

University of California
Santa Barbara

Electrostatics of single-stranded nucleic acids: Force spectroscopy and ion counting experiments

A dissertation submitted in partial satisfaction
of the requirements for the degree

Doctor of Philosophy
in
Physics

by

David R. Jacobson

Committee in charge:

Professor Omar A. Saleh, Co-Chair
Professor Philip A. Pincus, Co-Chair
Professor Everett A. Lipman

December 2016

The Dissertation of David R. Jacobson is approved.

Professor Everett A. Lipman

Professor Philip A. Pincus, Committee Co-Chair

Professor Omar A. Saleh, Committee Co-Chair

October 2016

Electrostatics of single-stranded nucleic acids: Force spectroscopy and ion counting
experiments

Copyright © 2016

by

David R. Jacobson

Acknowledgements

None of the work described in this dissertation would have been possible without the mentorship of my research advisor, Omar Saleh, who set the broad, initial course of my research and has, along the way, provided guidance, resources, and invaluable discussion.

I have also benefitted greatly from my day-to-day collaborations with the other members of the Saleh laboratory, with whom I have shared and discussed new ideas and collaborated to maintain the infrastructure of the lab. I am particularly indebted to Dustin McIntosh, who was my predecessor in working on the force spectroscopy of single-stranded nucleic acids and who taught me, at the beginning of my time in the lab, how to carry out and interpret those experiments.

I have received generous financial support throughout my graduate career, both in terms of stipend and research funding as well as having the opportunity to travel to conferences and summer schools around the world. I would like to particularly acknowledge the National Science Foundation, the International Center for Materials Research, and the Dow Chemical Company.

Finally, I would like to specifically acknowledge my co-authors and collaborators on each of the research projects discussed in this dissertation. Low-force elasticity of single-stranded RNA (Chapter 3): Andrew Dittmore, Dustin McIntosh, and Omar Saleh. Intermediate-force elasticity of single-stranded nucleic acids (Chapter 4): Dustin McIntosh, Philip Pincus, Michael Rubinstein, Mark Stevens, and Omar Saleh. Review of prior ion atmosphere studies (Chapter 5): David Draper, Alexander Grosberg, Changbong Hyeon, Philip Pincus, Omar Saleh, and Joan-Emma Shea. Ion atmosphere of single-stranded nucleic acids (Chapter 6): Rachel Behrens, David Draper, Alexander Grosberg, Luc Jaeger, Dan Nguyen, Philip Pincus, Omar Saleh, and Amanda Stromm. Extended theory of single-molecule ion counting experiments (Appendix D): David Draper, Luc

Jaeger, Jonathan Landy, and Omar Saleh. Force calibration for molecules exhibiting conformational switching (Appendix E): Sarah Innes-Gold, Bob Lansdorp, and Omar Saleh.

Curriculum Vitæ

David R. Jacobson

Education

- 2016 Ph.D. in Physics, University of California, Santa Barbara.
2014 M.A. in Physics, University of California, Santa Barbara.
2011 B.A. *cum laude* with Distinction in Physics and with Distinction in Biochemistry, University of Pennsylvania.

Honors and Awards

- 2017–2019 NIST NRC Postdoctoral Research Associateship
2013–2016 NSF Graduate Research Fellowship
2011 Chair’s Award, Biochemistry Program, University of Pennsylvania
2007–2011 Roy and Diana Vagelos Molecular Life Science Program, University of Pennsylvania.

Publications

D.R. Jacobson, O.A. Saleh, “Magnetic tweezers force calibration for molecules that exhibit conformational switching”, *Rev. Sci. Instrum.* **87**, 094302 (2016).

D.R. Jacobson, O.A. Saleh, “Quantifying the ion atmosphere of unfolded, single-stranded nucleic acids using equilibrium dialysis and single-molecule methods”, *Nucleic Acids Res.* **44**, 3763–3771 (2016).

C.-Y. Park, D.R. Jacobson, D.T. Nguyen, S. Willardson, O.A. Saleh, “A thin permeable-membrane device for single-molecule manipulation”, *Rev. Sci. Instrum.* **87**, 014301 (2016).

D.R. Jacobson, O.A. Saleh, “Measuring the differential stoichiometry and energetics of ligand binding to macromolecules by single-molecule force spectroscopy: An extended theory”, *J. Phys. Chem. B* **119**, 1930–1938 (2015).

D.R. Jacobson, D.B. McIntosh, O.A. Saleh, “The snakelike chain character of unstructured RNA”, *Biophys. J.* **105**, 2569–2576 (2013).

J.L. Sievers, *et al.*, “The Atacama Cosmology Telescope: Cosmological parameters from three seasons of data”, *J. Cosmol. Astropart. P.* **10**, 060 (2013).

