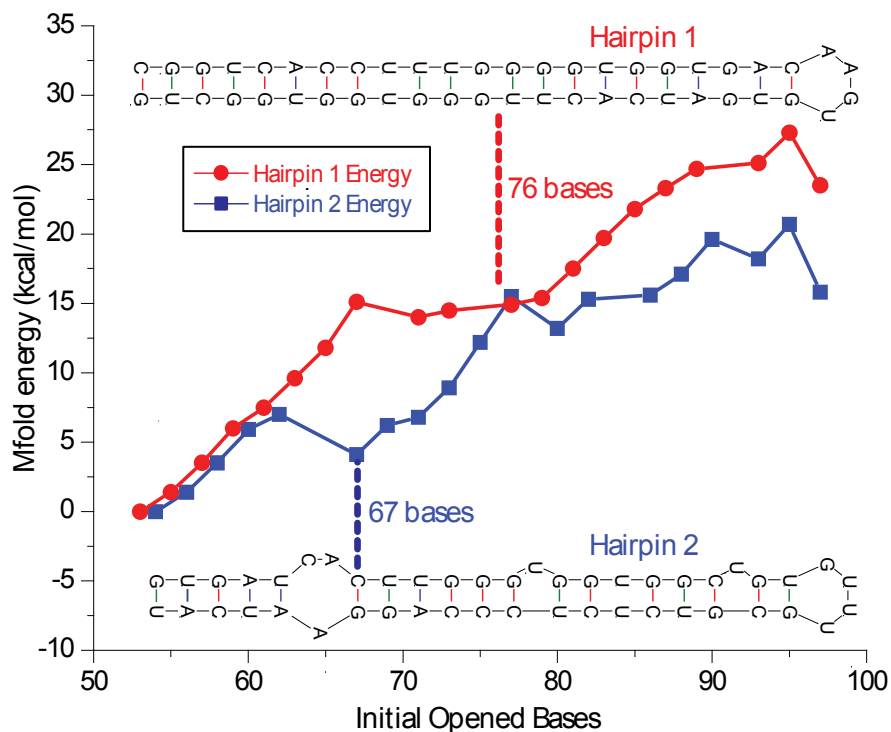
incorporates the full mRNA sequence. The high-force green conformation was suppressed when blocking the first 20 bases (5' end) but not when blocking the last 20 bases (3' end). This implies green represents a structure which does not include the last 20 bases, such as the pseudoknot shown in Fig. 3.2d or Fig. 3.2e. The pattern of substeps for the low-force red conformation changed when blocking the first 20 bases but not the last 20 bases. This indicates red is more consistent with *Hairpin 1* (Fig. 3.2b) which does not depend on the last 20 bases, than *Hairpin 2* (Fig. 3.2c) which does depend on the last 20 bases. This is also expected since *Hairpin 1* (-29.0 kcal/mol) is predicted to be more stable than *Hairpin 2* (-25.6 kcal/mol), determined by Mfold<sup>44</sup>.



**Figure 3.9 Inclusion of 1 nM 20-base DNA oligos which were complementary to either the first 20 bases (5' end) or last 20 bases (3' end) of the CCR5 mRNA. The 5' end of our single-stranded RNA begins immediately after the slippery site heptamer. (a) Control with no bases blocked. (b) Blocking the first 20 bases on the 5' end of the CCR5 sequence. (c) Blocking the last 20 bases on the 3' end.**

## Identification of the Pseudoknot Conformation

One of two hairpins (Fig. 3.2b and Fig. 3.2c) may represent the final disruption pathway of the mRNA, regardless of whether a pseudoknot initially formed. Figure 3.10 shows the energy required to disrupt the hairpins (determined by Mfold<sup>44</sup>). Only the final disruption path is shown, representing hypothetical disruption pathways after ~55 bases of RNA have been released.



**Figure 3.10** We compare the two hairpins from Fig. 3.2b and Fig. 3.2c and plot the free energy for disruption given by Mfold<sup>44</sup>. The blue and red dashed vertical lines correspond to the last transition points observed for the blue and red conformation described in Fig. 3.7.

We expect the observed blue conformation substep at 67 bases (Fig. 3.7) to correspond to a local energy minimum in Fig. 3.10. Comparison shows that the blue substep coincides with a local minimum for *Hairpin 2*, but is not compatible with *Hairpin 1* where a local maximum occurs. The reversible transition at 76 bases



(which appears for the red, green and yellow pathways in Fig. 3.7) lines up with a broad energy minimum for *Hairpin 1*. Therefore the reversible transition exhibited by the other pathways is more compatible with *Hairpin 1*, which is represented by Fig. 3.2b. Blue has high  $dE/dx$  values consistent with tertiary structure (Fig. 3.3a), depends on the full mRNA sequence (Figs. 3.7 and 3.9), and completes its disruption along a path more consistent with *Hairpin 2* than *Hairpin 1* (Fig. 3.10). These three features make the blue conformation consistent with the pseudoknot structure in Fig. 3.2a where the 1<sup>st</sup> stem corresponding to *Hairpin 1* (Fig. 3.2b) disrupts before the 2<sup>nd</sup> stem corresponding to *Hairpin 2* (Fig. 3.2c).

### ***Folding Times and Intermediate States***

Between scans the force was reduced to zero as quickly as possible by raising the stage. The force was then held at zero for 7.5 seconds, allowing the mRNA time to refold. The occupancy of the various conformations of the molecule is expected to depend on both the rate at which the force is decreased and the amount of time spent at zero force<sup>68</sup>. The occupancies are strongly dependent on refolding time, which provides further evidence that the different disruption pathways reflect different initial conformations, rather than different disruption pathways of the same initial conformation (Fig. 3.11). Blue initially has low occupancy, which increases over a measured refolding time of approximately 30 seconds. Green initially has high occupancy, which decreases at approximately the same rate that blue increases, suggesting green is a folding intermediate of blue. The red and yellow pathways vary slowly with refolding time, which may indicate that they are long-lived alternate conformations.

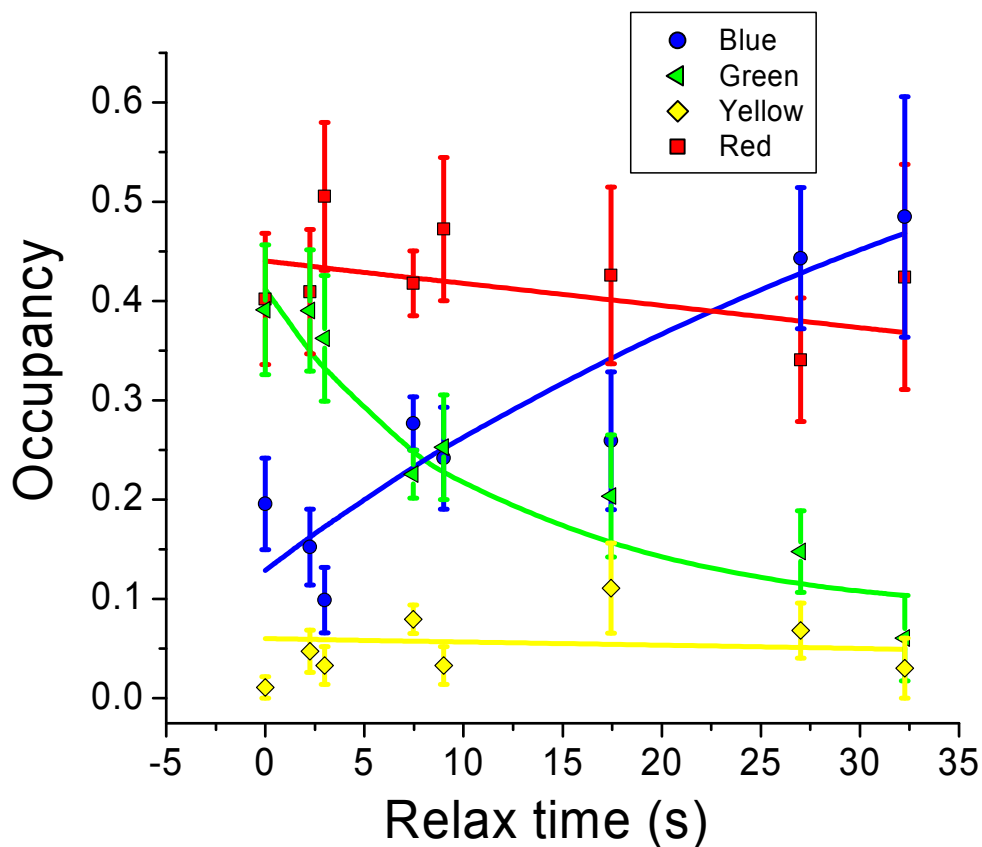
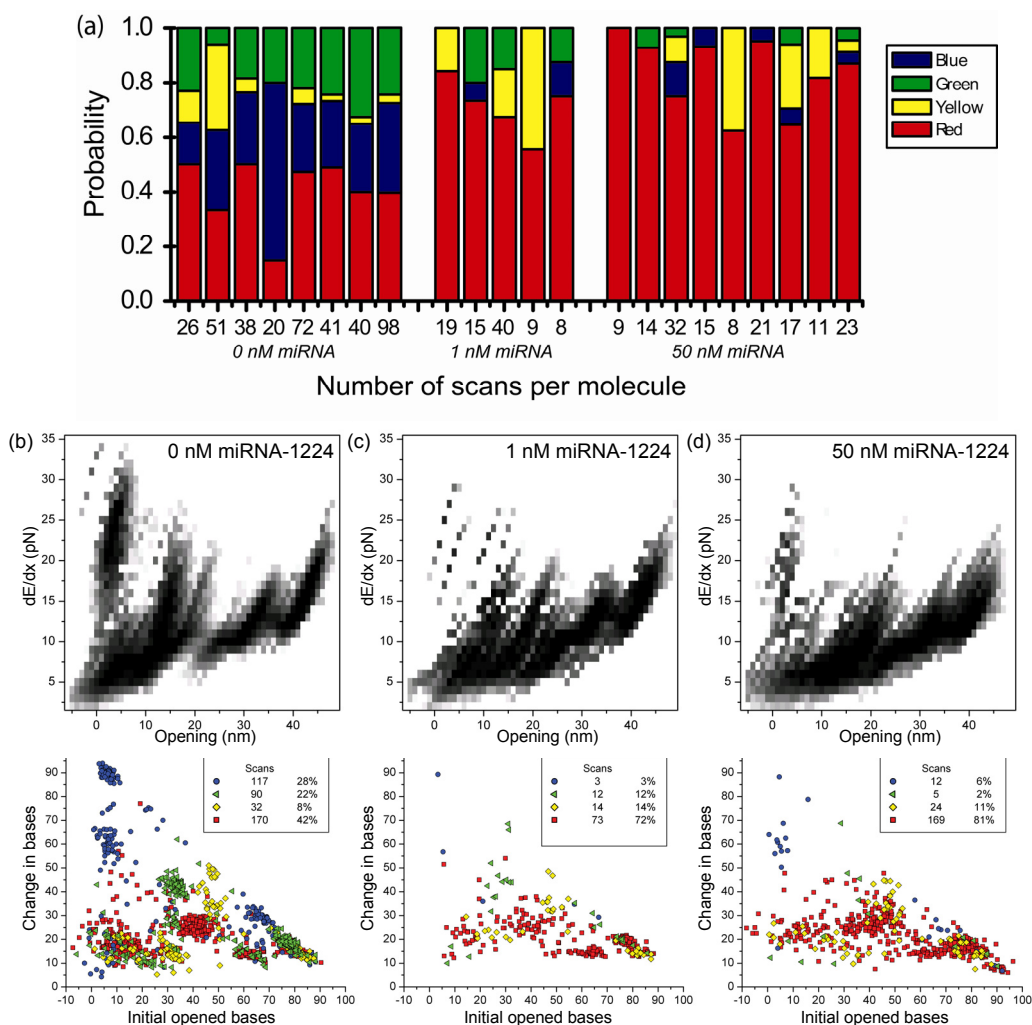


Figure 3.11 The CCR5 mRNA structure was relaxed at 0 force for the specified time and the distribution of the four conformations shown in Fig. 3.3 was measured. Data was fit with exponential decay functions as a visual guide.

### ***Role of miRNA-1224***

To investigate how miRNA may regulate the mRNA structure, we took data with and without miRNA-1224 (5'-GUGAGGACUCGGGAGGUGG-3'), which has been observed to increase  $-1$  PRF efficiency *in vivo*<sup>57</sup>. Overall the  $dE/dx$  graph with miRNA is less well defined, consistent with transient binding of the miRNA to the mRNA as it unfolds (Fig. 3.12).



**Figure 3.12 Results for different levels of miRNA-1224 (5'-GUGAGGACUCGGGAGGUGG-3').**

**(a)** Individual molecules were scanned repeatedly with 7.5 second refolding time and all molecules with at least 8 scans were represented as a bar distribution, indicating the percentage of each unfolding conformation observed for that individual molecule. The x-axis indicates the number of scans resulting for an individual molecule. The data in (a) shows 8 molecules without miRNA-1224 (left), 5 molecules with 1 nM miRNA (middle), and 9 molecules with 50 nM miRNA (right). Consistency between different molecules indicates that the different conformations were not a result of contamination or damaged mRNA. There is no correlation between the disruption pathways observed in successive pulls. **(b)**  $dE/dx$  density plot and distribution of substeps for data taken without miRNA, **(c)** with 1 nM miRNA, **(d)** with 50 nM miRNA. Adding miRNA caused the blue and green states to be suppressed.

The high-force conformations (blue and green) were significantly diminished in the presence of miRNA (Fig. 3.12) leaving a predominance of the low-force conformations (red and yellow). This demonstrates that the miRNA interacts directly with the mRNA structure and can alter its conformational ensemble when the mRNA folds in the presence of the miRNA. The miRNA has been demonstrated to promote -1 PRF<sup>57</sup> which suggests a correlation between the low-force conformations (red and yellow) and -1 PRF. This contrasts with previous research which has demonstrated a direct correlation between disruption force and -1 PRF<sup>51,69</sup>. However, hairpin loop structure<sup>70</sup>, complete blockage of translation by pseudoknots with high mechanical stability<sup>71</sup>, and geometric factors such as stem lengths<sup>72</sup> are all potentially important factors which can influence -1 PRF efficiency. An additional consideration is that the single-molecule experiments pull on both the 5' and 3' ends of the mRNA simultaneously while a translocating ribosome only directly interacts with the 5' end of the mRNA. Also, the miRNA experiments were performed after relaxing the molecule for only 7.5 seconds while *in vivo* experiments for frameshifting would effectively represent much longer equilibration time.

## **Methods**

Double-stranded DNA handles were generated by PCR, purified using a DNA purification kit (Qiagen), and then digested with BtgI and BstXI enzymes respectively. The handles were then run on a 2% agarose gel, excised, and gel purified (Qiagen) to extract 194 base-pair and 234 base-pair handles with 4-base sticky ends resulting from the digestion. The handles were then ligated (T4 Ligase New England Biolabs) to single-stranded DNA overlaps of length 51 bases and 85

bases respectively (synthesized by Invitrogen) using a 10:1 ratio of overlap to handle. This was subsequently gel purified as before. The resulting handles with ligated single-stranded DNA overlaps were typically measured to be about 40 ng/μl, though it was difficult to differentiate the 194 base-pair handle from the ligated construct with the extra 51 base overlap.

The RNA was transcribed by Belew<sup>57</sup> using an Ambion Megascript kit, DNase digested, phenol extracted, passed through a sephadex G-25 column, and ethanol precipitated. In its native context, this sequence directly follows the *UUUAAA* slippery site heptamer. The RNA included overlap regions on either end which matched the DNA single-stranded overlaps ligated to the handles. The central segment of the RNA is the CCR5 mRNA. The full RNA sequence is given below where red and blue indicate overlaps regions that will form hybrid DNA/RNA and black indicates the CCR5 mRNA sequence with total length 97 bases.

*GGGAAAUAUAUCAAAUCGUUCGUUGAGCGAGUUCUCAAAAAUGAACAAAUGUCGACUGU  
CGUCCAUGCUGUGUUUGCUUUAAGCCAGGACGGUCACCUUUGGGGUGGUGACAAGUG  
UGAUCACUUGGGUGGUGGUGUGUUUGCGUCUCUCCAGGAAUCAUUUACCAGAUCU  
CAAAAAGAAGGUCUUCAUUACACCGCAGCUCUCAUUUCCAUACAGUCAGUAUC*

The annealing procedure was as follows. 33% formamide, 1 mM EDTA, 20 mM TRIS pH 8.0, 533 mM NaCl, 50 ng/μl mRNA, 1/300 dilution RNasin® Plus RNase Inhibitor (2500u) (Promega), 1.3 ng/μl Handle 1 and Handle 2 with overlaps. The mixture was briefly vortexed and then annealed using the following protocol<sup>51</sup>: 10 minutes 85 °C, 90 minutes 62°C, 90 minutes 52°C, then cool to 4°C at a rate of 4°C/minute.

GODCAT<sup>37</sup> oxygen scavenger was prepared by mixing 20 mg glucose oxidase (Sigma) and 4 mg catalase (Sigma) with 200 μl of Tris buffer (10 mM Tris

pH 7.5, 1 mM EDTA, and 50 mM NaCl). The resulting mixture was vortexed and aliquoted for later use, stored at -20°C.

For sample chamber construction, cover slips were scrubbed with Windex, rinsed, and dried with an air hose. Polyclonal anti-digoxigenin (Roche) was diluted to 20 µg/ml in phosphate buffered saline and incubated in the chamber for 30 minutes. Blotting buffer (1 mg/ml blotting-grade blocker from Bio-Rad) in sodium phosphate buffer (50 mM sodium phosphate buffer pH 7.0, 50 mM NaCl, 10 mM EDTA, and 0.02% Tween 20)<sup>36</sup> was flowed into the chamber for 3 cycles of 20 minutes each, in order to prevent the beads from sticking to the surface. A mixture of 20 µl buffer (50 mM sodium phosphate buffer pH 7.0, 50 mM NaCl), 4 µl of the annealed RNA sample, and 1 µl of 1/10 dilution RNasin® (2500u) (Promega) diluted in 50 mM Tris pH 7.5 was flowed into the chamber for 45 minutes. 820 nm diameter polystyrene beads (Bangs Labs PC03N) were washed, coated with streptavidin, and diluted by 1/10 from the original concentration in sodium phosphate buffer (50 mM sodium phosphate buffer pH 7.0, 50 mM NaCl, 10 mM EDTA, and 0.02% Tween 20)<sup>36</sup>. 20 µl of diluted beads were mixed with 1 µl of 1/10 dilution RNasin® and flowed into the chamber for 20 minutes. For the final buffer conditions, Tris-Mg buffer (20 mM TRIS pH 8.0, 50 mM NaCl, 20 mM MgCl<sub>2</sub> (except Fig. 3.6 which varied Mg), 0.02% Tween-20) was combined with GODCAT<sup>37</sup> oxygen scavenger described above for a final concentration of 721 µg/ml glucose oxidase, 144 µg/ml catalase, and 3.9 mg/ml glucose. We prepared 550 µl of the final buffer which was then syringe filtered (VWR 0.2 µm polyethersulfone membrane). The filtering results in approximately 400 µl of solution, to which we added 2 µl of 1/10 dilution RNasin® and then

vortexed briefly. The final solution was flowed into the chamber to replace all prior buffer. The chamber was subsequently sealed with nail polish.

The experimental apparatus has been described in Chapter 2<sup>73</sup>. Standard methods were used to calibrate the optical trap<sup>41</sup>. The sample was repeatedly stretched by lowering the stage at  $\sim 77$  nm/s and then relaxing the sample at zero force for 7.5 seconds between each scan (except Fig. 3.11 in which the relaxation time was varied). Data was collected at 60 kHz and filtered to 20 kHz for analysis.

### ***Details of Worm-Like Chain Theory***

Worm-like chain theory<sup>38</sup> was used to model the properties of the double-stranded DNA, hybrid DNA/RNA, and single-stranded mRNA.

**Double-stranded DNA properties:** contour length of 0.338 nm per base<sup>39</sup>, stretch modulus 1205 pN, persistence length 43.1 nm, 15 nm contour length added to estimate the length of the attachments<sup>40</sup>, 436 total bases of double-stranded DNA in handles.

**Hybrid DNA/RNA properties:** contour length of 0.26 nm per base, persistence length 12 nm, enthalpic contributions neglected<sup>29</sup>, 136 total bases of hybrid DNA/RNA in handles.

**Single-stranded mRNA properties:** contour length of 0.59 nm per base, persistence length 1 nm<sup>26</sup>, neglected enthalpic contributions<sup>26</sup>, 97 total bases of single-stranded mRNA when the structure is unfolded.

## **Conclusion**

The conclusion of our study is that the CCR5 mRNA occupies an ensemble of relatively stable conformations and the conformation (blue) with the highest disruption force is consistent with a predicted pseudoknot structure. Upon refolding the mRNA can become trapped in an intermediate conformation (green) for a time of order 30 seconds but the high-force conformation (blue) dominates after long equilibration time. We have also found that miRNA-1224, which enhances  $-1$  PRF *in vivo*<sup>57</sup>, decreases occupancy of the high-force conformations (blue and green) and increases occupancy of the low-force conformations (red and yellow). This suggests that the low-force conformations play a role in  $-1$  PRF or that the mechanism of  $-1$  PRF enhancement by miRNA-1224 involves a specific interaction between the ribosome and the mRNA/miRNA complex, rather than simple steric hinderance.



## Chapter 4: Theoretical Modeling of Energy Landscapes<sup>d</sup>

### *Introduction*

Discrepancies between the model and the data for results shown in chapter 2 motivated development of an improved theory by La Porta<sup>d</sup>, and confirmed by experiments in our lab. This new model became a foundation for some of the work that follows and opened up many new research possibilities. The model calculates the theoretical energy of a biopolymer as a function of end-to-end extension. It also includes corrections which are dependent on the hydrogen bond energy.

### *Methods*

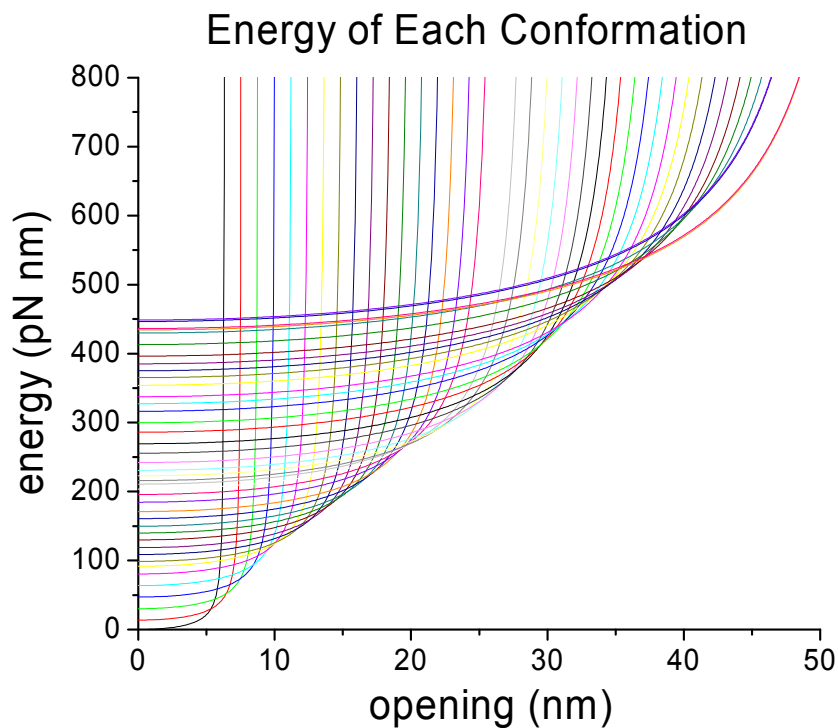
As an example, we will consider a DNA hairpin, though the principles described here would apply equally to RNA or any other biopolymer where the possible folding conformations are well-defined. The example sequence (Fig. 4.1) is taken from one of the samples to be discussed further in Chapter 5. As a first step, we obtain the binding energy for each base-pair in the hairpin duplex stack from a structure prediction program called Mfold<sup>44</sup> which is freely available online. Details of the parameters used are indicated in chapter 2. The resulting energy as a function of number of base-pairs released is plotted in Fig. 4.2.

---

<sup>d</sup>Adapted from La Porta, A. & de Messieres, M. Calculation of the energy of a biopolymer as a function of end-to-end extension. *In preparation* (2012).



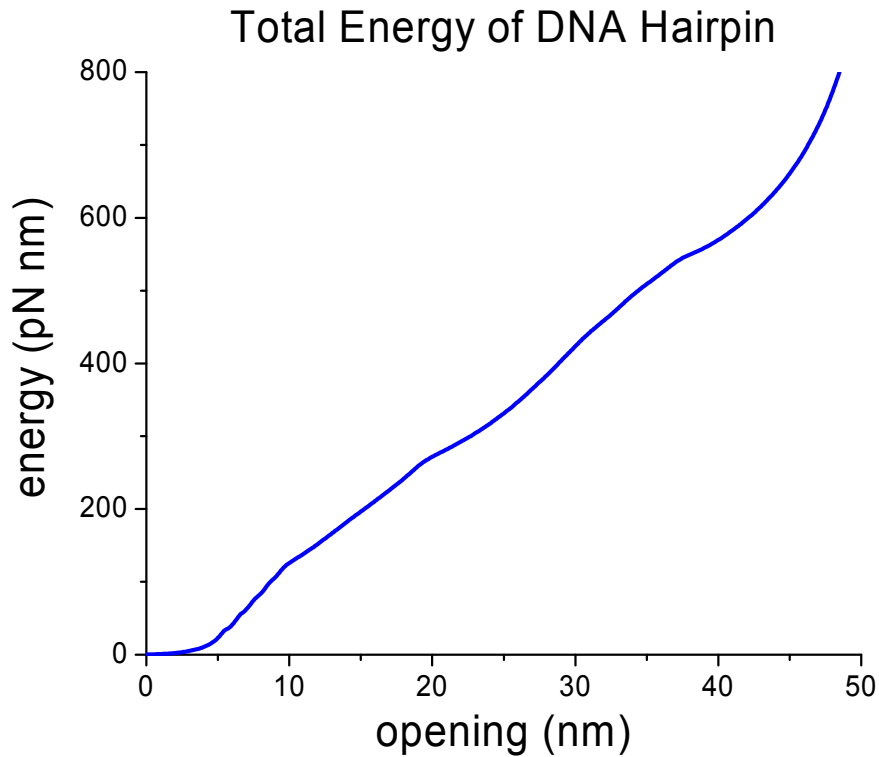
energy barrier for the number of base-pairs released, as shown in Fig. 4.2. However, a conformation with more bases released can be stretched to a longer distance more easily. The balance between the base-pair free energies and the stretching energy of the released single-stranded DNA determines which conformations will be present for each opening distance. The resulting plots of energy as a function of opening distance for each conformation are shown in Fig. 4.3. Each stretch curve, except for the final curve when all DNA is single-stranded, includes 1.85 nm of contour length to account for the helix width along with a correction distance (free fitting parameter) to account for curvature of single-stranded DNA extending from the helix<sup>12</sup>.



**Figure 4.3** Energy of each possible conformation corresponding to Fig. 4.2 as a function of opening distance  $x$ .

We may now calculate the total energy as a function of opening distance. For each opening distance  $x$ , we take the weighted sum of all conformations according to the Boltzmann distribution (Eq. 4.1). In general, the conformation with the lowest energy for each opening distance dominates. Note that some conformations dominate over a larger range of opening than others, while some conformations are never the lowest energy for any opening size. The final model for the DNA hairpin is shown in Fig. 4.4. Other details such as the convolution from attachment to DNA handles have been described previously in chapter 2.

$$\langle E \rangle = \frac{1}{Z} \sum_n E_n e^{-\frac{E_n}{k_B T}} \quad 4.1$$



**Figure 4.4** Energy for the DNA hairpin as a function of end-to-end distance, generated by weighting the energies of all conformations shown in Fig. 4.3 with the Boltzmann distribution for each opening size (Eq. 4.1).

## ***Hydrogen Bond Correction***

Here we consider a subtle correction which translates the energy parameters extracted from a program such as Mfold<sup>44</sup> into the proper format to describe disruption with the optical trap. Mfold uses a nearest-neighbor model<sup>34,43,74</sup> for base-pair binding where each possible block of 2x2 bases (GA/CT for example) is defined by a single energy parameter. Note that under the assumptions of the model, rotational symmetry gives GA/CT = TC/AG, but the directionality of DNA 5' to 3' ordering means TA/AT  $\neq$  AT/TA. There are 10 independent parameters which we measure in chapter 6 along with further discussion.

As described previously<sup>75</sup>, each parameter only includes  $\frac{1}{2}$  of the relevant base-pair binding energies so that the sum of all the energies is correct. Consider as an example the sequence GAC/CTG which has two blocks, GA/CT followed by AC/TG. The binding energy for the middle base-pair A/T is divided evenly for the two energy parameters representing those two blocks. However, during an optical trap experiment, the bases are disrupted as pairs, not 2x2 blocks. G/C pairs have 3 hydrogen bonds and A/T pairs have 2 hydrogen bonds. The nearest-neighbor model effectively smooths out the influence of the hydrogen bond whenever a 2-bond pair is followed by a 3-bond pair or vice versa. As a first approximation, we may add this correction by simply adding or subtracting  $\frac{1}{2}$  the hydrogen bond energy where appropriate. For example, Mfold gives us the energy for the block GA/CT which includes  $\frac{1}{2}$  of a G/C pair equivalent to  $\frac{3}{2} = 1.5$  hydrogen bonds and  $\frac{1}{2}$  of an A/T pair equivalent to  $\frac{2}{2} = 1$  hydrogen bonds. The total for the block is 2.5 hydrogen bonds. However, from the optical trap measurement, we should consider disruption of the

G/C bond to be a full 3 hydrogen bonds and therefore add a correction of  $\frac{1}{2}$  of a hydrogen bond. Likewise, other blocks may require subtraction of  $\frac{1}{2}$  of a hydrogen bond. Some blocks, TA/AT for example, require no correction. Previous studies considered the effect of duplex RNA terminal A-U base-pairs and the energy of a single hydrogen bond was determined to be approximately 0.9 kcal/mol in bulk experiments<sup>75</sup>.

Figure 4.5 shows the new model with and without a hydrogen bond correction, plotted as a function of  $dE/dx$  which highlights the subtle differences. Also shown are data for three different molecules compared to a model with hydrogen correction set to 1.3 kcal/mol (black) and no hydrogen bond correction (gray). Individual molecules were found to fit between 0.4 kcal/mol and 1.7 kcal/mol. A single molecule type is insufficient to establish whether the hydrogen bond correction is justified, but the general trend for this sequence and others indicates it may be a valuable addition. The hydrogen bond correction will be explored further in chapter 6.

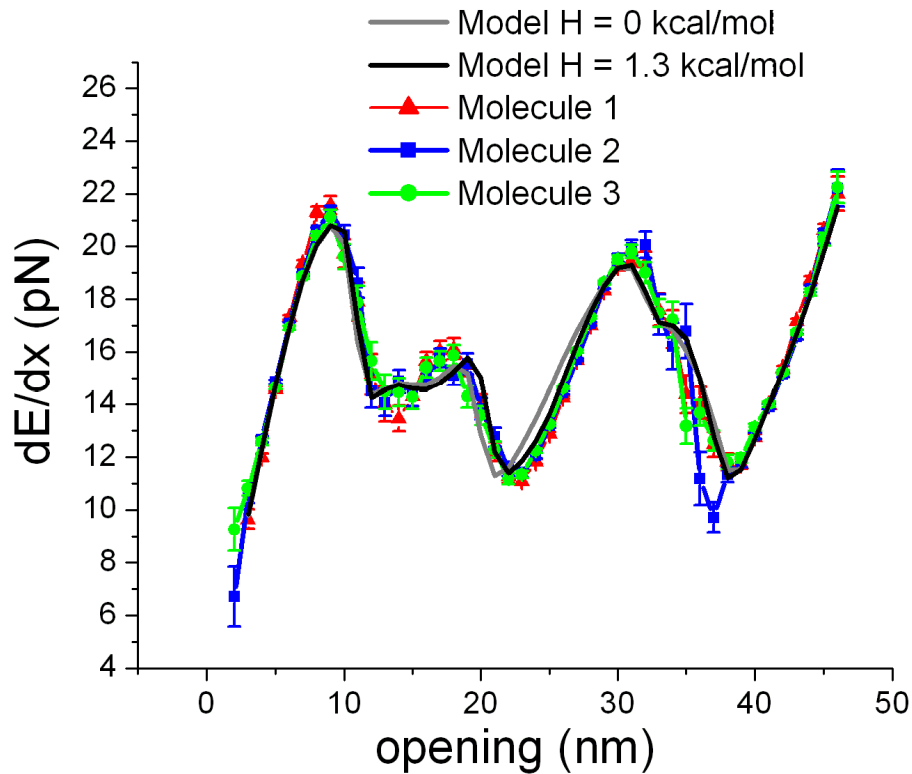


Figure 4.5 The theoretical model for the TA/AT control hairpin (Fig. 4.1) with a hydrogen bond correction of 1.3 kcal/mol (black) and no hydrogen bond correction (gray). Data collected for the TA/AT hairpin (red, blue, green) fits better with a hydrogen bond correction in the range of 0.4 to 1.7 kcal/mol.

### ***Loop Initiation***

The one free parameter of the model is a correction to the total contour length of the released DNA which results from the curvature of the single-stranded DNA extending from the duplex helix of the hairpin. When the loop is released, the DNA becomes completely single-stranded and then the curvature correction becomes 0. The effective free energy of initiation for folding of the loop is extremely sensitive to this correction distance. Since the correction is a free fitting parameter, it is therefore

difficult to accurately determine the free energy of loop initiation in a single-molecule stretching experiment.

## ***Conclusion***

The new model is derived from first-principles and resolves some discrepancies between a previously used model<sup>11,12,73</sup> and our experimental data. We will demonstrate in subsequent chapters that the combination of this model with the measurement techniques developed in chapter 2 results in precise agreement between data and theory. We will apply these methods to a system of biological interest, the GA/AG mismatch, and also conduct single-molecule measurements to determine the 10 nearest-neighbor energy parameters for duplex DNA.



## Chapter 5: Tandem GA Mismatch Energy<sup>e</sup>

### ***Introduction***

In chapter 2 we developed a new method to precisely measure the energy landscape of a biomolecule. In chapter 3 we studied CCR5 mRNA structure and developed the idea that looking at  $dE/dx$  rather than the energy landscape itself was a useful way to characterize single-molecule data. In chapter 4 we considered a new model which provides precise agreement between data and theory. Here we will combine all three of these developments to study a system of biological interest, the GA/AG mismatch.

### ***Background***

The DNA mismatch repair system (MMR) is a critical pathway by which cells maintain the integrity of the genome<sup>76</sup>. Dysfunction in the MMR mechanisms can induce cancer, so increased understanding of the MMR may provide drug targets for treatment of disease. G-A mismatches are of particular interest because they are often not efficiently recognized by the MMR<sup>77,78</sup>.

We apply our single-molecule experimental and analysis methods<sup>73</sup> to directly measure a GA/AG<sup>79</sup> tandem mismatch. We find significantly higher binding energy than would be expected for non-interacting bases, but less than Watson-Crick base-pairing as might be expected. Bulk melting experiments<sup>80</sup> often use data from multiple sequences to extract DNA binding energies. In contrast, our approach is direct and requires only a single sequence. By measuring the GA/AG mismatch

---

<sup>e</sup>Adapted from de Messieres, M., Chang, J.-C. & La Porta, A. Direct single-molecule measurement of GA mismatch energy. *In preparation* (2012).

energy, we demonstrate a practical application of our broadly applicable energy landscape reconstruction method<sup>73</sup>.

## ***Methods***

A hairpin sequence containing a GA/AG mismatch (Fig. 5.1a) was attached through double-stranded DNA handles to both the cover slip surface and a microsphere held by a single-beam optical trap. The hairpin was unfolded by moving the cover slip away from the optical trap while force and opening distance were recorded. We also took three control measurements where the GA/AG mismatch was replaced by TA/AT, GA/CT, or GC/CG.

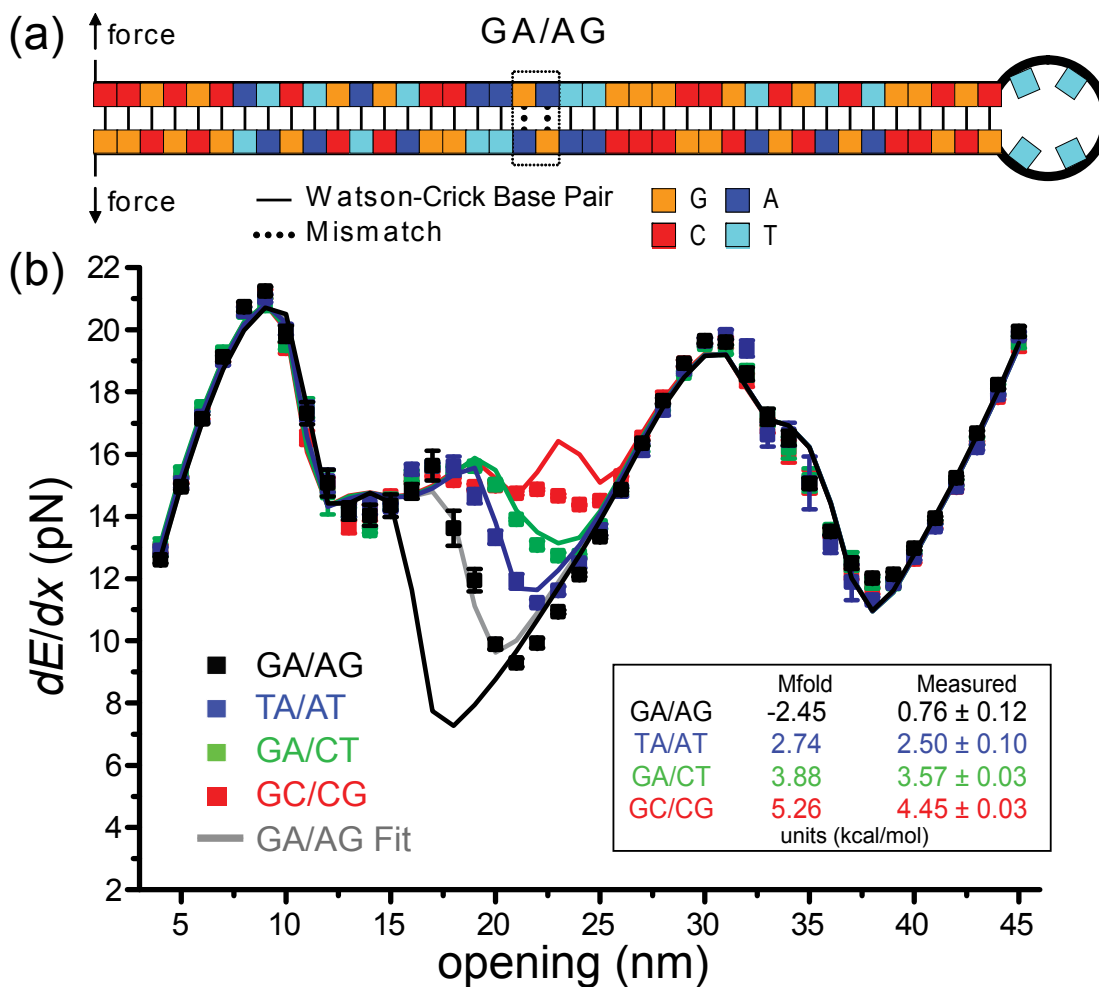


Figure 5.1 (a) Four hairpins varied only in two base-pairs, where GA/AG was replaced by TA/AT, GA/CT, and GC/CG for control measurements. (b)  $dE/dx$  plot shows data points (symbols) and theory (solid lines) for GA/AG treated as an unpaired loop (black), TA/AT (blue), GA/CT (green), and GC/CG (red). The discrepancy between the theory for the GA/AG mismatch (black line) and data points (black squares) is not unexpected since Mfold treats the GA/AG mismatch as an open non-interacting loop. Gray line: Resulting fit for GA/AG. The fits for the three controls are not shown. Inset: The Mfold prediction and measured values for the sum of energy parameter indices 18, 19, and 20 in the hairpin sequence. Uncertainties are the standard deviation of the mean from fits to single scans.