L. Laureano-Perez, R. Collé, D.R. Jacobson, R. Fitzgerald, N.S. Kahn, I.J. Dmochowski, “A novel application for ^{222}Rn emanation standards: Radon-cryptophane host chemistry”, *Appl. Radiat. Isotopes* **70**, 1997–2001 (2012).

D.R. Jacobson, N.S. Kahn, R. Collé, R. Fitzgerald, L. Laureano-Pérez, Y. Bai, I.J. Dmochowski, “Measurement of radon and xenon binding to a cryptophane molecular host”, *Proc. Natl. Acad. Sci. U.S.A.* **108**, 10969–10973 (2011).

Professional Appointments

2013–2016	Graduate research fellow, NSF Graduate Research Fellowship Program, University of California, Santa Barbara
2012–2013, 2016	Graduate student researcher, Materials Department, University of California, Santa Barbara
2011–2012	Teaching assistant, Department of Physics, University of California, Santa Barbara
2009–2011	Undergraduate researcher, Roy and Diana Vagelos Molecular Life Science Program and Department of Chemistry, University of Pennsylvania
2008	Undergraduate researcher, Provost’s Undergraduate Research Mentorship Program, University of Pennsylvania

Fields of Study

2012–2016	Single-molecule biophysics and polymer physics. <i>Supervisor:</i> Prof. Omar Saleh, Materials Department, University of California, Santa Barbara
2009–2011	Physical chemistry and biomedical imaging. <i>Supervisor:</i> Prof. Ivan Dmochowski, Department of Chemistry, University of Pennsylvania
2008	Observational cosmology. <i>Supervisor:</i> Prof. Mark Devlin, Department of Physics and Astronomy, University of Pennsylvania

Abstract

Electrostatics of single-stranded nucleic acids: Force spectroscopy and ion counting experiments

by

David R. Jacobson

The strong negative charge of nucleic acids adds additional complexity to their behavior, beyond what would be expected for neutral polymers, through the addition of long-range self-repulsive interactions and through the screening of those interactions by ions in solution. Among the techniques that can probe the electrostatic behavior of a charged polymer are single-molecule force spectroscopy, which reports on the elasticity of the polymer (i.e., extension as a function of applied force), and ion counting, which reports on the stoichiometry of positive and negative ions constituting the polymer’s “ion atmosphere”. While extensive studies of both types have been made of double-stranded DNA, a strongly charged and moderately stiff polymer, less is known about the single-stranded nucleic acids (ssNAs), which are also strongly charged but are comparatively flexible.

In my dissertation, I use single-molecule magnetic tweezers to measure the elasticity of ssRNA as a function of monovalent and divalent salt concentration under 0.1–10 pN applied force and observe behavior characteristic of flexible polyelectrolytes. I then extend the measurements up to 100 pN to probe the intermediate-force domain connecting behavior dominated by electrostatics and by the chemical structure of the backbone. These new data allow the critical testing of a wormlike chain-derived model treating electrostatics through a mean-field, salt-dependent tension. This model is shown to quantitatively account for the data and to bridge the established behaviors at lower and higher force. I also quantify the ion atmosphere stoichiometry of both ssDNA and ssRNA using

three complementary experimental techniques, including both bulk and single-molecule approaches. These results are compared with Poisson-Boltzmann models incorporating varying degrees of structural complexity, demonstrating the inherently short-ranged (i.e., sub-Debye screening length) nature of nucleic acid-ion interactions.

Contents

Curriculum Vitae	vi
Abstract	viii
1 Introduction	1
1.1 Charge and polymer behavior	1
1.2 Single-stranded nucleic acids: Biologically significant model flexible poly-electrolytes	3
1.3 Single-molecule force spectroscopy	7
1.4 Ion counting experiments	10
1.5 Structure of the dissertation	13
2 Elasticity of single-stranded nucleic acids: Theory and existing results	14
2.1 Scaling theory of uncharged polymers	14
2.2 Theories describing polyelectrolyte effects	19
2.3 Elasticity measurements	22
2.3.1 Double-stranded nucleic acids	23
2.3.2 Single-stranded nucleic acids	25
3 Low-force elasticity of single-stranded RNA	29
3.1 Results	31
3.1.1 Elasticity measurements	31
3.1.2 Regime crossover analysis	33
3.1.3 Quantification of associated ions	35
3.2 Discussion	38
3.2.1 Elastic response	38
3.2.2 Associated ion excess	40
3.2.3 Comparing ssRNA and ssDNA	41
3.3 Conclusion	42

4	Intermediate-force elasticity: Experiments and mean-field model	45
4.1	Results	47
4.1.1	Elasticity measurements	47
4.1.2	Force-extension model	47
4.1.3	Electrostatic tension	51
4.1.4	Three-parameter global fit	52
4.2	Discussion	52
4.2.1	Force-independence of effective charge spacing	52
4.2.2	Salt-independence and magnitude of effective charge spacing	54
4.2.3	Recovery of bare persistence length	55
4.3	Regimes of elastic behavior	56
5	The nucleic acid ion atmosphere: Theory and existing results	60
5.1	Quantifying the ion atmosphere	61
5.1.1	Three-dimensional distribution	61
5.1.2	Ion counting	62
5.2	Theoretical models	65
5.2.1	Limiting behavior: Counterion condensation and Debye-Hückel theories	66
5.2.2	Poisson-Boltzmann theory	67
5.2.3	Simulations	70
5.3	Ion counting experiments	70
5.3.1	Absolute ion atmosphere stoichiometry	72
5.3.2	Competition between ions	73
5.3.3	Changes with denaturation	75
5.4	Discussion	78
6	Measurements of the ssNA ion atmosphere	80
6.1	Results	82
6.1.1	Coil ion excess measured by D-AES	82
6.1.2	PB calculations	84
6.1.3	Single-molecule measurement of $\Delta\Gamma_{\text{ds-stretch}}$	86
6.1.4	Completing the thermodynamic cycle	88
6.2	Discussion	89
6.2.1	Uniform mean-field model	90
6.2.2	Non-uniform mean-field models	92
6.3	Conclusion	94
7	Outlook	96

A	Materials and methods	101
A.1	Sample preparation	101
A.1.1	Long homopolymeric ssNAs	101
A.1.2	Short ssRNA/ssDNA oligonucleotides for D-AES	102
A.1.3	RNA hairpin molecular construct	102
A.2	Magnetic tweezers experiments	105
A.2.1	Flow cell preparation	109
A.2.2	Instrumentation	110
A.2.3	Force calibration	112
A.2.4	Elasticity measurements	117
A.2.5	Equilibrium folding/unfolding studies	118
A.2.6	Data analysis	118
A.3	Dialysis/atomic emission spectroscopy	122
A.3.1	Procedure	122
A.3.2	AES calibration	123
A.4	Poisson-Boltzmann calculations	124
A.5	Analysis of duplex thermal denaturation data	127
B	Hairpin $\Delta\Gamma$ control experiments	128
B.1	Hairpin control experiments	128
B.2	Agreement between $\Delta\Gamma_{\text{ds-stretch}}$ methods	131
B.3	High-salt linear limiting behavior of Γ^{molar}	132
C	Derivation of the uniform mean-field model	134
D	Theory of differential ion counting measurements	139
D.1	Equilibrium stoichiometry measurements	141
D.1.1	Thermodynamic description of the isotensional experiment	142
D.1.2	Ligand binding relation for a continuously deformable macromolecule (Maxwell relation)	145
D.1.3	Ligand binding relations for a macromolecule inhabiting discrete conformational states	147
D.1.4	Thermodynamic description of the isometric experiment	150
D.2	Non-equilibrium measurements	151
D.3	Extraction of free energies of ligand binding	154
D.4	$\Delta\Delta G$ from RNA and DNA data	157
D.4.1	ssRNA, ssDNA and dsDNA	157
D.4.2	DNA hairpin	161
E	Force calibration for molecules that exhibit conformational switching	164
E.1	Theory	167
E.1.1	Effect of conformational switching on measures of bead fluctuations	167
E.1.2	Effect of two-state switching kinetics	169

E.2	Simulations	170
E.2.1	Validation of predicted effects	171
E.2.2	Force calibration bias	171
E.3	Sample experiment	174
E.4	Conclusion	176
F	Notation used	178
F.1	Abbreviations used	178
F.2	Mathematical symbols used	179
	Bibliography	184

Chapter 1

Introduction

1.1 Charge and polymer behavior

Polymer physics is concerned with the study of long, flexible objects; in particular, chain-like molecules made up of many monomers chemically linked one to the next [1].

The development of polymer physics occurred throughout the twentieth century—motivated, to a considerable extent, by the desire to explain the thermomechanical behavior of rubber [2]—and was marked by a series of theoretical milestones [3]. Among these were Staudinger’s hypothesis (1920) that *fadenmoleküle* (fiber molecules) “consist of very long molecules, in which a hundred or more single molecules are bound by means of main valencies into chains,” [4,5] Kuhn’s theory (1934) on the size of an ideal polymer chain in solution [6], Flory’s theory (1949) on the size of “real” (i.e., self-avoiding) polymer chains in good solution [7], and Kratky and Porod’s theory (1949) of a continuously deformable (wormlike) chain [8]. This theoretical edifice built upon the gradual accumulation of experimental knowledge of polymer systems, beginning with natural rubber and extending to synthetic systems such as polystyrene and polyethylene glycol [1].