## **Analysis**

The change in energy per opening distance ( $dE/dx$ ) is calculated using techniques developed previously<sup>73</sup> and compared to a theoretical model described in chapter 4.  $dE/dx$  represents the force required to open the molecule an additional unit of distance. The weighted-mean  $dE/dx$  plots from multiple molecules for each hairpin type are shown in Fig. 5.1. The three control sequences without mismatches give good agreement with theoretical models (red, green, and blue solid lines) derived from Mfold parameters<sup>44</sup>. The GC/CG control (red) showed some deviation from the theoretical prediction and further study may lead to new models to account for these discrepancies. In comparison to theory which treats the GA/AG mismatch as an open loop (black line), we measured significantly higher binding energy (black squares), consistent with previous bulk experiments<sup>78</sup>.

To quantify the  $dE/dx$  plots, we used a simplex method to fit the three energy parameters which are dependent on the two base-pairs we modified (indices 18, 19, 20) and compare the sum to Mfold (Fig. 5.1b inset). From symmetry, indices 18 and 20 are equal, except for the GA/CT control. The determination of individual energy parameters is limited by the  $\sim 1.1$  nm uncertainty in position measurements. However, the data establishes sensitivity to single base-pair changes in the sequence, as demonstrated by the progression between the three control hairpins in Fig. 5.1.

## **Conclusion**

Direct measurement using single-molecule techniques demonstrates that the GA/AG mismatch is significantly more stable than would be expected from a model where the bases are assumed to be non-interacting. Extensions of this work can

consider subtle effects such as nearest-neighbor assumptions in currently used models<sup>80</sup> or salt and pH dependency. Bulk experiments cannot differentiate variations in a sample and may average in other sources of error. In contrast, the single-molecule techniques presented here provide direct and less ambiguous access to the relevant energies.

## Chapter 6: Nearest-Neighbor Free Energy Parameters for Duplex DNA<sup>f</sup>

### *Introduction*

In chapter 4 we discussed a new model which allows us to precisely predict the theoretically expected energy landscape for an optical trap experiment. The model depends on 10 nearest-neighbor free energy values<sup>42,43</sup> which we obtained from the structure prediction program Mfold<sup>44</sup>. We will now invert the process and fit our data to determine those parameters based on a data set for hairpin sequences designed to cover all possible configurations. This single-molecule technique can explore subtle features of base-pair binding which cannot be measured in bulk experiments. This research is currently in progress.

### *Background*

Secondary-structure prediction software packages can predict the folding conformation for DNA and RNA sequences and provide a suite of useful analysis tools. The popular Mfold<sup>44</sup> program is one example, with thousands of citations across a range of disciplines. These programs require a library of binding energies for the different possible combinations of base-pairs. Based on the nearest-neighbor (NN) model<sup>43,74,81,82</sup>, the library contains all possible blocks of 2x2 base-pairs. For Watson-Crick base-pairing of DNA we have 10 possible combinations: GC/CG, CG/GC, GG/CC, CA/GT, GT/CA, GA/CT, CT/GA, AA/TT, AT/TA, and TA/AT<sup>43</sup>. These numbers have previously been established using bulk methods<sup>43</sup> and also

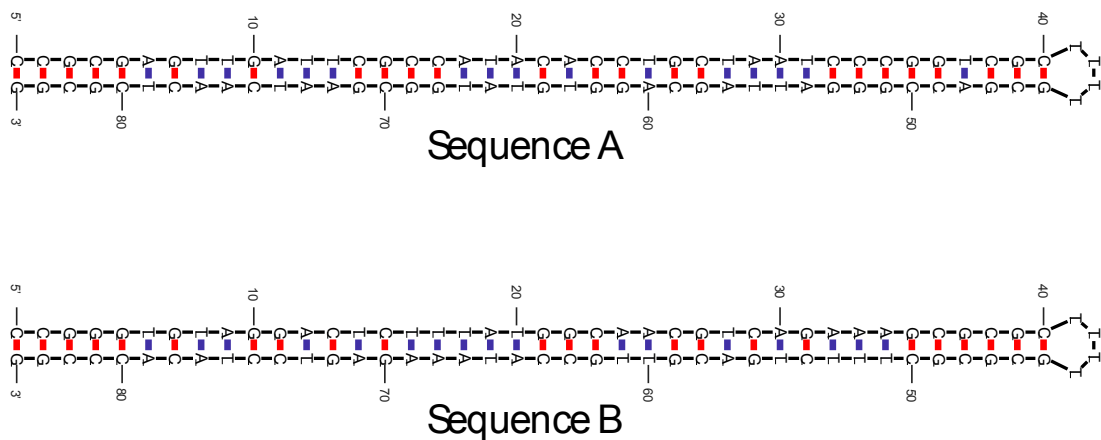
---

<sup>f</sup>Adapted from de Messieres, M., Chang, J.-C. & La Porta, A. Single-molecule measurement of 10 nearest-neighbor energy parameters for duplex DNA. *In preparation* (2012).

measured by unzipping long duplex DNA of length 2.2kb to 6.8kb<sup>34</sup>. Here we demonstrate that single-molecule methods can directly measure these energy parameters with sufficient resolution to distinguish single-base mutations and other sources of error.

## **Methods**

The optical trap setup and general methods of data collection and analysis have been described previously in chapter 2 and chapter 4. We used a random search algorithm to generate sequences of 66 bases which contain exactly one of each of the 64 possible triplets of DNA. For example, a 4-base sequence GCCA contains two triplets, GCC and CCA. The purpose of this method is to obtain DNA hairpins with even exposure to all possible sequence patterns. Of many possibilities, we chose a sequence that had a fairly even energy landscape profile as predicted by Mfold<sup>44</sup>. Excessively sharp peaks and valleys in the energy landscape would create regions of low and high statistical convergence respectively. This is not necessarily problematic but would result in some regions of the landscape having more weight in the final fit than others. We split the selected sequence into two parts and inserted each into a hairpin with flanking bases that we expected would give the hairpin a well-defined stem base and loop structure. Splitting the sequence was necessary due to the approximately 100-base limit for synthesis (Invitrogen), with PAGE purification. We will refer to the two sequences as *Sequence A* and *Sequence B*, which are shown in Fig. 6.1. We found that ~11% of *Sequence A* and ~14% of *Sequence B* appear self consistent and also have the highest measured  $dE/dx$  values. Other samples show sharp drops in  $dE/dx$ , consistent with the assumption that they represent mutations.



**Figure 6.1** Structures for the two hairpins used in this study, generated by Mfold<sup>44</sup>.

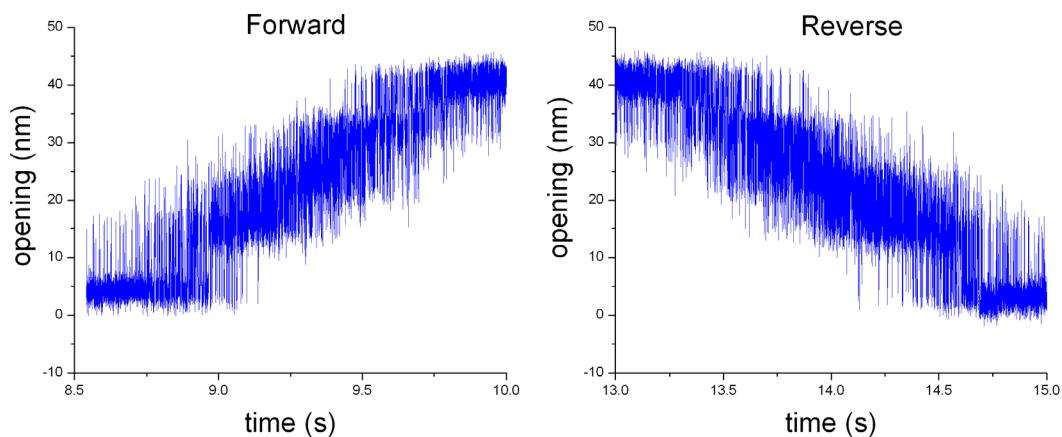
*Sequence A:*

**CCGCGAGTTGATTGCCATACACCTGCTAATCCCGGTCGCTTTTGCAGCCGGGATTAGCAGGTG  
TATGGCGAATCAACTCGCGG**

*Sequence B:*

**CCGGGTGTAGGACTCTTTATGGCAACGTCAGAAAGCGCGCTTTTGCAGCGCTTTCTGACGTTGCC  
ATAAAGAGTCCTACACCCGG**

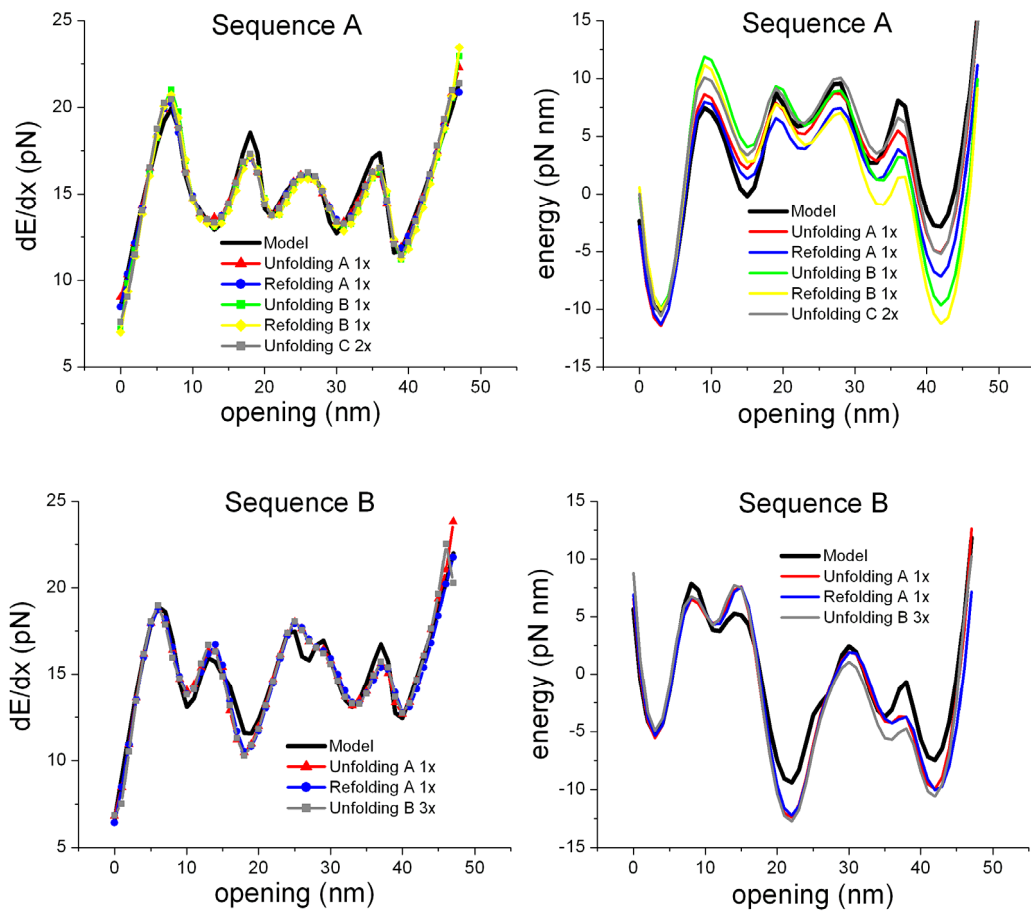
## Reversibility



**Figure 6.2** Example forward and reverse scan for a single molecule of type *Sequence A*.



To explore the possibility that our results were dependent on irreversible effects, we analyzed data from both unfolding and folding pathways. An example of unfolding followed by folding for one molecule of *Sequence A* is shown in Fig. 6.2. We plot  $dE/dx$  results for *Sequence A* (Fig. 6.3a) along with the corresponding energy plot (Fig. 6.3b) for two different molecules in the unfolding (stage speed 23 nm/s) and refolding (16 nm/s) directions, along with a 3<sup>rd</sup> set of data for unfolding at twice the speed (42 nm/s). We fit the stiffness calibration for different molecules to match the model for comparison. A high degree of precision is apparent from the consistency between unfolding and refolding, between different molecules, and between different trap movement rates. This sequence is also in excellent agreement with a theoretical model, shown in black. We also tested *Sequence B* and plot  $dE/dx$  (Fig. 6.3c) and the energy plots (Fig. 6.3d) for unfolding, refolding, and unfolding of a different set of molecules at 3x the original speed (76 nm/s). We have dropped the error bars for clarity.



**Figure 6.3** (a,b)  $dE/dx$  results and the corresponding landscape for one molecule of *Sequence A* for the unfolding ( $\sim 23$  nm/s) and refolding ( $\sim 16$  nm/s) directions (red and blue), a 2<sup>nd</sup> molecule at the same speeds for unfolding and refolding (green and yellow), and a 3<sup>rd</sup> data set (gray) from several molecules unfolded at approximately twice the speed (42 nm/s). (c,d)  $dE/dx$  results and the corresponding landscape for one molecule of *Sequence B* for the unfolding (23 nm/s) and refolding (16 nm/s) directions (red and blue) and a 2<sup>nd</sup> data set (gray) from several different molecules unfolded at approximately three times the speed (76 nm/s). Error bars not shown for clarity.

## ***Fitting Method and Results***

We found that fitting only one of the two sequences gave ambiguous results. However, if we fit one trace of *Sequence A* simultaneously with one trace of *Sequence B*, we obtained good agreement with the Mfold library. This is expected since neither sequence was designed to have full coverage of the possible combinations. To determine statistical uncertainty, we used the first three traces from each of three molecules giving 9 total traces for each of the two sequence types. We then ran our fit for all possible pair combinations ( $9 \times 9 = 81$  combinations) of *Sequence A* (unfolding speed 42 nm/s) and *Sequence B* (unfolding speed 76 nm/s), and reported the standard deviation between different fits as shown in Fig. 6.4 red. We also tested a different molecule for each type at slower unfolding speed (23 nm/s) and slower refolding speed (16 nm/s) shown as blue and black respectively. Fits for the hydrogen bond are reasonable and similar to a previously reported value of 0.9 kcal/mol for RNA melting due to Xia<sup>75</sup>. For comparison, the TA control described in chapter 4 gave fits for the hydrogen bond correction in the range of 0.4-1.7 kcal/mol. The Xia value was obtained as a correction to terminal RNA A-U pairs so it is not directly comparable to internal DNA hydrogen bonds. However, the magnitude of the hydrogen bond is compelling and further experiments across a wider range of sequences may determine the validity of this model more precisely.

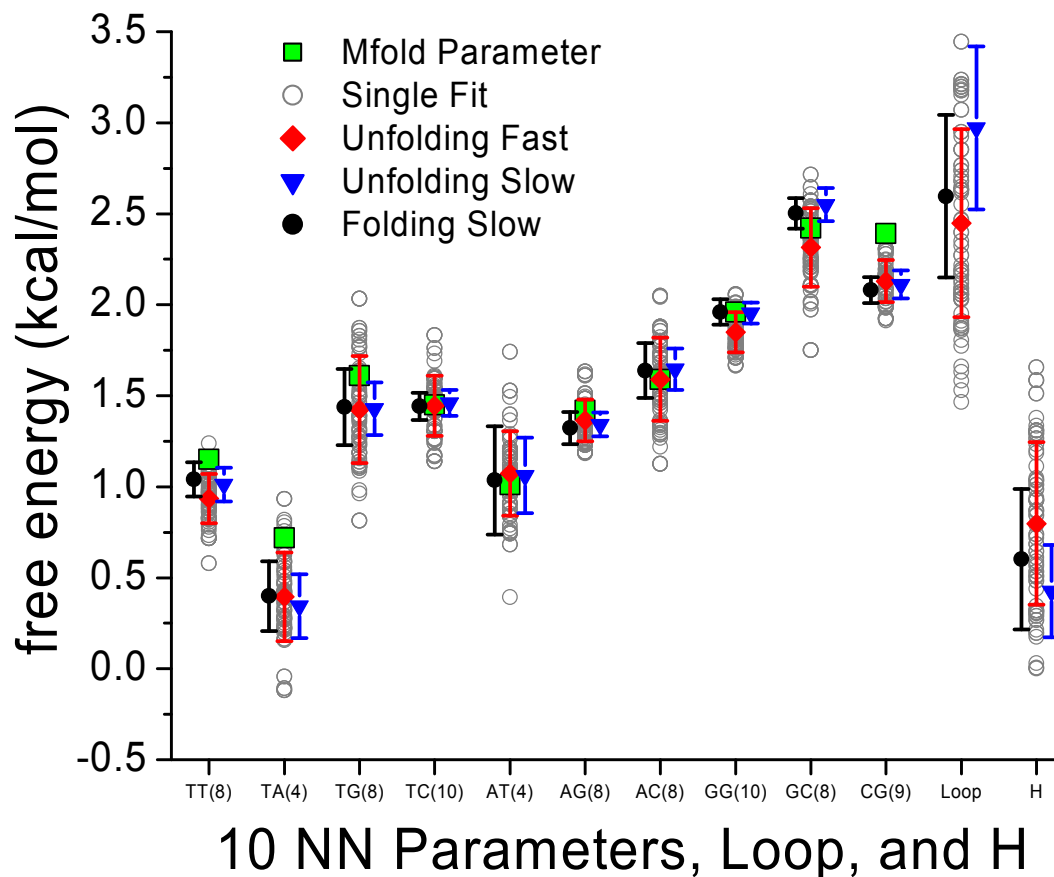


Figure 6.4 Green shows Mfold NN free energy parameters obtained at 24°C and 140 mM Na<sup>+</sup>. Each fit (gray) represents a global fit for one scan of *Sequence A* and one scan of *Sequence B*. There were 3 scans for each of 3 molecules for both types, giving 9x9 = 81 combinations. Red indicates the mean and standard deviation for these fits. We also tested one molecule of each type at slower speed in the unfolding (blue) and unfolding (black) directions.

In general we found excellent agreement between the Mfold values (Fig. 6.4 green) and our measured fit values (Fig. 6.4 red, blue, or black). A global error in trap calibration stiffness likely accounts for some of the differences. The loop parameter is not expected to be precise for reasons discussed in chapter 4. We have

also found that the Mfold values do not correctly predict all sequences suggesting that in some instances, a more detailed model is necessary. For example, *Sequence B* has a consistent error at approximately 30 nm which is apparently not measurement error. Other sequences we have studied (see the control GC/CG in chapter 5) also show consistent deviation from the Mfold model. These may be caused by effects beyond the nearest-neighbor assumption which are not accounted for by the current model. This illustrates a crucial advantage of the single-molecule techniques, because these discrepancies would potentially pass unnoticed for bulk experiments.

## **Conclusion**

In conclusion, we have demonstrated that single-molecule techniques can precisely measure the 10 NN parameters. We find that the hydrogen bond is a potentially valuable addition to the NN model, as previously suggested from bulk experiments<sup>75</sup>. These are the first single-molecule measurements with near base resolution of this parameter set. This research is currently in progress and will explore other models, such as a 16 NN parameter model (with all 16 possible 2x2 blocks) and a 14 NN model (where TT/AA = AA/TT and GG/CC = CC/GG). The single-molecule techniques presented here provide access to subtleties which are potentially lost when using bulk techniques.

## Chapter 7: G-Quadruplex Dynamic Force Spectroscopy<sup>g</sup>

### ***Introduction***

We now consider an alternative method of reconstructing energy landscapes through Dynamic Force Spectroscopy (DFS). For some molecules, especially those with short barrier distances, it may not be possible to reconstruct the landscape using the method described in chapter 2. DFS allows us to measure the barrier distance and the barrier height, though we obtain only coarse-grained results. We used our optical trap to disrupt single molecules of single-stranded DNA G4-quadruplex and analyzed the results using DFS to infer the nature of the transition state barrier for unfolding of the structure.

### ***Background***

Guanine-rich sequences in nucleic acids can fold into G-quadruplexes, in which four guanines on a single strand combine to form G-tetrad planes stabilized by metallic ions. G-quadruplex conformations exist throughout the genome, often in pre-transcribed regions of genes<sup>83-88</sup>. Recent research provides evidence that G-quadruplex forms *in vivo*<sup>89</sup>, has important regulatory functions<sup>90</sup>, and could serve as a drug target<sup>91</sup>. G-quadruplex formation in the human telomeric repeat sequence is of particular interest because it can inhibit telomerase, over expression of which is associated with cancers<sup>92</sup>. G-quadruplex occurs in different forms including RNA, single-stranded DNA, and double-stranded DNA accompanied by the corresponding

---

<sup>g</sup>Adapted from de Messieres, M., Brawn-Cinani, B., Chang, J.-C. & La Porta, A. Single-molecule study of G-quadruplex disruption using dynamic force spectroscopy. *Under Review* (2012).

C-rich i-motif structure<sup>93,94</sup>. In contrast to more typical stem-loop structures which can be disrupted progressively, G-quadruplex structures disrupt irreversibly when subject to a sufficiently large external force<sup>95</sup>.

We used optical tweezers to study a single-stranded DNA G4-quadruplex from the insulin-linked polymorphism region (ILPR), which is located upstream of the insulin gene's promoter. Many variants have been identified in this region but the GGGG segments tend to be conserved, suggesting quadruplex formation plays an important regulatory role for the insulin gene<sup>96</sup>. G4-quadruplex molecules were subject to a force which increased linearly in time until disruption was observed. Dynamic force spectroscopy (DFS) was used to characterize the transition state energy barrier between the folded and unfolded conformation in terms of the distribution of disruption forces measured. Two populations of disruptions were resolved in the data, which have been identified in prior work as corresponding to parallel and antiparallel conformations<sup>95</sup>. The distance and height of the principal energy barriers were extracted for the two conformations and the energy barrier was found to be close to the folded conformation, resulting in a high disruption force due to the steepness of the barrier, despite the fact that the barrier is relatively low.

### ***Theoretical Models***

Several theoretical models have been put forward to predict the distribution of rupture forces that would be observed when a system with a single transition state barrier is subject to a force that increases linearly with time. A model proposed by Evans<sup>23,24</sup> has two free parameters, the barrier distance  $d$  and the zero-force unfolding rate  $k_0$ . More general models which include the zero-force transition state barrier

energy  $G$ <sup>97,98</sup> were subsequently refined by Dudko et al<sup>99</sup>. In the latter model the probability  $p(F)$  that the structure ruptures at force  $F$  is given by

$$p(F) = \frac{k(F)}{r} \exp \left[ \frac{k_o k_B T}{r d} - \frac{k(F) k_B T}{r d} \left( 1 - \frac{\nu F d}{G} \right)^{1-\frac{1}{\nu}} \right]$$

$$k(F) = k_o \left( 1 - \frac{\nu F d}{G} \right)^{\frac{1}{\nu}-1} \exp \left[ \frac{G}{k_B T} \left( 1 - \left[ 1 - \frac{\nu F d}{G} \right]^{\frac{1}{\nu}} \right) \right] \quad 7.1$$

where  $k_B$  is the Boltzmann constant,  $T$  is the temperature,  $r$  is the rate at which the force is increased, and  $\nu$  is a dimensionless constant which parameterizes the shape of the energy barrier. Fig. 7.1 represents the relationship between the assumed shape of the energy barrier and the value of the parameter  $\nu$ . In the limit  $\nu \rightarrow 1$ , Eq. 1 reduces to the Evans model<sup>23</sup>. The Dudko model has several advantages over the Evans model. It takes into account the variation of the transition state distance with applied force and it determines the transition state energy  $G$ . However, the presence of an additional fitting parameter can make the analysis more sensitive to statistical fluctuations or measurement errors.



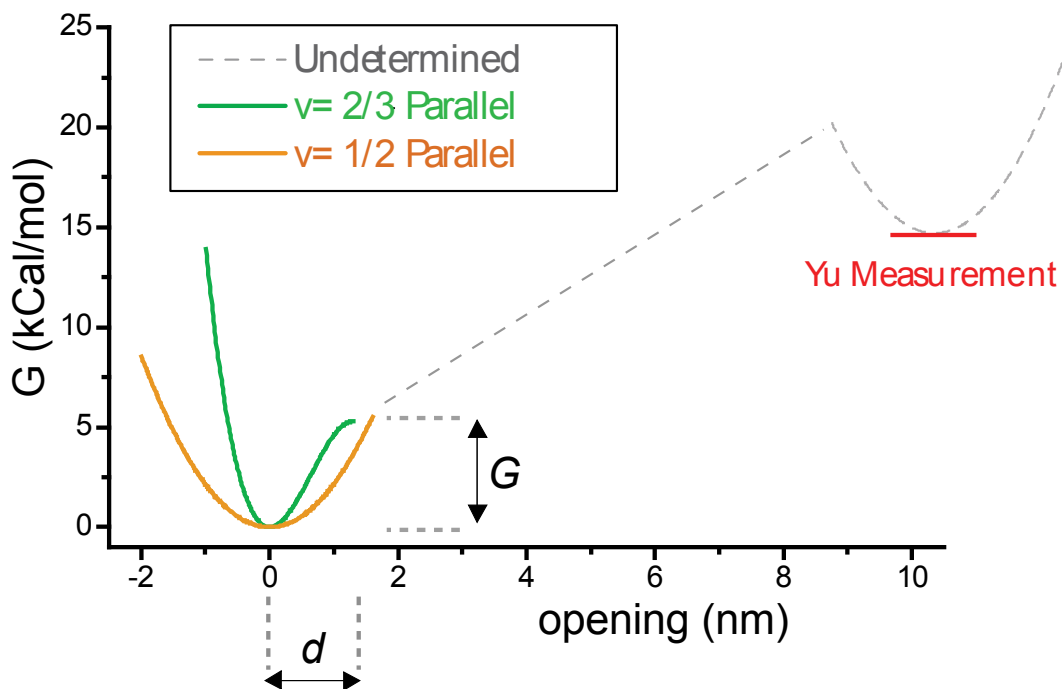
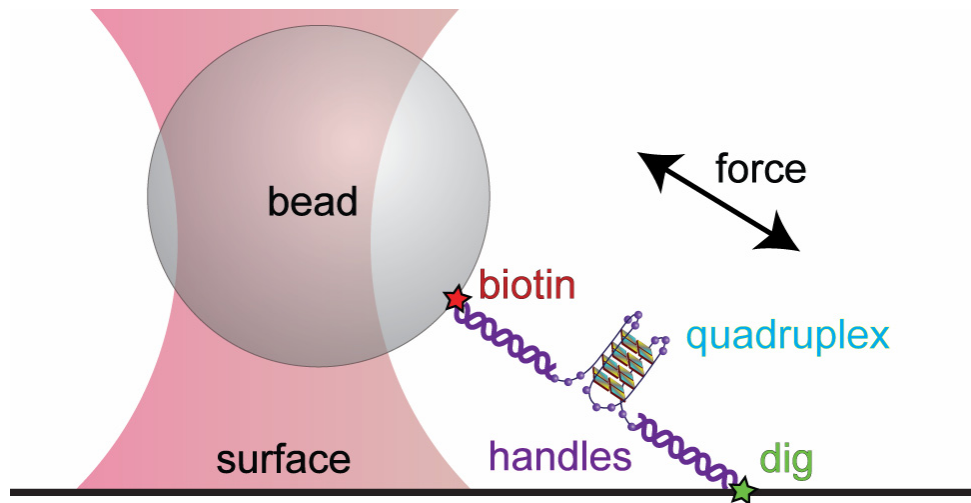


Figure 7.1 The Dudko models for  $\nu = 2/3$  (green) and  $\nu = 1/2$  (yellow) using parameters for  $d$  and  $G$  from fitting results in Table 7.1 for the parallel configuration. Dashed gray line indicates the undetermined region of the full energy landscape. DFS determines the primary barrier (yellow or green) but also constrains the rest of the barrier to have a slope less than that of the primary barrier. Red shows the total free energy to unfold the parallel conformation, determined by Yu<sup>95</sup> (14.5 kcal/mol from CD melting or 14 kcal/mol using the Jarzynski equality<sup>100</sup>). For the antiparallel conformation, fit values in Table 7.1 are similar for the Dudko models with final free energy of 28 kcal/mol (CD melting) or 23 kcal/mol (Jarzynski).

## Methods

In prior work<sup>95</sup> the G4-quadruplex was studied at a single loading rate of 5.5 pN/s and it was determined that the force distribution contains components from parallel and antiparallel forms of the structure. We took data at three loading rates (2.1 pN/s, 7 pN/s, and 23.9 pN/s) for analysis using DFS.

Double-stranded DNA handles were generated by PCR with digoxigenin and biotin labeled primers and digested with BtgI or BstXI enzymes to produce 4-base sticky ends. Data was taken using two different handle configurations with total lengths 3716 or 2432 base-pairs. A control hairpin studied previously<sup>12</sup> was synthesized as ***cggttttgagtc****aacgtctggatcctgtttcaggatccagacgttgactctttc****gat***, where bold indicates sticky ends to be used for ligation. The quadruplex was synthesized with sequence ***cgtg***(*acaggggtgtgggg*)<sub>2</sub>***acacgat***, identical to that used in previously reported results<sup>95,101</sup>. The quadruplex and control hairpin were ligated between the handles and then gel purified. The surface of the sample chamber was coated with blotting buffer to prevent interactions with the surface. Samples were attached to the surface through the digoxigenin label and attached to 0.82  $\mu\text{m}$  diameter beads using the biotin label as shown in Fig. 7.2. Final buffer conditions were 100 mM KCl, 2 mM EDTA, 10 mM Tris buffer (pH 8.0), 0.02% Tween 20, and oxygen-scavenging solution<sup>37</sup> (721  $\mu\text{g}/\text{ml}$  glucose oxidase, 144  $\mu\text{g}/\text{ml}$  catalase, and 3.9 mg/ml glucose) which increased the lifetime of the samples under exposure to the trapping beam. Data was collected on a single-beam optical trap at 10 kHz using an 8-pole 5 kHz Bessel filter. Standard methods were used to calibrate the position and force measurements<sup>41</sup>. The apparatus has been described elsewhere<sup>73</sup>.

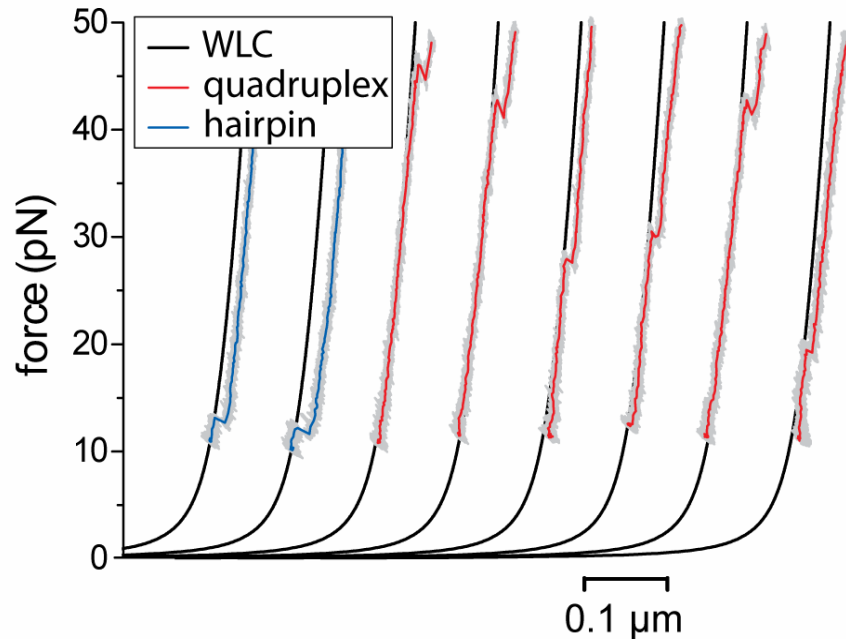


**Figure 7.2 Experimental setup.** The quadruplex is ligated between double-stranded DNA handles and attached to the surface and bead by digoxigenin and biotin labels respectively. Force is applied along the direction of the tether.

To provide a reliable force loading rate, feedback to the stage position was used to maintain the bead at a fixed position relative to the optical trap as the optical power was increased at a constant rate using an acousto-optic modulator. For each individual scan the initial force, the maximum accessible force, and the force of the observed disruption, if any, were recorded. To maintain accurate length and force calibration throughout a scan, the range of forces was restricted to lie below 50 pN.

Each sample was scanned repeatedly until the tether broke or the sample stopped showing any further disruptions. Between scans the sample was brought to a relaxed state at zero force for 5 seconds to allow the quadruplex time to refold. In order to ensure that the disruptions in our statistical sample came from quadruplexes that had formed under identical circumstances, we discarded the first disruption from each molecule.

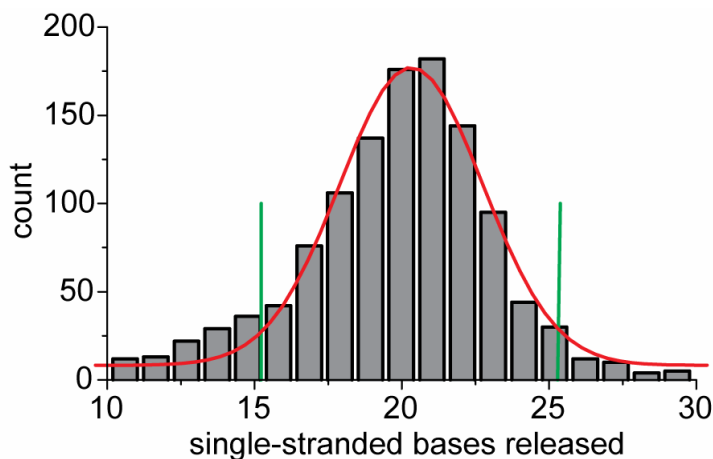
In order to verify the force and displacement calibration of our measurements and confirm slide-to-slide consistency, each sample chamber was prepared so that roughly half of the tethers had the G4-quadruplex molecule and the remainder had the control hairpin. The type of molecule linked to any given tether is easily determined from the nature of the disruptions observed. Measurements of the control hairpin yielded opening size 17.5 nm with standard deviation 1.7 nm, and disruption force of approximately 13 pN, in agreement with previously reported values<sup>12</sup>. Typical scans of the control hairpin and quadruplex are shown in Fig. 7.3.



**Figure 7.3** Raw data scans at 24 pN/s for quadruplex (red) and control hairpin (blue) overlaid on a worm-like chain model with persistence length 41 nm, stretch modulus 1277 pN<sup>40</sup>, and contour length per base of 0.338 nm<sup>39</sup>. Each scan was preceded by a relaxation period at zero force for 5 s.

## Analysis

We expect the 25-mer G4-quadruplex to release about 21 bases after accounting for the width of the quadruplex structure<sup>95</sup>. Opening distances were converted to the number of single-stranded bases using a worm-like chain model<sup>38</sup> for the single-stranded DNA with persistence length 1.25 nm, contour length of 0.625 nm/base, and neglecting enthalpic contributions<sup>12</sup>. We measured an opening size of 20.3 bases, standard deviation 2.5 bases, shown in Fig. 7.4. To assure that each tether analyzed had a single-molecule attachment we excluded tethers that failed to produce the expected low-force stretch curve or that exhibited asymmetry for  $x$  and  $y$  stretching. To exclude improperly folded molecules from the sample we excluded disruptions for which the number of bases released was more than 2 standard deviations from the mean value or for which multiple disruptions were observed on a single scan.



**Figure 7.4** Opening distance for the G4-quadruplex with mean 20.3 single-stranded bases released and standard deviation 2.5 bases. The data was filtered two standard deviations from the mean (green boundaries).

We fit our results to the Evans model ( $\nu = 1$ ) and Dudko models ( $\nu = 1/2$  and  $\nu = 2/3$ ) using maximum likelihood ( $L_m$ ). Models using these values of  $\nu$  can be solved analytically and are described in more detail by Dudko et al<sup>99</sup>. Since our disruptions exhibited the two peaks that have been identified as originating from parallel and antiparallel conformations, we based our analysis on a dual-component distribution. The likelihood  $L$  that an ensemble of measurements would be observed is defined as:

$$L = \prod_i^N \left[ w P_i(F_i | d_p, G_p, k_p; r_i, \nu) + (1 - w) P_i(F_i | d_a, G_a, k_a; r_i, \nu) \right] \quad 7.2$$

where the product is indexed over all scans and  $P_i$  is the theoretical probability density of observing a disruption event at force  $F_i$  and loading rate  $r_i$  for one of the two populations. The weight  $w$  indicates the fraction of the molecules in the parallel conformation (subscript p), while the fraction in the antiparallel conformation (subscript a) is  $1-w$ . Using one weight or separate weights for the three loading rates gives essentially the same results. The theoretical probability density function was normalized for each scan based on the range of forces actually measured. The best fit was determined by maximizing the natural log of  $L$  based on the dual population model defined by Eq. 7.2.