By the end of the twentieth century, a firm understanding had emerged of the prop-

erties of uncharged, chain-like molecules, and three Nobel prizes had been awarded in the area: to Staudinger in 1953, to Flory in 1974, and to de Gennes in 1991. Charged polymers (polyelectrolytes), however, remain a different matter. For example, de Gennes wrote in *Scaling Concepts in Polymer Physics* (1979) that:

“The effects of Coulomb repulsions between the ionized groups are strong and complex. From a scaling point of view, the properties of polyelectrolyte solutions are not yet fully understood.” [1]

While progress has certainly been made since that time, the understanding of polyelectrolyte physics remains incomplete (see Chapter 2).

In the canonical picture of polymer scaling, different behaviors are separated by different characteristic length scales. These length scales include the overall size of the polymer, the “thermal blob” size (described in Section 2.1), and the statistical monomer size (Kuhn length) [3, 9]. The presence of charge on the polymer introduces a new length scale characteristic of electrostatic self-repulsion. This interaction is extremely long-range, falling off only as $1/r$ in vacuum and thus interfering with all of the other characteristic length scales. Biologically relevant polyelectrolyte studies, however, are performed in solutions containing some amount of salt, dissociated into positive and negative ions. These ions have the effect of screening the electrostatic potential of the polyelectrolyte and causing it to fall off more rapidly with distance. Thus, in solution the characteristic length scale of electrostatic interactions is salt concentration dependent.

An additional complication is that it is not the bulk salt concentration that is responsible for the charge screening, but rather the local concentration in the vicinity of the polyelectrolyte; these two concentrations are not the same, since the salt ions are free to move in solution, with ions of one charge being attracted to the polyelectrolyte and ions of the other charge being repelled. The region surrounding the polyelectrolyte where the

ion concentration is perturbed from its bulk value is known as the *ion atmosphere*.

The study of polyelectrolytes must thus consider many additional degrees of complexity on top of the well-established theories of uncharged polymer physics. Some of this additional complexity can be dealt with by simply-stated analytical theories, while the rest can only be handled, to date, using complicated computer simulations.

1.2 Single-stranded nucleic acids: Biologically significant model flexible polyelectrolytes

Many important molecules in biology are electrically charged, including proteins, polysaccharides, certain lipids, and the nucleic acids (DNA and RNA) [10]. In an evolutionary sense, charge provides these macromolecules with a means of long-range attractive or repulsive interaction that can be tuned by pH, via the protonation or de-protonation of certain chemical groups, and by the bulk salt concentration via charge screening. DNA provides an example of the tunable nature of the electrostatic interactions. Whereas at low-to-moderate salt concentrations, DNA occurs as an easily accessible, self-repulsive, straight chain, in the presence of +3 or higher valence ions, it condenses into dense, self-attractive, chromatin-like conformations [11].

The nucleic acids (NAs) consist of a sugar-phosphate polymer backbone with attached nitrogenous bases (Figure 1.1). Because they can each have one of several chemical compositions (denoted A, T, C, and G for DNA; A, U, C, and G for RNA), these bases accomplish the NA's role in the storage of genetic information. Double-stranded NAs form by hydrogen bonding between the bases on two separate strands.

Nucleic acids are responsible for the storage of genetic information in cells. The “central dogma” of molecular biology holds that double-stranded DNA (dsDNA) is the