$\nu$	$\ln(L_m)/L_0$	$w$	$d_p(\text{\AA})$	$d_a(\text{\AA})$	$\log k_p(\text{s}^{-1})$	$\log k_a(\text{s}^{-1})$	$G_p$ (kcal/mol)	$G_a$ (kcal/mol)
1	0	$0.33 \pm 0.05$	$6.8 \pm 0.9$	$6.1 \pm 0.9$	$-1.7 \pm 0.2$	$-2.7 \pm 0.3$	-	-
2/3	$+2.2 \pm 1.0$	$0.29 \pm 0.05$	$13 \pm 3$	$10 \pm 2$	$-2.4 \pm 0.3$	$-3.3 \pm 0.5$	$5.3 \pm 0.4$	$5.1 \pm 1.2$
1/2	$+2.3 \pm 0.9$	$0.31 \pm 0.09$	$16 \pm 3$	$14 \pm 4$	$-2.7 \pm 0.3$	$-4.0 \pm 0.8$	$5.5 \pm 0.5$	$7 \pm 2$

**Table 7.1 Summary of global maximum-likelihood fits for different values of  $\nu$ .  $\ln(L_m)/L_0$  indicates the log of the maximum-likelihood score relative to the result for  $\nu = 1$ , normalized by  $L_0 = \ln(N)/2$  where  $N = 1014$  data points. Uncertainties were determined by bootstrapping. Weight ( $w$ ) indicates the percentage allocated to the parallel conformation (subscript p) while the antiparallel conformation (subscript a) is allocated percentage  $(1-w)$ .**

The best fit parameters for the three models as well as the corresponding likelihoods are summarized in Table 7.1, based on a total of  $N = 1014$  disruptions. The model which gives the largest value of the likelihood is preferred. However, when comparing models which do not have the same number of parameters we expect an increase in the log likelihood as additional parameters are introduced. The Bayesian information criterion specifies that an increase of  $\ln(N)/2$  is required to justify each additional parameter<sup>102,103</sup>. The Dudko models have two additional parameters,  $G_1$  and  $G_2$ , and result in likelihoods which marginally satisfy this criterion.

Fitting the distributions of disruption force in detail, rather than simply analyzing the most likely or mean disruption force as a function of loading rate, gives a more exhaustive comparison of data with theory, and allows the overlapping distributions associated with the parallel and anti-parallel conformations to be identified. In Fig. 7.5, histograms of disruptions at three different loading rates are compared to the global fit obtained for  $\nu = 2/3$  (Table I). Blue and red represent the

contribution from the parallel and antiparallel conformations respectively and green represents the sum of the two conformations.

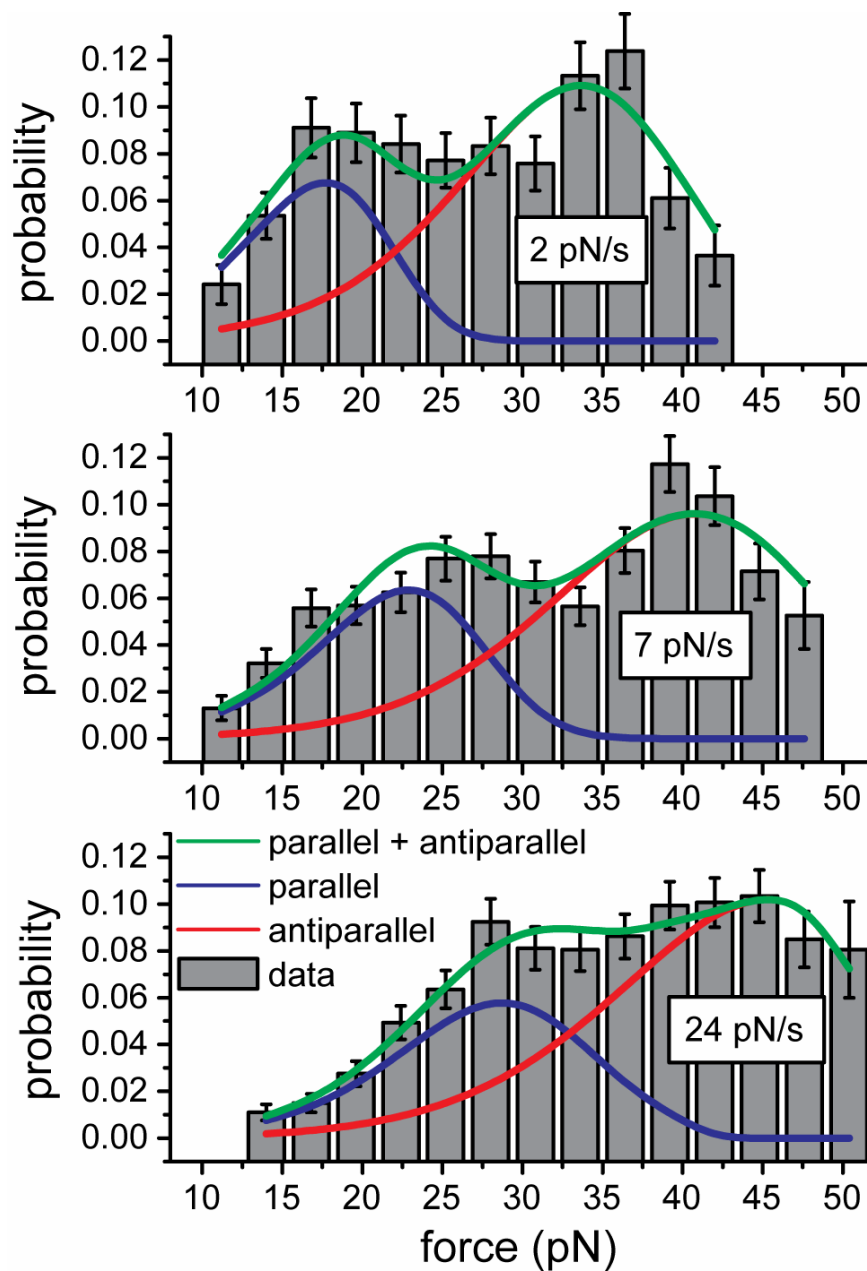


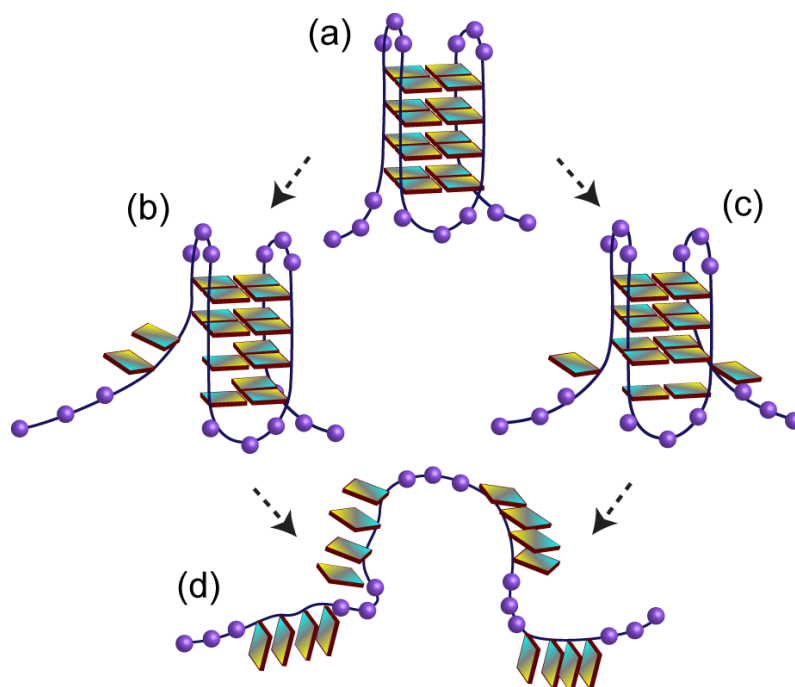
Figure 7.5 Global fit of G4-quadruplex data at three loading rates (2, 7, and 24 pN/s) for  $\nu = 2/3$ . Blue and red show the theoretical distribution for the parallel and antiparallel conformations respectively. Green represents the sum of the two conformations. The binned results provide a visual representation of the data and the best fits from Table 7.1. However, the maximum likelihood analysis used to generate the fits in Table 7.1 does not require binning of the data.



## ***Interpretation***

We now consider interpretation of the results reported in Table 7.1. The physical parameters obtained are relatively insensitive to the choice of  $\nu$ , allowing us to conclude that the G4-quadruplex energy barrier distance is approximately 10-16 Å, for both parallel and antiparallel conformations. The release of a nucleotide of single-stranded DNA will increase the tether length by approximately 5 Å, so this indicates that the transition state is reached after 2-3 bases have been pulled out of the structure. The transition state distance is a small fraction of the total amount of DNA released, and indicates that the transition state is much closer to the fully folded state than to the unfolded state.

A schematic model for disruption of the antiparallel conformation is shown in Fig. 7.6. The model illustrates hypothetical symmetrical and nonsymmetrical disruption pathways for the antiparallel conformation, in which the first two base-pairs are released from the same side or opposite sides of the structure, respectively. (Analogous models can be formulated for the parallel case.) Formation of the quadruplex appears to be highly cooperative, in that a relatively small perturbation to the structure is required to reach the transition state and progress to complete dissociation.



**Figure 7.6** A possible model of G4-quadruplex disruption. (a) Original antiparallel conformation. Applied force causes two bases to be released from the G4-quadruplex, either antisymmetrically (b) or symmetrically (c). (d) Upon release of the 3rd base, the entire structure disrupts.

The Dudko models indicate that the height of the energy barrier is approximately  $10 k_B T$ . An applied force will tilt the energy landscape, adding a term  $-Fx$  to the effective energy as a function of the opening distance  $x$ , and disruption typically occurs when the force cancels the slope of the native landscape. The quadruplex has a high disruption force because of the short distance and steep slope to the principal barrier, despite its modest height. However, DFS measurements will not be sensitive to a secondary barrier beyond  $d$  with a more shallow energy slope<sup>104</sup>. In Fig. 7.1, we consider a model which includes our measured primary barrier followed by a more shallow energy barrier slope. The final barrier height is

determined from previous measurements<sup>95</sup> using CD melting or Jarzynski's equality<sup>100</sup>. The rate of passing the principal barrier is  $k_{\text{off}} \approx A \exp(-G/k_{\text{B}}T)$ <sup>1</sup>, where  $A$  is a coefficient which depends on the effective diffusion coefficient and the shape of the barrier. It has been demonstrated that quadruplex structures can hinder molecular motors<sup>105</sup>. Processive enzymes which are potentially stalled by the high mechanical stability of the principal barrier may be assisted by thermal fluctuations and then progress more easily along the shallower regime of the energy barrier.

Alternative methods of measuring G-quadruplex kinetics include nanopores<sup>106</sup>, CD<sup>95</sup>, FRET<sup>87</sup>, and AFM<sup>25</sup>. The results of Table 7.1 are comparable with those reported in a prior AFM study of a bi-molecular form of another G4-quadruplex sequence, though the previous study assumed a single unknown conformation type<sup>25</sup>. Although experiments based on optical traps cannot duplicate the very high loading rates achieved in AFM studies, the high precision with which disruption force can be measured results in very accurate measurement of the distribution of rupture forces. In contrast to Lynch, this has allowed us to resolve the parallel and antiparallel conformations of the molecule, and to determine the transition state distances of each conformation with sub-nanometer precision.

## ***Conclusion***

In conclusion, we have measured the disruption force and opening distance of a G4-quadruplex at three loading rates, resolving two force distributions which have previously been shown to originate in two distinct quadruplex conformations<sup>95</sup>. The high mechanical stability previously demonstrated originates from the short steep barrier we measured for both forms of the quadruplex. These parameters have implications for regulatory processes. The high-force resistance associated with the short barrier distance would make G-quadruplex difficult for a processive enzyme to disrupt. However, thermal fluctuations may assist an enzyme beyond the principal barrier where disruption can progress more easily.

## Chapter 8: Noise Associated With Nonconservative Forces in Optical Traps<sup>h</sup>

### ***Introduction***

An important component of our model for energy landscapes is a convolution parameter which accounts for uncertainty in our measurements due to fluctuations of the DNA handles as well as general sources of noise. It is therefore essential to fully understand the mechanisms of noise in the optical trap. This research is theoretical and investigates the effect of noise due to the nonconservative scattering force from the laser. The results are fundamentally dependent on derivations of Eq. 8.4 and 8.5 by our collaborator, Natalia Denesyuk<sup>107</sup>.

### ***Background***

It is known that for a particle held in an optical trap, the interaction of thermal fluctuations with a nonconservative scattering force can cause a persistent nonequilibrium probability flux in the particle position. We investigate position fluctuations associated with this nonequilibrium flux analytically and through simulation. We introduce a model which reproduces the nonequilibrium effects, and in which the magnitude of additional position fluctuations can be calculated in closed form. The ratio of additional nonconservative fluctuations to direct thermal fluctuations scales inversely with the square root of trap power, and is small for typical experimental parameters. In a simulated biophysical experiment the

---

<sup>h</sup>Adapted from de Messieres, M., Denesyuk, N. A. & La Porta, A. Noise associated with nonconservative forces in optical traps. *Physical Review E* **84**, 031108 (2011).

nonconservative scattering force does not significantly increase the observed fluctuations in the length of a double-stranded DNA tether.

In a single-beam optical trap<sup>2,4</sup>, a dielectric particle is confined to the focal volume by a force  $\vec{F}_g$  proportional to the optical intensity gradient. The force exerted on the particle and the resulting displacement can be measured by detecting the deflection of the trapping beam<sup>9,108,109</sup>. Using this technique, the movement of a biological macromolecule attached to a trapped particle can be measured with near-ångström precision<sup>110</sup>. However, the trapped particle is also subject to a scattering force  $\vec{F}_s$ , first noted by Ashkin<sup>111</sup>, which cannot be expressed as the gradient of any function. Assuming the geometrical center of the trapping beam is defined as the origin of coordinates, with beam propagation in the positive  $z$  direction, the scattering force acts mainly in the  $+z$  direction and causes the effective trap center to lie at positive  $z$ , typically within a wavelength of the geometrical trap center. In the limit of a small particle, the scattering force is proportional to the local intensity and decreases with distance from the beam axis. For particles that are large compared with the wavelength, the scattering force can be estimated using ray tracing, and increases with small displacement from a diffraction-limited trap center<sup>111</sup> (although it must ultimately decrease for large displacements). For small displacements both curves, illustrated in Fig. 8.1a, have a roughly quadratic dependence on distance from the beam axis.

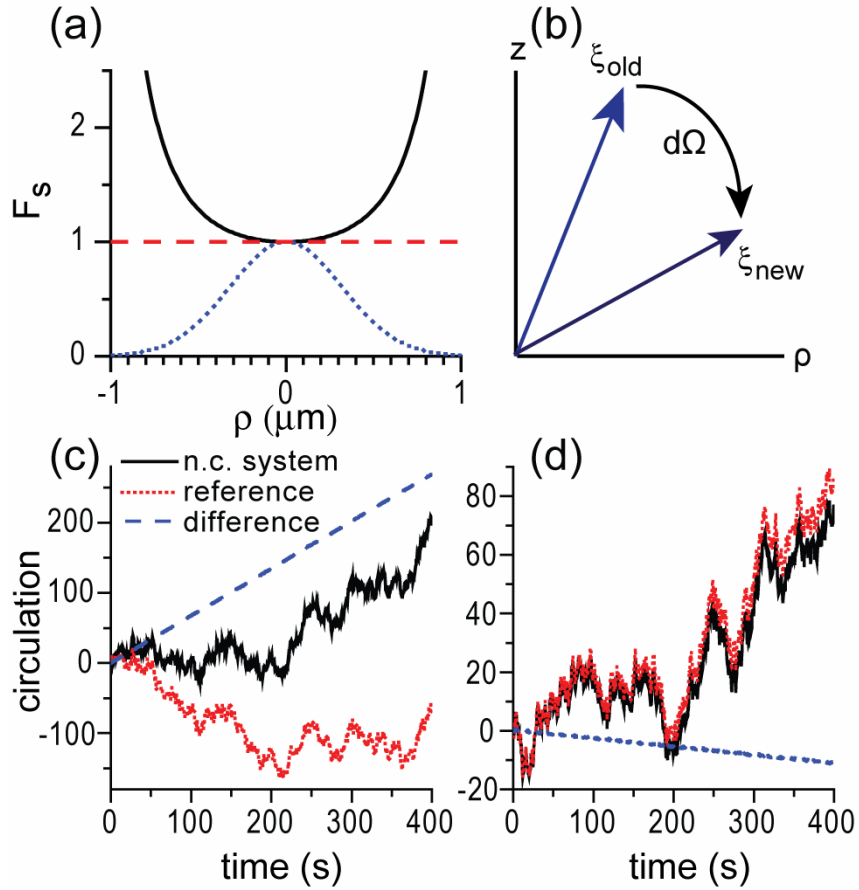


Figure 8.1 (a) Scattering force profiles for a small particle (blue dash line) and large particle (solid black line). The small particle curve is a Gaussian with standard deviation  $\sigma = 0.32 \mu\text{m}$  and the large particle curve is based on the ray tracing result with a particle of radius  $1 \mu\text{m}$ . (b) Circulation is defined as the area swept out in the  $\rho$ - $z$  plane, where clockwise motion is considered positive by convention. (c) Circulation obtained from the nonconservative system (solid black) and reference system (short dash) driven by identical thermal force fluctuations. The difference is shown in the long dashed line. Physical parameters:  $T = 298 \text{ K}$ ,  $\alpha_x = \alpha_y = 2.0 \text{ pN}/\mu\text{m}$ ,  $\alpha_z = 0.4 \text{ pN}/\mu\text{m}$ , time step  $dt = 0.0001 \text{ s}$ , particle radius  $r_p = 0.25 \mu\text{m}$  and nominal scattering force given by the Gaussian function in panel (a) with  $F_0 = 0.04 \text{ pN}$  on the beam axis. (d) Same but using the large particle form of the scattering force with  $r_p = 1.0 \mu\text{m}$  and  $F_0 = 0.04 \text{ pN}$  on the beam axis.

## ***Simplified Model***

Our goal is to define a model system which reproduces the behavior of the nonconservative system with as simple a force field as possible. We model the gradient force as a harmonic restoring force and assume that the scattering force is axial,

$$\vec{F}(\vec{r}) = -\alpha_x x \hat{x} - \alpha_y y \hat{y} - \alpha_z z \hat{z} + F_s(\rho) \hat{z} \quad 8.1$$

where  $\alpha_i$  is the stiffness of the harmonic force along the  $i$  axis and  $F_s(\rho)$  is the magnitude of the scattering force as a function of  $\rho = \sqrt{x^2 + y^2}$ . We will expand the scattering force as  $F_s(\rho) = S_0 + S_2 \rho^2$ , where the second-order term is nonconservative because it cannot be derived from a scalar potential and does net work on a particle that moves on a closed path. For the small particle, assuming a scattering term of the form  $F_0 \exp(-\rho^2 / 2\sigma^2)$ , we have  $S_2 = -F_0 / 2\sigma^2$ , where  $\sigma$  is the standard deviation of the scattering force profile, which depends on both the beam profile and particle radius. For the large particle, Taylor expansion of the Ashkin form around the effective trap center gives  $S_2 = 0.8F_0 / r_p^2$ , where  $r_p$  is the particle radius<sup>111</sup>. We will show that this model reproduces the circulation effect observed in previous studies while permitting important simplifications in the simulations and analytical analysis. We will also show that the analytical solution reproduces the fluctuations observed, not only in a simulation of the simplified system, but in a realistic simulation which takes into account the 3-dimensional nature of the intensity gradient and scattering force fields.