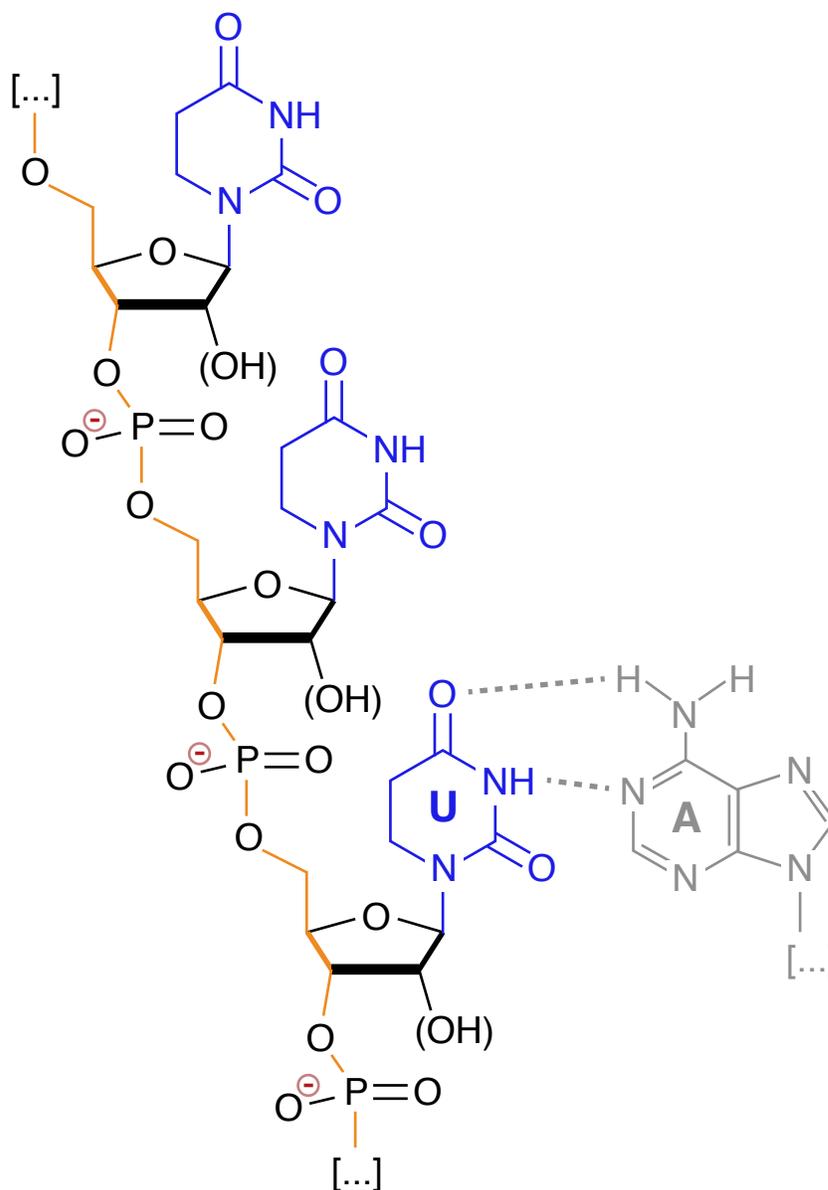


Figure 1.1: Structural diagram of three nucleotides along the single-stranded NA backbone. Each nucleotide consists of the charge-bearing phosphate group, the sugar ring, and the base (blue). The backbone, which is under tension in force spectroscopy experiments, is highlighted in orange. The 2' hydroxyl group of RNA, shown in parentheses, is replaced by a hydrogen in DNA. Shown is a polyuridine chain like that studied in Chapter 3. A Watson-Crick base pair between one uridine (U) and an adenine (A) is shown in gray as an example of the interaction that would stabilize a two-stranded double-helix.

repository of genetic information on long time scales, and that this information is then transcribed into single-stranded RNA (ssRNA) for short-term transfer to the ribosome, where it is then translated into protein [12]. The incredible complexity of, for example, human life is reflected not in the total number of proteins encoded by the genome, of which there are only about 20,000–25,000 [13], but in the myriad ways in which the expression of those proteins is regulated across different types of cells and across different time-points in the cell cycle and in the development of the organism. This regulation occurs both at the level of the DNA, through structural organization achieved by binding to proteins such as histones [14] and through transcriptional regulation achieved by promoter/repressor binding [15] and interaction with polymerases [16,17], and at the level of the RNA, through self-catalytic “ribozyme” activity [18] and through translational regulation including “riboswitch” activity [19]. Electrostatics play a particularly important role in understanding the riboswitch and ribozyme activities of RNA. A ribozyme, for example, is a stable, three-dimensional structure—including considerable double-stranded character—that folds spontaneously from an initially single-stranded RNA sequence; the intrinsic electrostatic self-repulsion of the negatively charged RNA must be overcome in order for this to happen.

The negative charge of nucleic acids is considerable, as seen by inspecting the values of structural charge spacing, b , listed in Table 1.1. For the ssNAs, these values represent the charge spacings along the backbone of the polymer when fully extended; for the dsNAs, they represent the average spacing between charges projected onto the helical axis. These charge spacing values can be contextualized by comparison with the Bjerrum length:

$$l_B \equiv \frac{e^2}{4\pi\epsilon_0\epsilon k_B T}, \quad (1.1)$$

where e is the elementary charge, ϵ_0 is the permittivity of free space, ϵ is the dielectric

Table 1.1: Electrostatic and structural parameters of the single- and double-stranded nucleic acids. Reported are bare (i.e., non-electrostatic) persistence lengths, l_p , measured in the high-salt limit where electrostatic interactions are screened away. A dsRNA l_p measurement of 63.8 ± 0.7 nm has been made at 10 mM ionic strength by Abels *et al.* [27], but I am not aware of any studies conducted in the high-salt limit.