The nonequilibrium probability flux in the particle position  $\vec{r}$  that results from the nonconservative component of the scattering force was characterized by Roichman et al. in terms of the circulation  $\Omega$ , as illustrated in Fig. 8.1b<sup>112</sup>. Each discrete measurement of  $\vec{r}$  contributes a differential circulation of

$$d\Omega = \frac{\vec{\xi}_{old} \times \vec{\xi}_{new}}{2A} \quad 8.2$$

where  $\vec{\xi}$  is a 2-dimensional vector defined by  $\vec{\xi} = \{\rho, z\}$ . The characteristic area  $A$  defines a unit of circulation, and the cross product defines the sign of circulation to be positive for clockwise motion in the  $\rho$ - $z$  plane.

Previous work has shown that circulation is dominated by the Brownian motion of the particle, requiring very long simulation times to achieve statistical convergence<sup>113</sup>. However, the contribution of the nonconservative force to the circulation can be efficiently calculated by comparing results of the nonconservative system with a conservative reference system, in which the particle is subject to identical Brownian force fluctuations, but in which the nonconservative scattering force is replaced by a constant force. By looking at the difference in circulation between the original system and the reference system, circulation fluctuations arising directly from thermal forcing cancel out, isolating the circulation arising from the nonconservative force.

We use a Langevin equation of the form  $\vec{F}_T(t) + \vec{F}(\vec{r}) - \beta\dot{\vec{r}} = 0$  to simulate the model optical trap, where  $\beta$  is the Stokes drag,  $\vec{F}(\vec{r})$  is obtained from Eq. 8.1, and  $\vec{F}_T(t)$  is a delta function-correlated stochastic noise term whose components have

spectral density  $|\tilde{F}_T(\omega)|^2 = 2k_B T\beta$ . We define two primary nonconservative systems using the two scattering curves in Fig. 8.1a and compare them with conservative reference systems which are driven by identical thermal force fluctuations. The results are shown in Fig. 8.1c and 8.1d. Thermal diffusion dominates for short times, causing large fluctuations in circulation, while the nonconservative contribution manifests itself as a linear increase in net circulation over longer times, with opposite signs for large and small particles as expected. There is a large degree of cancellation of the circulation fluctuations when the primary and reference system are subtracted, revealing the excess circulation arising from the nonconservative term. The excess circulation obtained from subtraction of the reference system from the primary system is exactly the same as the value obtained over long time scales using a single simulation, but comparison with a matched reference system gives us additional insight into the generation of circulation, since it reveals that even over short time intervals the excess circulation accumulates at a nearly constant rate. This supports the view that the effect of the nonconservative force is a continuous biasing of fluctuations. More importantly, Figs. 8.1c and 8.1d confirm that our simplified system reproduces the nonequilibrium flux which is the essential feature of the nonconservative scattering force.

### ***Calculation of Excess Fluctuations***

We now consider whether the nonconservative force results in additional fluctuations in the particle position. In the harmonic approximation, and assuming that the scattering force acts only in the positive  $z$  direction (Eq. 8.1), the  $x$ - $y$

components of the total force acting on the particle do not depend on the  $z$  coordinate. As a result, the evolution of the  $x$ - $y$  coordinates is independent of the  $z$  motion, and is unaffected by the axial scattering force.

The independence of the transverse dynamics from the axial motion allows us to calculate the excess position fluctuations in the nonconservative system. For a simple harmonic trap, Brownian fluctuations in  $z$  are described by a thermal force  $F_T$  acting in a strongly-damped equation of motion  $F_T(t) - \alpha_z z - \beta \dot{z} = 0$ . Using the known spectral density of the thermal force term, this gives rise to fluctuations in the  $z$  coordinate with Fourier spectral density

$$|\tilde{z}_T(\omega)|^2 = \frac{|\tilde{F}_T(\omega)|^2}{\alpha_z^2 + \beta^2 \omega^2} = \frac{2k_B T \beta}{\alpha_z^2 + \beta^2 \omega^2} \quad 8.3$$

where similar results apply to  $x$  and  $y$ . However, in the nonconservative system the  $z$  coordinate has an additional source of noise since the scattering force depends on  $\rho^2$  and fluctuates in time as the distance between the particle and the beam axis varies, despite the fact that the scattering force itself has no explicit time dependence. This gives an additional fluctuating driving term for the  $z$  motion which originates entirely in the  $x$ - $y$  dynamics of the particle. Since the  $x$ - $y$  dynamics are independent of the  $z$  motion, this additional term, which is formally a function of  $\rho$ , is effectively a stochastic function of time, a pseudo-thermal forcing term which is uncorrelated with the thermal force fluctuations  $F_T(t)$  driving the  $z$  thermal motion. In contrast to  $F_T(t)$ , the scattering force fluctuations manifest finite time autocorrelations determined by the continuous  $x$ - $y$  motion, or equivalently, have a spectral density which is related to the spectral density of fluctuations in  $\rho^2$ . An extensive calculation which is detailed

in the supplemental material indicates that the nonconservative driving term arising from  $x$  fluctuations has Fourier spectral density<sup>107</sup>

$$|\tilde{F}_{\text{nc}}(\omega)|^2 = \frac{8(k_{\text{B}}T)^2 S_2^2}{\alpha_x \beta \left( \omega^2 + \frac{4\alpha_x^2}{\beta^2} \right)} \quad 8.4$$

where similar results are obtained for  $y$  fluctuations. The equation of  $z$  motion is linear, so the solution can be decomposed into components arising from the thermal driving term (Eq. 8.3) and the nonconservative driving term. The spectral density of  $z$  fluctuations resulting from the nonconservative force is

$$|\tilde{z}_{\text{nc}}(\omega)|^2 = \frac{16(k_{\text{B}}T)^2 S_2^2}{\alpha_x \beta^3 \left( \omega^2 + \frac{4\alpha_x^2}{\beta^2} \right) \left( \omega^2 + \frac{\alpha_z^2}{\beta^2} \right)} \quad 8.5$$

where we set  $\alpha_x = \alpha_y$  for clarity.

The Fourier spectra of the thermal and nonconservative forces are shown for the case of a large particle in Fig. 8.2a and for a small particle in Fig. 8.3a (using both the full scattering form and its second-order expansion). The Fourier spectra of the fluctuations in  $z$  driven by the thermal and nonconservative forces are shown in Figs. 8.2b and 8.3b. Precise agreement is found between the analytical spectra, indicated by solid curves, and the simulations. It is clear from these spectra that the characteristic time scale of nonconservative fluctuations is determined by the  $x$ - $y$  dynamics of the particle, rather than by the circulation time inferred from data in Fig. 8.1c and 8.1d. The uncertainty in a measurement of the particle position is found by integrating the spectral density over the bandwidth of the measurement. However, the total root mean square (RMS) fluctuations of the thermal and nonconservative

components of the displacement,  $z_T$  and  $z_{nc}$  (integrated over all frequencies) are a useful practical measure of noise in experiments where the localization of the particle is of primary importance. We find that the standard deviation for thermal fluctuations is  $\sigma_T = \left(\frac{k_B T}{\alpha_z}\right)^{1/2}$ , and for nonconservative fluctuations is  $\sigma_{nc} = \frac{10 S_2 k_B T}{\sqrt{11} \alpha_x^2}$  (where we set  $\alpha_x = \alpha_y = 5 \alpha_z$ ). These are plotted as a function of  $x$  stiffness in the Fig. 8.2b and 3b inserts, with the assumption that the scattering force and trap stiffness both scale linearly with the trap optical power  $P$ . The ratio of the standard deviations of nonconservative and thermal fluctuations scales as  $P^{-1/2}$ , indicating that the importance of the nonconservative effect decreases as the trap becomes stiffer. Using the physical parameters chosen for the spectra in Fig. 8.2, the RMS fluctuations arising from the direct thermal fluctuations are approximately 1000 times larger than those arising from the scattering force. Since the thermal and nonconservative noise sources are uncorrelated and add in quadrature, the augmentation of the amplitude of total fluctuations is less than 1 part in  $10^6$ .

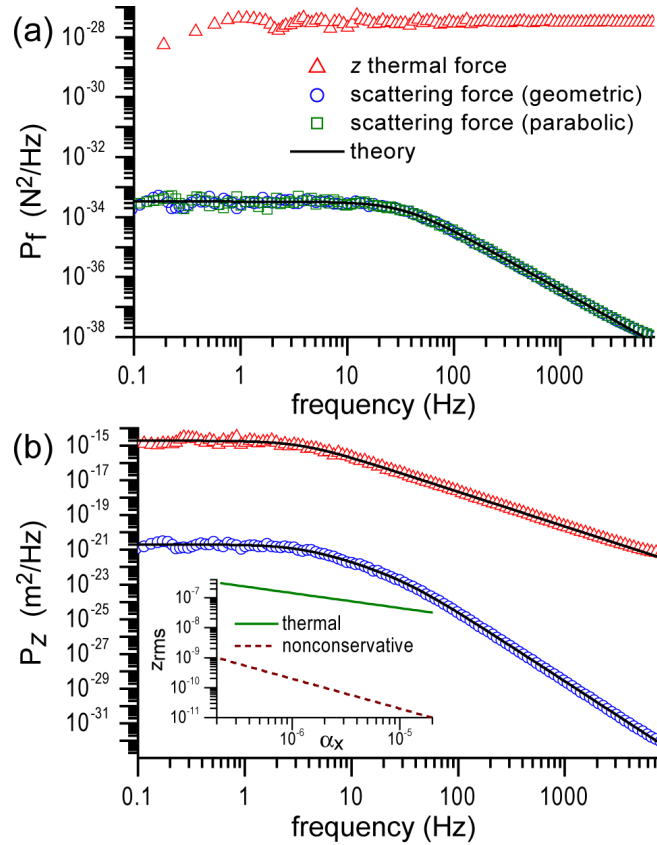


Figure 8.2 (a) Fourier spectral density of the thermal driving force (triangles) and the nonconservative  $z$  scattering force using the large-particle profile of Ashkin (circles) and its second-order Taylor expansion (squares) for dynamics obtained from numerical simulation of a Langevin equation based on Eq. 8.1 with the same parameters as Fig. 1. Black line: The analytical prediction for scattering force fluctuations (Eq. 8.4). (b) Fourier spectral density of the  $z$  motion driven by thermal forcing (triangles) and by the nonconservative scattering force (circles). Black lines: Analytical predictions for  $z$  fluctuations driven by thermal forcing (Eq. 8.3) and the scattering force (Eq. 8.5). Inset: RMS  $z$  fluctuations arising from the thermal forcing and from the scattering force as a function of  $\alpha_x$ , assuming  $\alpha_z$  and  $F_0$  are linearly related to  $\alpha_x$ .

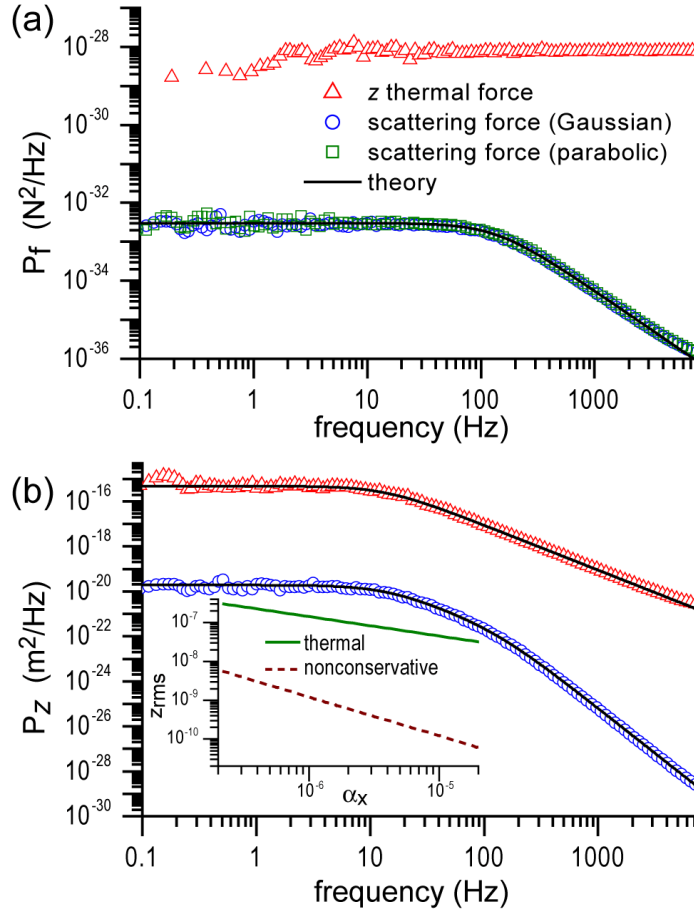


Figure 8.3 Similar to Fig. 2, except the particle is assumed to be small compared with the wavelength, and the scattering profile is assumed to have a Gaussian form which corresponds to the intensity profile of the trapping beam at the trap center. Physical parameters are identical to those used in the small particle simulation in Fig. 1.

### ***Application to a Realistic Force Field***

In the section above, we have shown that the simplified system, consisting of a harmonic trapping force and an axial scattering force which depends only on  $\rho$ , reproduces the nonequilibrium probability flux which is the main feature of the nonconservative scattering force. The simplifications facilitated an analytical calculation of the level of excess noise introduced by the nonconservative force. In

this section we consider a realistic system, in which the gradient (trapping) and scattering forces are derived from the 3-dimensional Gaussian optical mode. Although this realistic system contains nonlinearities, crosstalk between axial and transverse dynamics, and a scattering force which is no longer purely axial, we will show that the analytic calculations derived from the simplified model accurately predict scattering force fluctuations observed in simulations of the realistic system.

In the realistic simulation we take the gradient force to be proportional to the gradient of the optical intensity and the scattering force to be proportional to the local Poynting vector, which is proportional to the intensity itself and acts along the phase gradient direction<sup>114</sup>. This form is appropriate for a small particle, but can also be used as a highly accurate parameterization of the measured force on a finite-size particle as a function of displacement in experiments. The 3-dimensional, nonlinear force functions introduce new effects not present in the linearized system. Differential stiffness decreases with distance from the geometrical center of the trap. In addition, the optical mode spreads as it propagates past the trap center, and as a result there is a decrease in the transverse restoring force with increasing  $z$  in the neighborhood of the effective trap center. The spreading of the optical mode also causes the scattering force to weaken with increasing  $z$  as well as with increasing  $\rho$ . The  $z$  dependence of the intensity gradient and scattering forces will cause transverse force fluctuations to arise from axial position fluctuations, which was not included in the simpler model introduced above. The question arises whether these additional effects result in an increase in fluctuations of the particle position.



To address this issue we run the simulation of a particle subject to the scattering force vector and gradient force based on the 3-dimensional Gaussian mode<sup>114</sup>. However, in this case the scattering force includes a conservative contribution which can be represented as the gradient of a potential, as we explain below. In the neighborhood of the effective trap center the  $z$  component of the scattering force decreases with increasing  $z$ . This predominantly linear dependence of the axial scattering force on  $z$  serves as an enhancement of the axial restoring force experienced by the particle – the particle is more strongly trapped in the presence of the  $z$ -dependent scattering force than it would be if the scattering force were independent of  $z$ . The  $z$  dependence of the axial scattering force can therefore be represented by a potential, and does not contribute to the nonequilibrium effects that result from nonconservative forces. We therefore decompose the scattering force into conservative and nonconservative components and calculate the power spectra of the two components separately. The decomposition is done as follows. The magnitude of the full scattering force is of the form

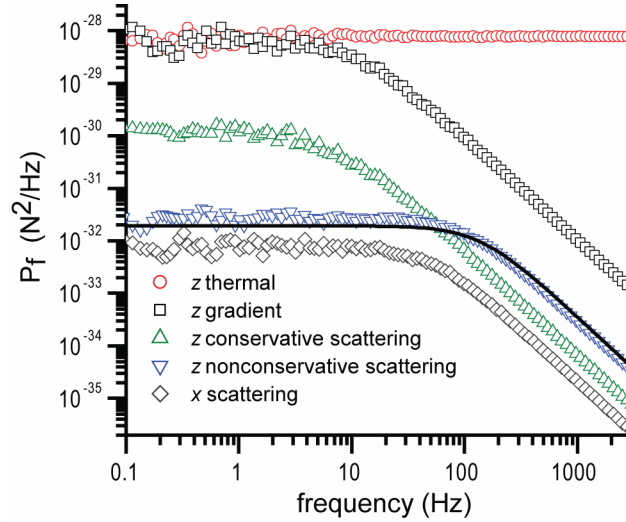
$$\vec{F}_s(x, y, z) = F_0 \left( \frac{1}{1 + z^2/z_0^2} \right) \exp \left[ \frac{-2(x^2 + y^2)}{\omega_0^2 (1 + z^2/z_0^2)} \right] \hat{S} \quad 8.6$$

where  $\hat{S}$  is a unit vector in the direction of the Poynting vector (which corresponds to the phase gradient direction),  $\omega_0$  is the Gaussian spot size and  $z_0$  is the Rayleigh distance,  $z_0 = \frac{\pi \omega_0^2 n}{\lambda}$ . We wish to calculate the potential whose gradient most closely matches the total scattering force in the region visited by the particle. The difference between the actual scattering force and this conservative approximation

would represent the nonconservative component of the scattering force. Since the particle is confined to a small volume centered on the beam axis, it is sufficient to choose the conservative component of the scattering force to duplicate the  $z$  dependence of the  $z$  component of the scattering force on the beam axis,  $\vec{F}_{sc}(z) = \vec{F}_s(0,0,z)$ , where the lack of  $x$  and  $y$  dependence is determined by the criteria that  $\vec{F}_{sc}(z)$  is conservative. The nonconservative component is obtained by subtracting the conservative component from the full scattering force,  $\vec{F}_{snc}(x,y,z) = \vec{F}_s(x,y,z) - \vec{F}_{sc}(z)$ . Making use of this decomposition of the scattering force, the effective conservative force acting on the particle is the combination of the intensity gradient force and the conservative component of the scattering force. The remainder of the scattering force,  $\vec{F}_{snc}(x,y,z)$ , contains the transverse dependence of the scattering force and is the driving term for nonconservative effects.

Fig. 8.4 shows the fluctuations in the scattering force in a simulation using the 3-dimensional restoring and scattering forces. The decomposition into conservative and nonconservative components only applies to the calculation of the power spectra of the scattering force, the particle dynamics are generated from the full scattering force. The system parameters are comparable to those used to produce Fig. 8.3, although the effective trap power has been increased by a factor of approximately 5% to maintain the same effective trap stiffness at the equilibrium point in the nonlinear simulation. The scattering profile, adopting the small-particle limit, is strictly defined by the 3-dimensional optical mode, resulting in a Gaussian profile with standard deviation  $\sigma = \omega(z)/2$  where  $\omega(z)^2 = \omega_0(1 + z^2/z_0^2)$ . The value of  $\omega_0$ , has been set

to  $0.411 \mu\text{m}$  in order to obtain the nominal 5:1 ratio between axial and transverse stiffness that we use to evaluate our analytical theory. The resulting scattering profile is narrower than the one that was used in the previous simulations and results in somewhat stronger scattering fluctuations in Fig. 8.4.