species	b [23] (nm)	l_p (nm)
dsDNA	0.17	35–50 [24]
dsRNA	0.12	—
ssDNA	0.70	0.60 [25]
ssRNA	0.59	0.83 [26]

constant, and $k_B T$ is the thermal energy. The Bjerrum length is the characteristic length scale over which two elementary charges will have Coulomb interaction energy equal to the thermal energy, and has a value of approximately 0.7 nm in water at 25°C [20]. Thus, we see that all of the nucleic acids are “strongly” charged, in that $b \leq l_B$ in all cases, with the dsNAs being the more strongly charged. As will be further explained in Section 5.2.1, the phenomenon of counterion condensation [21, 22] renders the structural value of b in many cases unimportant once it exceeds the critical value of l_B , and so even though the ssNAs and dsNAs differ considerably in their charge density, they can be expected to behave, from an electrostatic point of view, similarly.

Another way in which the single- and double-stranded nucleic acids differ, however, is in their stiffness, which is listed in Table 1.1 in the form of the persistence length, l_p . The persistence length is a property of the wormlike chain (WLC) model of polymer elasticity [8, 28], which will be further discussed in Section 2.1. Briefly, this model treats the polymer as a continuously deformable rod; the correlation between vectors tangent to the polymer is taken to decrease exponentially as a function of separation along the backbone, and l_p is the length scale of this exponential decay. Thus, l_p can be thought of as the length scale over which the polymer is “stiff” against thermal fluctuations. The values of l_p given in Table 1.1 are “bare” persistence lengths; i.e., persistence lengths

measured under sufficiently high salt concentrations such that all of the electrostatic interactions, which will further stiffen the chain, are screened away. Inspection of these values reveals that the dsNAs (dsRNA is expected to behave similarly to dsDNA based on the the study of Abels *et al.* [27]; although, good high-salt data are not available) are roughly two orders of magnitude stiffer than the ssNAs. The dsNAs are therefore said to be “semi-flexible” (persistence length much larger than the physical monomer size), whereas the ssNAs are said to be “flexible” (persistence length comparable to the physical monomer size).

1.3 Single-molecule force spectroscopy

In this dissertation, I will discuss the use of two major classes of experiments to study the electrostatic properties of ssNAs. In the first of these, single-molecule force spectroscopy (SMFS) is used to measure ssNA elasticity.

There are several types of widely adopted SMFS methodologies, including atomic force microscopy (AFM), optical trapping (OT), and magnetic tweezers (MT) [29]. The general scheme of all of these experiments is the same, and is illustrated in Figure 1.2A. The biomolecule under study is prepared so as to have chemical functionalization at both ends; this functionalization is used to bind the molecule, at one end, to a solid surface (a glass coverslip, in the case of MT) and, at the other end, to a force transducer. In AFM this transducer is a micro-fabricated cantilever, in OT it is a dielectric sphere suspended in a focused laser beam, and in MT it is a super-paramagnetic sphere in a non-uniform magnetic field.

In all three modalities, a force, f , is applied, which will tend to elongate the molecule. It is thus the extension, X , of the molecule that is read out of the experiment, as a function of f or time. Such measurements, referred to as *force-extension curves* or *force spectra*,

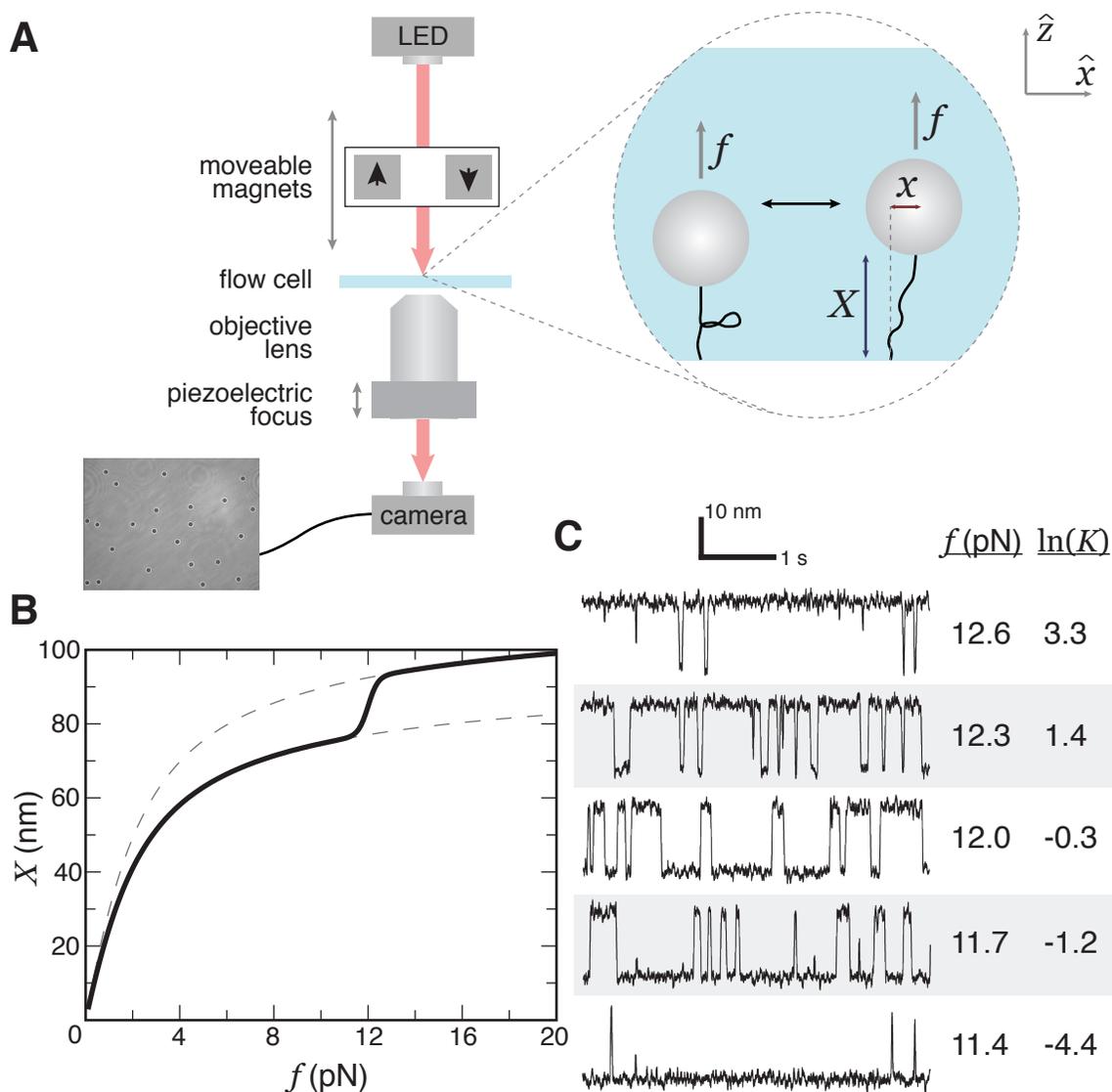


Figure 1.2: (A) Schematic representation of an MT SMFS experiment. A single biomolecule is chemically modified at its ends to facilitate attachment to a substrate, at one end, and to the paramagnetic bead, at the other. This bead applies a force to the molecule, leading to its deformation. The extension of the molecule is measured as a function of the applied force and of time. (B, C) Simulated data for a molecule exhibiting a two-state folding/unfolding transition. (B) Force-extension curve. The dashed lines indicate WLC elasticity for both the folded and unfolded states, separately; the solid line applies to the simulated molecule, which experiences a structural transition at about 12 pN. (C) Constant-force trajectories recorded as a function of time. As the force is varied, the relative amount of time spent in each of the two states changes; this can be quantified by the equilibrium constant, K . Adapted from Refs. [30] and [31].

constitute a measurement of the molecular elasticity.

The behavior of the molecule as a function of applied force (as recorded in the force-extension curve) can be related to the spatial conformation of that molecule through the tensile screening length: [32]

$$\xi \equiv \frac{k_B T}{f}. \quad (1.2)$$

The basic idea of the tensile screening length is illustrated in Figure 1.3. In the absence of applied force, an unstructured polymer will exist as a random coil in solution (see Section 2.1) in which any two monomers in the chain can interact with each other. However, as force is applied, this coil is deformed until, at high enough force, it becomes a straight chain and only adjacent monomers interact. The free energy associated with stretching any part of the chain by length X is fX . Since the chain is at equilibrium with a thermal bath, it fluctuates with characteristic energy $k_B T$. Thus, elements of the chain for which $fX < k_B T$ can still interact, whereas those for which $fX > k_B T$ can not. Therefore, ξ is the characteristic length scale, for a particular force, beyond which monomer-monomer interactions are screened away by the applied force [32]. That is to say that features in the force-extension curve (e.g., Figure 1.2B) occurring at high force will correspond to short-length-scale features in the chain conformation, and features in the force-extension curve occurring at low force will correspond to long-length-scale features in the chain conformation.