**Figure 8.4** Power spectra of force components acting on a small particle assuming the gradient force and scattering force are proportional to the intensity gradient and Poynting vector, respectively, of the 3-dimensional Gaussian optical mode. Fourier spectral density of the thermal driving force (circles), axial gradient force (squares), conservative axial component of the scattering force (triangles), nonconservative axial component of the scattering force (inverted triangles) and the  $x$  component of the scattering force (diamonds). The solid line represents the analytical calculation of the nonconservative scattering force contribution predicted by Eq. 8.4. The time step is  $5 \mu\text{s}$ , and the physical parameters are  $\alpha_x = \alpha_y = 2.01 \text{ pN}/\mu\text{m}$ , and  $\alpha_z = 0.360 \text{ pN}/\mu\text{m}$  and  $F_0 = 0.04 \text{ pN}$  (evaluated at the effective trap center) using the point particle approximation with  $\omega_0 = 0.411 \mu\text{m}$ . The effective  $z$  stiffness contains a contribution of  $0.026 \text{ pN}/\mu\text{m}$  from the conservative component of the scattering force.

Fig. 8.4 has several interesting features. The thermal force spectrum  $|\tilde{F}_T(\omega)|^2$  and the intensity gradient force spectrum  $|\tilde{F}_g(\omega)|^2$  correspond closely at low frequency. Comparison of the phase of the two spectra (not shown) indicates they are anti-coherent. This is consistent with the fact that low frequency wandering of the particle is largely suppressed by the optical trap, which implies that the gradient force neutralizes the thermal force at low frequency. Above the characteristic frequency  $f_0 = \alpha_z/2\pi\beta$  of axial fluctuations, the spectrum of the gradient force falls below that of the thermal forcing, consistent with the fact that high frequency jiggling of the particle is relatively unaffected by the trapping potential. The spectrum of the conservative component of the scattering force  $|\tilde{F}_{sc}(\omega)|^2$  is similar to that of the gradient force, but at lower amplitude. These spectra are phase coherent and confirm a slight enhancement of the restoring force by the conservative component of the scattering force. The spectrum of the  $z$  component of the nonconservative scattering force component  $|\tilde{F}_{snc}(\omega)|^2$  represents the driving force for nonconservative fluctuations. Unlike  $|\tilde{F}_{sc}(\omega)|^2$ , it is generated by  $x$ - $y$  motion of the particle and its dependence on the  $z$  coordinate has been subtracted off. It therefore acts as a generator of independent axial fluctuations, rather than as a component of the restoring force. The solid line shows the prediction of Eq. 4 for  $|\tilde{F}_{snc}(\omega)|^2$ , using the effective stiffness (including contributions from the gradient and conservative scattering forces) and the on-axis scattering force  $F_0$ , both evaluated at the effective

trap center. There is good agreement between the analytical calculation based on the linearized system and the dynamics found in the fully nonlinear system. The level of fluctuations in the simulation is slightly in excess of the analytical prediction, apparently because the trap is barely stable for these parameters and the particle transiently wanders from the effective trap center to regions where the effective stiffness is smaller than the nominal value. This small discrepancy vanishes if the trap power is increased by a factor of 2 or more. The spectrum of force fluctuations associated with the nonconservative component of the scattering force is nearly four orders of magnitude below the direct thermal forcing. Fig. 8.4 also shows the spectrum of fluctuations in the  $x$  (tangential) component of the scattering force, which is smaller than both the conservative and nonconservative components of the axial scattering force, and is therefore not a significant source of additional fluctuations.

The decomposition of *position* fluctuations into components arising from distinct forcing terms that was performed in the linear system is no longer possible in the nonlinear system. However, the relative strengths of forcing from conservative and nonconservative forces have been calculated with high precision, and there is no reason to believe that the nonlinear effects or cross-terms in the realistic model would preferentially amplify force fluctuations originating from nonconservative relative to those originating from thermal fluctuations. Therefore the increase in position fluctuations in the presence of nonconservative effects should be in proportion to the increase in force fluctuations. No significant excess fluctuations were detected in the spectra of  $x$ ,  $x^2$  and  $z$  when comparing the simulation of the nonlinear system to one based on an equivalent linear system (data not shown), confirming that the additional

dynamic processes present in the nonlinear model (nonlinearities and cross-terms in the equations of motion) do not change the character or amount of fluctuations introduced by the nonconservative force term.

### ***Additional Fluctuations When Stretching a DNA Tether***

The results above apply to the case where there is no external force on the trapped particle and it remains in the neighborhood of the effective trap center. In single-molecule biophysics experiments, the optical trap is typically used to apply a substantial force to a biological macromolecule, and the equilibrium position, taking into account the external force, will no longer lie on the beam axis. In this regime the scattering force will no longer be an even function of the displacement  $x$  or  $y$  from the equilibrium position and may have a substantial non-axial component. As in the previous section, the nonlinearity of the system prevents us from decomposing the motion into components arising from distinct forcing terms, we can gain insight into the level of nonequilibrium fluctuations by measuring fluctuations of the scattering force itself in a realistic simulation.

We model an experiment in which the optical trap is used to measure variations in the length of a DNA tether. We assume the DNA is anchored to a cover slip, the tether making a  $45^\circ$  angle with the beam axis. As in Fig. 8.4, we assume that the scattering force and gradient force are proportional to the local Poynting vector and intensity gradient vector, respectively, of the 3-dimensional Gaussian mode of the trapping beam<sup>114</sup>. The attachment point is chosen so that the force exerted on the tether is  $3/4$  of the maximum force that the optical trap can produce at  $45^\circ$ , well past the regime of a linear restoring force. Unlike the simulation illustrated in Fig. 8.4, the

particle is sufficiently far from the trap center that cross terms, such as the dependence of non-axial force on axial position, and vice versa, are substantial. The force vs. extension curve of the DNA tether is calculated using the Marko-Siggia worm-like chain model with contour length of  $2.0 \mu\text{m}$ <sup>38</sup>.

Force fluctuations in the simulation of a tether experiment are shown in Fig. 8.5. The curves shown in Fig. 8.5b correspond to those shown in Fig. 8.4, except that the effective potential experienced by the particle includes the tension in the DNA tether. As in Fig. 8.4, the spectrum of the combined restoring force (including the gradient force and the tether force) cancels the thermal force at low frequency. The conservative part of the scattering force opposes the gradient force (since the particle is pulled to negative  $z$  by the DNA tether) but is small compared to the gradient and tether forces. The spectra of the non-conservative components of the scattering force (axial and non-axial) are again far below the level of the thermal force spectrum.

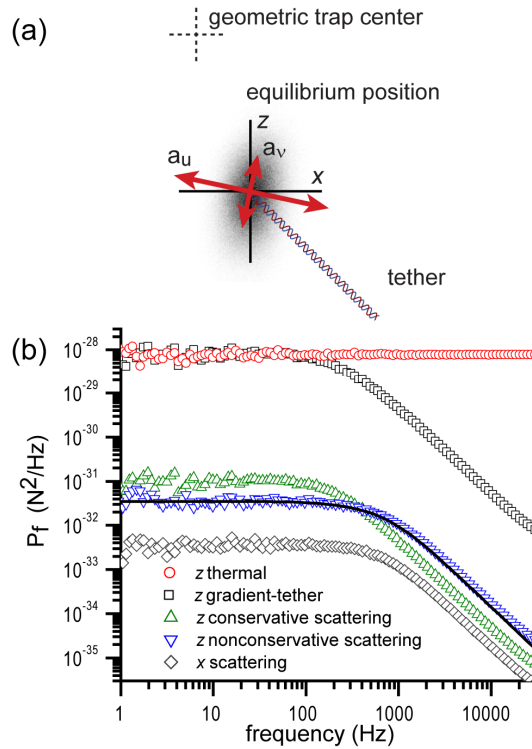


Figure 8.5 Simulation of a particle tethered by a DNA molecule at  $45^\circ$ , in which the tether attachment point is displaced from the trap center in the  $x$ - $z$  plane such that the tension in the DNA is  $3/4$  of the maximum force the trap can exert at  $45^\circ$ . (a) Schematic illustrates the effective conservative force characteristics for the particle. The combined gradient force, conservative component of the scattering force and tether force were linearized around the equilibrium point. The principal axes of the resulting elasticity tensor are illustrated by arrows which are superimposed on the probability density function of particle position obtained in the simulation. (b) Fourier spectral density of the thermal driving force (circles), combined gradient and tether force (squares), conservative axial component of the scattering force (triangles), nonconservative axial component of the scattering force (inverted triangles), and the  $x$  component of the scattering force (diamonds). The solid curve represents the analytical calculation of the nonconservative scattering force contribution, as described in the text. Simulations performed at tension  $1.2$  pN with  $r_p = 0.250$   $\mu\text{m}$ , Gaussian spot size  $\omega = 0.411$   $\mu\text{m}$  and  $F_0 = 0.4$  pN. The effective stiffnesses are  $a_u = 21.7$  pN/ $\mu\text{m}$  and  $a_v = 6.4$  pN/ $\mu\text{m}$  with  $\phi = -10^\circ$ .



The theoretical framework defined above can be used to calculate the nonconservative nonaxial scattering force fluctuations. In the present case, assuming the attachment pulls the particle away from the beam axis along the  $x$  direction, the expansion of the axial scattering force about equilibrium position will have a nonzero first-order term,  $\vec{F}_s = (S_0 + S_1 x)\hat{z}$ . To first order in the displacement  $x$ , the spectrum of scattering force fluctuations can be obtained by simply multiplying the spectrum of  $x$  fluctuations by  $S_1$ . However, the spectrum of  $x$  fluctuations is more complex for a trapped particle which is tethered to the surface than for an otherwise unconstrained particle. For a free particle trapped on the beam axis the motion is easily resolved into axial and transverse components, which have different effective stiffnesses. The tethered particle experiences a more complex force field which is the combination of the optical trap and DNA elasticity. By combining the effective potential energy arising from the optical trap, the tether elasticity and the conservative portion of the scattering force we can define the total potential energy of the particle as a function of position,  $U(x,y,z)$ . To first order the force arising from a small displacement in an arbitrary direction is found by multiplying the displacement vector by the Hessian matrix of  $U(x,y,z)$ . However, the Hessian evaluated at the equilibrium position is not diagonal, meaning a displacement along the  $x$ ,  $y$  or  $z$  coordinate axis results in a force which is generally not parallel to the displacement. The eigenvectors of the Hessian matrix, however, define a coordinate system in which the Hessian matrix is diagonal, and in which its diagonal elements are the stiffnesses for displacement along the corresponding eigenvectors. The dynamics along the three eigenvectors are independent and governed by the corresponding stiffness. (This procedure is not

necessary for the particle trapped on the beam axis because the axial and transverse directions themselves are eigenvectors.) The result of this analysis is illustrated in Fig. 8.5a. Two eigenvectors lying in the  $x$ - $z$  plane are superimposed on a probability density function of the particle position in the simulation. The axis of maximum stiffness, labeled  $u$ , makes an angle  $\phi = -10^\circ$  with the positive  $x$  axis while the stiffness along the eigenvector  $v$  is a factor of 3 smaller (see Fig. 8.5 caption for values). As expected, the maximum observed particle fluctuations are along the axis of smallest stiffness.

The spectrum of nonconservative axial force fluctuations is predicted as follows. The power spectrum of position fluctuations along each of the principal axes is calculated from Eq. 8.3, using the corresponding values of  $\alpha$ . The projections of these fluctuations on the  $x$  direction are calculated, and since the dynamics corresponding to different principal axes are independent, the projected power spectra are additive. The resulting combined spectrum is multiplied by  $S_1$  to obtain the spectrum of nonconservative axial force fluctuations. The dashed line shows this calculated spectra, which agrees precisely with the measured fluctuation spectrum of the nonconservative component of the axial scattering force.

The same considerations applying to Fig. 8.4 apply here. Due to the nonlinear nature of the equation of motion fluctuations of particle position cannot be decomposed into components arising from conservative and nonconservative forcing terms. However the additional force fluctuations introduced by the  $\rho$  dependence of the axial and non-axial components of the scattering are very small compared with the other fluctuating forces. Since there is no reason to assume that nonconservative

fluctuations will be amplified disproportionately compared to fluctuations arising from conservative forces, the nonconservative forces are not expected to cause an appreciable degradation in a measurement of the tether length.

Our general conclusion is that the method of calculating additional fluctuations associated with the nonconservative scattering force that we developed in the context of the simplified model can be applied to this experimental configuration, provided the system is linearized about the equilibrium point and that transverse dependence of the scattering force is represented by a Taylor expansion. As in the case of the simplified model, the additional fluctuations associated with the nonconservative force are small. The main effect of the scattering force is a slight reduction in the effective stiffness of the optical trap in the neighborhood of the equilibrium point.

## ***Conclusion***

In conclusion, we have introduced a simplified model of trapping with a nonconservative scattering force which fully reproduces the nonequilibrium flux reported previously, and which allows us to identify a distinct noise mechanism associated with the nonequilibrium flux. The model correctly predicts fluctuations in simulations of the simplified system, of the fully nonlinear system, and of a system including a DNA tether, provided that the transverse dependence of the scattering force is represented as a Taylor series in  $\rho$  and the effective potential experienced by the particle is linearized about the effective equilibrium point. The additional fluctuations are extremely broadband, in contrast to the extremely long time period that characterizes the net circulation rate. Although the effect is small for typical

configurations, the formulae obtained establish the dependence of this additional noise on the system parameters and can be used to guide the design of an experiment that would maximize or minimize the effect. A realistic simulation of an experiment in which the length of a DNA molecule is measured indicates that the additional RMS fluctuations associated with the scattering force in a typical experimental configuration do not significantly increase the experimental uncertainty in this common type of measurement. The most significant effect of the scattering force in such experiments is modification of the effective stiffness of the trap.

## Bibliography

- 1 Kramers, H. A. Brownian motion in a field of force and the diffusion model of chemical reactions. *Physica* **7**, 284-304 (1940).
- 2 Ashkin, A. Acceleration and trapping of particles by radiation pressure. *Phys. Rev. Lett.* **24**, 156-159 (1970).
- 3 Ashkin, A. & Dziedzic, J. M. Optical levitation by radiation pressure. *Applied Physics Letters* **19**, 283-285 (1971).
- 4 Ashkin, A., Dziedzic, J. M., Bjorkholm, J. E. & Chu, S. Observation of a single-beam gradient force optical trap for dielectric particles. *Opt. Lett.* **11**, 288-290 (1986).
- 5 Ashkin, A. & Dziedzic, J. M. Optical trapping and manipulation of viruses and bacteria. *Science* **235**, 1517-1520 (1987).
- 6 Neuman, K. C. & Block, S. M. Optical trapping. *Review of Scientific Instruments* **75**, 2787-2809 (2004).
- 7 Moffitt, J. R., Chemla, Y. R., Smith, S. B. & Bustamante, C. Recent advances in optical tweezers. *Annu. Rev. Biochem.* **77**, 205-228 (2008).
- 8 Shaevitz, J. W., Abbondanzieri, E. A., Landick, R. & Block, S. M. Backtracking by single RNA polymerase molecules observed at near-base-pair resolution. *Nature* **426**, 684-687 (2003).
- 9 Pralle, A., Prummer, M., Florin, E. L., Stelzer, E. H. K. & Hörber, J. K. H. Three-dimensional high-resolution particle tracking for optical tweezers by forward scattered light. *Microsc. Res. Tech.* **44**, 378-386 (1999).
- 10 Neuman, K. C. & Nagy, A. Single-molecule force spectroscopy: optical tweezers, magnetic tweezers and atomic force microscopy. *Nat. Methods* **5**, 491-505 (2008).
- 11 Woodside, M. T., Anthony, P. C., Behnke-Parks, W. M., Larizadeh, K., Herschlag, D. & Block, S. M. Direct measurement of the full, sequence-dependent folding landscape of a nucleic acid. *Science* **314**, 1001-1004 (2006).
- 12 Woodside, M. T., Behnke-Parks, W. M., Larizadeh, K., Travers, K., Herschlag, D. & Block, S. M. Nanomechanical measurements of the sequence-dependent folding landscapes of single nucleic acid hairpins. *Proc. Natl. Acad. Sci.* **103**, 6190-6195 (2006).

- 13 Arrhenius, S. On the rate of reaction of the inversion of sucrose by acids. *Zeitschrift für Physikalische Chemie* **4**, 226-248 (1889).
- 14 Eyring, H. The activated complex in chemical reactions. *Journal of Chemical Physics* **107**, 107-115 (1935).
- 15 Laidler, K. J. & King, M. C. Development of transition-state theory. *The Journal of Physical Chemistry* **87**, 2657-2664 (1983).
- 16 Dill, K. A., Ozkan, S. B., Shell, M. S. & Weikl, T. R. The protein folding problem. *Annual Review of Biophysics* **37**, 289-316 (2008).
- 17 Onuchic, J. N. & Wolynes, P. G. Theory of protein folding. *Curr. Opin. Struct. Biol.* **14**, 70-75 (2004).
- 18 Plotkin, S. S. & Onuchic, J. N. Understanding protein folding with energy landscape theory Part I: basic concepts. *Q. Rev. Biophys.* **35**, 111-167 (2002).
- 19 Wolynes, P. G., Onuchic, J. N. & Thirumalai, D. Navigating the folding routes. *Science* **267**, 1619-1620 (1995).
- 20 Borgia, A., Williams, P. M. & Clarke, J. Single-molecule studies of protein folding. *Annu. Rev. Biochem.* **77**, 101-125 (2008).
- 21 Evans, E. Looking inside molecular bonds at biological interfaces with dynamic force spectroscopy. *Biophys. Chem.* **82**, 83-97 (1999).
- 22 Dudko, O. K., Hummer, G. & Szabo, A. Theory, analysis, and interpretation of single-molecule force spectroscopy experiments. *Proc. Natl. Acad. Sci.* **105**, 15755-15760 (2008).
- 23 Evans, E. Probing the relation between force - lifetime - and chemistry in single molecular bonds. *Annu. Rev. Biophys. Biomol. Struct.* **30**, 105-128 (2001).
- 24 Koch, S. J. & Wang, M. D. Dynamic force spectroscopy of protein-DNA interactions by unzipping DNA. *Phys. Rev. Lett.* **91**, 028103 (2003).
- 25 Lynch, S., Baker, H., Byker, S. G., Zhou, D. & Sinniah, K. Single molecule force spectroscopy on G-quadruplex DNA. *Chemistry - A European J.* **15**, 8113-8116 (2009).
- 26 Liphardt, J., Onoa, B., Smith, S. B., Tinoco Jr., I. & Bustamante, C. Reversible unfolding of single RNA molecules by mechanical force. *Science* **292**, 733-737 (2001).

- 27 Hyeon, C., Morrison, G. & Thirumalai, D. Force-dependent hopping rates of RNA hairpins can be estimated from accurate measurement of the folding landscapes. *Proc. Natl. Acad. Sci.* **105**, 9604-9609 (2008).
- 28 Manosas, M., Wen, J.-D., Li, P. T. X., Smith, S. B., Bustamante, C., Tinoco Jr, I. & Ritort, F. Force unfolding kinetics of RNA using optical tweezers. II. Modeling experiments. *Biophys. J.* **92**, 3010-3021 (2007).
- 29 Wen, J.-D., Manosas, M., Li, P. T. X., Smith, S. B., Bustamante, C., Ritort, F. & Tinoco Jr., I. Force unfolding kinetics of RNA using optical tweezers. I. Effects of experimental variables on measured results. *Biophys. J.* **92**, 2996-3009 (2007).
- 30 Torrie, G. M. & Valleau, J. P. Nonphysical sampling distributions in Monte Carlo free-energy estimation: Umbrella sampling. *Journal of Computational Physics* **23**, 187-199 (1977).
- 31 Kumar, S., Rosenberg, J. M., Bouzida, D., Swendsen, R. H. & Kollman, P. A. The weighted histogram analysis method for free-energy calculations on biomolecules. I. The method. *Journal of Computational Chemistry* **13**, 1011-1021 (1992).
- 32 Denesyuk, N. A. & Weeks, J. D. Equilibrium and nonequilibrium effects in the collapse of a model polypeptide. *Phys. Rev. Lett.* **102**, 108101 (2009).
- 33 Greenleaf, W. J., Woodside, M. T., Abbondanzieri, E. A. & Block, S. M. Passive all-optical force clamp for high-resolution laser trapping. *Phys. Rev. Lett.* **95**, 208102 (2005).
- 34 Huguet, J. M., Bizarro, C. V., Forns, N., Smith, S. B., Bustamante, C. & Ritort, F. Single-molecule derivation of salt dependent base-pair free energies in DNA. *Proc. Natl. Acad. Sci.* **107**, 15431-15436 (2010).
- 35 Junier, I., Mossa, A., Manosas, M. & Ritort, F. Recovery of free energy branches in single molecule experiments. *Phys. Rev. Lett.* **102**, 070602 (2009).
- 36 Koch, S. J., Shundrovsky, A., Jantzen, B. C. & Wang, M. D. Probing protein-DNA interactions by unzipping a single DNA double helix. *Biophys. J.* **83**, 1098-1105 (2002).
- 37 Landry, M. P., McCall, P. M., Qi, Z. & Chemla, Y. R. Characterization of photoactivated singlet oxygen damage in single-molecule optical trap experiments. *Biophys. J.* **97**, 2128-2136 (2009).
- 38 Marko, J. F. & Siggia, E. D. Stretching DNA. *Macromolecules* **28**, 8759-8770 (1995).

- 39 Saenger, W., *Principles of Nucleic Acid Structure*. (Springer-Verlag, New York, 1988).
- 40 Wang, M. D., Yin, H., Landick, R., Gelles, J. & Block, S. M. Stretching DNA with optical tweezers. *Biophys. J.* **72**, 1335-1346 (1997).
- 41 Svoboda, K. & Block, S. M. Biological applications of optical forces. *Annu. Rev. Biophys. Biomol. Struct.* **23**, 247-285 (1994).
- 42 Peyret, N., Prediction of nucleic acid hybridization: parameters and algorithms, Thesis, Wayne State University, 2000.
- 43 SantaLucia Jr, J. A unified view of polymer, dumbbell, and oligonucleotide DNA nearest-neighbor thermodynamics. *Proc. Natl. Acad. Sci.* **95**, 1460-1465 (1998).
- 44 Zuker, M. Mfold web server for nucleic acid folding and hybridization prediction. *Nuc. Acids Res.* **31**, 3406-3415 (2003).
- 45 Crick, F. H. C., Barnett, L., Brenner, S. & Watts-Tobin, R. J. General nature of the genetic code for proteins. *Nature* **192**, 1227-1232 (1961).
- 46 Wen, J.-D., Lancaster, L., Hodges, C., Zeri, A.-C., Yoshimura, S. H., Noller, H. F., Bustamante, C. & Tinoco, I. Following translation by single ribosomes one codon at a time. *Nature* **452**, 598-603 (2008).
- 47 Tsuchihashi, Z. & Kornberg, A. Translational frameshifting generates the gamma subunit of DNA polymerase III holoenzyme. *Proc. Natl. Acad. Sci.* **87**, 2516-2520 (1990).
- 48 Farabaugh, P. J. Programmed translational frameshifting. *Annu. Rev. Genet.* **30**, 507-528 (1996).
- 49 Harger, J. W., Meskauskas, A. & Dinman, J. D. An integrated model of programmed ribosomal frameshifting. *Trends Biochem. Sci.* **27**, 448-454 (2002).
- 50 Liao, P.-Y., Choi, Y. S., Dinman, J. D. & Lee, K. H. The many paths to frameshifting: kinetic modelling and analysis of the effects of different elongation steps on programmed -1 ribosomal frameshifting. *Nuc. Acids Res.* **39**, 300-312 (2011).
- 51 Chen, G., Chang, K.-Y., Chou, M.-Y., Bustamante, C. & Tinoco, I. Triplex structures in an RNA pseudoknot enhance mechanical stability and increase



- efficiency of -1 ribosomal frameshifting. *Proc. Natl. Acad. Sci.* **106**, 12706-12711 (2009).
- 52 Jacks, T. & Varmus, H. E. Expression of the Rous sarcoma virus pol gene by ribosomal frameshifting. *Science* **230**, 1237-1242 (1985).
- 53 Dragic, T., Litwin, V., Allaway, G. P., Martin, S. R., Huang, Y., Nagashima, K. A., Cayan, C., Maddon, P. J., Koup, R. A., Moore, J. P. & Paxton, W. A. HIV-1 entry into CD4+ cells is mediated by the chemokine receptor CC-CKR-5. *Nature* **381**, 667-673 (1996).
- 54 Pleij, C. W. A., Rietveld, K. & Bosch, L. A new principle of RNA folding based on pseudoknotting. *Nuc. Acids Res.* **13**, 1717-1731 (1985).
- 55 Giedroc, D. P. & Cornish, P. V. Frameshifting RNA pseudoknots: Structure and mechanism. *Virus Res.* **139**, 193-208 (2009).
- 56 Namy, O., Moran, S. J., Stuart, D. I., Gilbert, R. J. C. & Brierley, I. A mechanical explanation of RNA pseudoknot function in programmed ribosomal frameshifting. *Nature* **441**, 244-247 (2006).
- 57 Belew, A. T., Sulima, S. O., Meskauskas, A., Musalgaonkar, S. & Dinman, J. D. A programmed -1 ribosomal frameshift signal in the human CCR5 mRNA operates as an mRNA destabilizing element through the nonsense-mediated mRNA decay pathway and is regulated by miR-1224. *In preparation* (2012).
- 58 Bartel, D. P. MicroRNAs: Genomics, biogenesis, mechanism, and function. *Cell* **116**, 281-297 (2004).
- 59 Jacks, T., Power, M. D., Masiarz, F. R., Luciw, P. A., Barr, P. J. & Varmus, H. E. Characterization of ribosomal frameshifting in HIV-1 gag-pol expression. *Nature* **331**, 280-283 (1988).
- 60 Dinman, J. D. & Wickner, R. B. Ribosomal frameshifting efficiency and gag/gag-pol ratio are critical for yeast M1 double-stranded RNA virus propagation. *J. Virol.* **66**, 3669-3676 (1992).
- 61 Biswas, P., Jiang, X., Pacchia, A. L., Dougherty, J. P. & Peltz, S. W. The human immunodeficiency virus type 1 ribosomal frameshifting site is an invariant sequence determinant and an important target for antiviral therapy. *J. Virol.* **78**, 2082-2087 (2004).
- 62 Brierley, I., Pennell, S. & Gilbert, R. J. C. Viral RNA pseudoknots: versatile motifs in gene expression and replication. *Nat. Rev. Micro.* **5**, 598-610 (2007).

- 63 Plant, E. P., Rakauskaitė, R., Taylor, D. R. & Dinman, J. D. Achieving a golden mean: Mechanisms by which coronaviruses ensure synthesis of the correct stoichiometric ratios of viral proteins. *J. Virol.* **84**, 4330-4340 (2010).
- 64 Lopinski, J. D., Dinman, J. D. & Bruenn, J. A. Kinetics of ribosomal pausing during programmed -1 translational frameshifting. *Mol. Cell. Biol.* **20**, 1095-1103 (2000).
- 65 Belew, A. T., Hepler, N. L., Jacobs, J. L. & Dinman, J. D. PRFdb: A database of computationally predicted eukaryotic programmed -1 ribosomal frameshift signals. *BMC Genomics* **9**, 1-7 (2008).
- 66 Brierley, I. & Dos Ramos, F. J. Programmed ribosomal frameshifting in HIV-1 and the SARS-CoV. *Virus Res.* **119**, 29-42 (2006).
- 67 Green, L., Kim, C.-H., Bustamante, C. & Tinoco Jr, I. Characterization of the mechanical unfolding of RNA pseudoknots. *J. Mol. Biol.* **375**, 511-528 (2008).
- 68 Biyun, S., Cho, S. S. & Thirumalai, D. Folding of human telomerase RNA pseudoknot using ion-jump and temperature-quench simulations. *J. Amer. Chem. Soc.* **133**, 20634-20643 (2011).
- 69 Hansen, T. M., Reihani, S. N. S., Oddershede, L. B. & Sørensen, M. A. Correlation between mechanical strength of messenger RNA pseudoknots and ribosomal frameshifting. *Proc. Natl. Acad. Sci.* **104**, 5830-5835 (2007).
- 70 Yu, C.-H., Noteborn, M. H., Pleij, C. W. A. & Olsthoorn, R. C. L. Stem loop structures can effectively substitute for an RNA pseudoknot in -1 ribosomal frameshifting. *Nuc. Acids Res.* **39**, 8952-8959 (2011).
- 71 Tholstrup, J., Oddershede, L. B. & Sørensen, M. A. mRNA pseudoknot structures can act as ribosomal roadblocks. *Nuc. Acids Res.* **40**, 303-313 (2012).
- 72 Pallan, P. S., Marshall, W. S., Harp, J., III, F. C. J., Wawrzak, Z., II, B. A. B., Rich, A. & Egli, M. Crystal structure of a luteoviral RNA pseudoknot and model for a minimal ribosomal frameshifting motif. *Biochemistry* **44**, 11315-11322 (2005).
- 73 de Messieres, M., Brawn-Cinani, B. & La Porta, A. Measuring the folding landscape of a harmonically constrained biopolymer. *Biophys. J.* **100**, 2736-2744 (2011).

- 74 Tinoco, I., Borer, P. N., Dengler, B., Levine, M. D., Uhlenbeck, O. C., Crothers, D. M. & Gralla, J. Improved estimation of secondary structure in ribonucleic acids. *Nature New Biology* **246**, 40-41 (1973).
- 75 Xia, T., SantaLucia, J., Burkard, M. E., Kierzek, R., Schroeder, S. J., Jiao, X., Cox, C. & Turner, D. H. Thermodynamic Parameters for an Expanded Nearest-Neighbor Model for Formation of RNA Duplexes with Watson-Crick Base Pairs. *Biochemistry* **37**, 14719-14735 (1998).
- 76 Li, G.-M. Mechanisms and functions of DNA mismatch repair. *Cell Res* **18**, 85-98 (2008).
- 77 Fersht, A. R., Knill-Jones, J. W. & Tsui, W.-C. Kinetic basis of spontaneous mutation: Misinsertion frequencies, proofreading specificities and cost of proofreading by DNA polymerases of *Escherichia coli*. *J. Mol. Biol.*, 37-51 (1982).
- 78 Ebel, S., Lane, A. N. & Brown, T. Very stable mismatch duplexes: Structural and thermodynamic studies on tandem G\*A mismatches in DNA. *Biochemistry* **31**, 12083-12086 (1992).
- 79 Kan, L. S., Chandrasegaran, S., Pulford, S. M. & Miller, P. S. Detection of a guanine X adenine base pair in a deoxyribonucleotide by proton magnetic resonance spectroscopy. *Proc. Natl. Acad. Sci.* **80**, 4263-4265 (1983).
- 80 SantaLucia, J., Allawi, H. T. & Seneviratne, P. A. Improved nearest-neighbor parameters for predicting DNA duplex stability. *Biochemistry* **35**, 3555-3562 (1996).
- 81 Freier, S. M., Kierzek, R., Jaeger, J. A., Sugimoto, N., Caruthers, M. H., Neilson, T. & Turner, D. H. Improved free-energy parameters for predictions of RNA duplex stability. *Proc. Natl. Acad. Sci.* **83**, 9373-9377 (1986).
- 82 Crothers, D. M. & Zimm, B. H. Theory of the melting transition of synthetic polynucleotides: Evaluation of the stacking free energy. *J. Mol. Biol.* **9**, 1-9 (1964).
- 83 Gellert, M., Lipsett, M. N. & Davies, D. R. Helix formation by guanylic acid. *Proc. Natl. Acad. Sci.* **48**, 2013-2018 (1962).
- 84 Gilbert, D. E. & Feigon, J. Multistranded DNA structures. *Curr. Opin. Struct. Biol.* **9**, 305-314 (1999).
- 85 Huppert, J. L. Four-stranded nucleic acids: structure, function and targeting of G-quadruplexes. *Chem. Soc. Rev.* **37**, 1375-1384 (2008).

- 86 Sen, D. & Gilbert, W. Formation of parallel four-stranded complexes by guanine-rich motifs in DNA and its implications for meiosis. *Nature* **334**, 364-366 (1988).
- 87 Shirude, P. S., Okumus, B., Ying, L., Ha, T. & Balasubramanian, S. Single-molecule conformational analysis of G-quadruplex formation in the promoter DNA duplex of the proto-oncogene C-Kit. *J. Am. Chem. Soc.* **129**, 7484-7485 (2007).
- 88 Smith, F. W. & Feigon, J. Strand orientation in the DNA quadruplex formed from the Oxytricha telomere repeat oligonucleotide d(G4T4G4) in solution. *Biochemistry* **32**, 8682-8692 (1993).
- 89 Lipps, H. J. & Rhodes, D. G-quadruplex structures: in vivo evidence and function. *Trends Cell Biol.* **19**, 414-422 (2009).
- 90 Huppert, J. L. & Balasubramanian, S. G-quadruplexes in promoters throughout the human genome. *Nuc. Acids Res.* **35**, 406-413 (2007).
- 91 Balasubramanian, S., Hurley, L. H. & Neidle, S. Targeting G-quadruplexes in gene promoters: a novel anticancer strategy? *Nat. Rev. Drug Disc.* **10**, 261-275 (2011).
- 92 Neidle, S. & Parkinson, G. Telomere maintenance as a target for anticancer drug discovery. *Nat. Rev. Drug Disc.* **1**, 383-393 (2002).
- 93 Dhakal, S., Schonhoft, J. D., Koirala, D., Yu, Z., Basu, S. & Mao, H. Coexistence of an ILPR i-Motif and a partially folded structure with comparable mechanical stability revealed at the single-molecule level. *J. Am. Chem. Soc.* **132**, 8991-8997 (2010).
- 94 Wanrooij, P. H., Uhler, J. P., Simonsson, T., Falkenberg, M. & Gustafsson, C. M. G-quadruplex structures in RNA stimulate mitochondrial transcription termination and primer formation. *Proc. Natl. Acad. Sci.* **107**, 16072-16077 (2010).
- 95 Yu, Z., Schonhoft, J. D., Dhakal, S., Bajracharya, R., Hegde, R., Basu, S. & Mao, H. ILPR G-quadruplexes formed in seconds demonstrate high mechanical stabilities. *J. Am. Chem. Soc.* **131**, 1876-1882 (2009).
- 96 Schonhoft, J. D., Das, A., Achamyelah, F., Samdani, S., Sewell, A., Mao, H. & Basu, S. ILPR repeats adopt diverse G-quadruplex conformations that determine insulin binding. *Biopolymers* **93**, 21-31 (2010).
- 97 Hummer, G. & Szabo, A. Kinetics from nonequilibrium single-molecule pulling experiments. *Biophys. J.* **85**, 5-15 (2003).

- <sup>98</sup> Dudko, O. K., Filippov, A. E., Klafter, J. & Urbakh, M. Beyond the conventional description of dynamic force spectroscopy of adhesion bonds. *Proc. Natl. Acad. Sci.* **100**, 11378-11381 (2003).
- <sup>99</sup> Dudko, O. K., Hummer, G. & Szabo, A. Intrinsic rates and activation free energies from single-molecule pulling experiments. *Phys. Rev. Lett.* **96**, 108101 (2006).
- <sup>100</sup> Jarzynski, C. Nonequilibrium equality for free energy differences. *Phys. Rev. Lett.* **78**, 2690-2693 (1997).
- <sup>101</sup> Schonhofs, J. D., Bajracharya, R., Dhakal, S., Yu, Z., Mao, H. & Basu, S. Direct experimental evidence for quadruplex-quadruplex interaction within the human ILPR. *Nuc. Acids Res.* **37**, 3310-3320 (2009).
- <sup>102</sup> Dudko, O. K., Mathé, J., Szabo, A., Meller, A. & Hummer, G. Extracting kinetics from single-molecule force spectroscopy: nanopore unzipping of DNA hairpins. *Biophys. J.* **92**, 4188-4195 (2007).
- <sup>103</sup> Schwarz, G. Estimating the Dimension of a Model. *The Annals of Statistics* **6**, 461-464 (1978).
- <sup>104</sup> Hyeon, C. & Thirumalai, D. Measuring the energy landscape roughness and the transition state location of biomolecules using single molecule mechanical unfolding experiments. *J. of Phys. Cond. Matt.* **19**, 113101 (2007).
- <sup>105</sup> Liu, J.-Q., Chen, C.-Y., Xue, Y., Hao, Y.-H. & Tan, Z. G-quadruplex hinders translocation of BLM helicase on DNA: A real-time fluorescence spectroscopic unwinding study and comparison with duplex substrates. *Journal of American Chemical Society* **132**, 10521–10527 (2010).
- <sup>106</sup> Shim, J. W., Tan, Q. & Gu, L.-Q. Single-molecule detection of folding and unfolding of the G-quadruplex aptamer in a nanopore nanocavity. *Nuc. Acids Res.* **37**, 972-982 (2009).
- <sup>107</sup> de Messieres, M., Denesyuk, N. A. & La Porta, A. Noise associated with nonconservative forces in optical traps. *Physical Review E* **84**, 031108 (2011).
- <sup>108</sup> Denk, W. & Webb, W. W. Optical measurement of picometer displacements of transparent microscopic objects. *Appl. Opt.* **29**, 2382-2391 (1990).
- <sup>109</sup> Rohrbach, A. & Stelzer, E. H. K. Three-dimensional position detection of optically trapped dielectric particles. *Journal of Applied Physics* **91**, 5474-5488 (2002).

- <sup>110</sup> Abbondanzieri, E. A., Greenleaf, W. J., Shaevitz, J. W., Landick, R. & Block, S. M. Direct observation of base-pair stepping by RNA polymerase. *Nature* **438**, 460-465 (2005).
- <sup>111</sup> Ashkin, A. Forces of a single-beam gradient laser trap on a dielectric sphere in the ray optics regime. *Biophys. J.* **61**, 569-582 (1992).
- <sup>112</sup> Roichman, Y., Sun, B., Stolarski, A. & Grier, D. G. Influence of nonconservative optical forces on the dynamics of optically trapped colloidal spheres: The fountain of probability. *Phys. Rev. Lett.* **101**, 128301 (2008).
- <sup>113</sup> Sun, B. & Grier, D. G. Response to Huang, Wu and Florin, "Comment on Influence of non-conservative optical forces on the dynamics of optically trapped colloidal spheres: The fountain of probability" [arxiv.org/abs/0807.1242](http://arxiv.org/abs/0807.1242) (2008).
- <sup>114</sup> Yariv, A., *Optical Electronics, 3rd ed.* (John Wiley & Sons, New York, 1988).