In another type of SMFS experiment, the force is held at a constant value and the extension of the molecule is monitored as a function of time. Constant force is a natural consequence of the MT setup, but must be maintained using feedback loops in AFM and OT. Certain structured molecules, like a folded DNA hairpin, will exhibit equilibrium folding/unfolding behavior, as illustrated in Figure 1.2C, when the applied force is tuned to the correct value. The equilibrium constant, K , describing the multi-state equilibrium

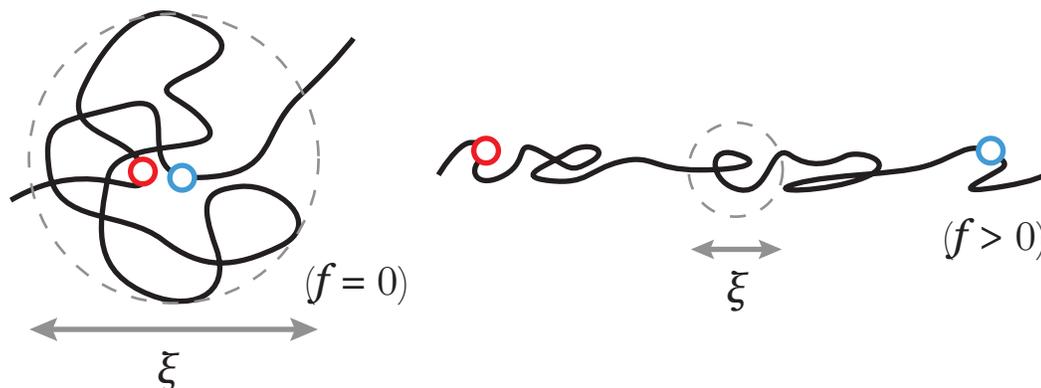


Figure 1.3: Illustration of the tensile screening length, $\xi = k_B T / f$. In the absence of force, polymers occur as random coils in solution, and so any two monomers are able to interact with each other. As force is applied, the coil is deformed and monomers that are sufficiently far apart along the backbone no longer interact. For a particular force, ξ is the length scale separating interacting and non-interacting monomers.

can be read out, as a function of force, by comparing the relative amounts of time spent occupying each state.

1.4 Ion counting experiments

In addition to observing the effect of electrostatics on the elasticity of ssNAs, I will also discuss experiments that directly probe the interaction of ssNAs with ions in solution by measuring the stoichiometry of the ion atmosphere. As mentioned above, electrostatic screening of NAs occurs not just due to the bulk salt concentration, but also due to the perturbed local salt concentration arising from attractive interactions between the NA and cations and repulsive interactions between the NA and anions. While this phenomenon is most fully described by specifying the ion concentration as a function of spatial position—and it is such a quantity that is often the output of theoretical models—in experimental practice it is only the net total number of excess cations or excluded anions that can be measured. This quantity is known as the *preferential ion*

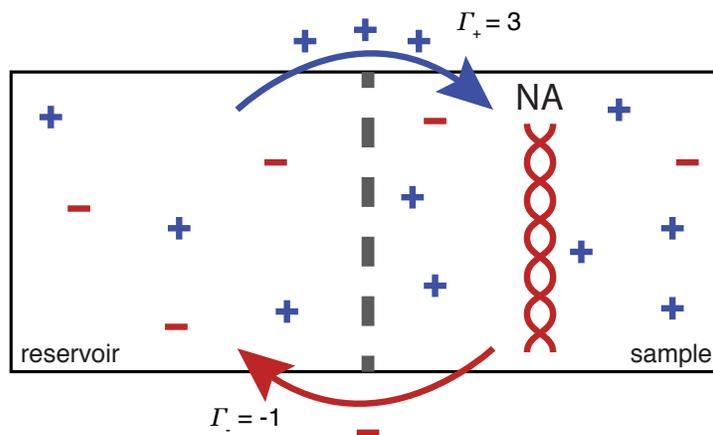


Figure 1.4: Simple diagram of a Donnan dialysis experiment, in which two compartments, one containing the NA sample and the other a reservoir with a certain bulk salt concentration, are separated by a membrane permeable to mobile ions but impermeable to the NA. Preferential ion interactions with the NA cause an imbalance in the ion concentrations in each compartment, which can be measured and related to Γ .

interaction coefficient or *ion excess*, Γ , and will be formally defined in Section 5.1.2. Succinctly, Γ can be thought of as the number of excess ions of a particular species surrounding the NA above the number that would have been expected due to the bulk salt concentration alone.

The ion excess is traditionally measured through dialysis experiments taking advantage of the phenomenon of Donnan equilibrium [33]. The basic setup of such an experiment is sketched in Figure 1.4. The experiment consists of two compartments, one containing the NA sample and the other a large reservoir with a particular bulk salt concentration, which are separated by a membrane permeable to the ions but impermeable to the NA. The ions are in grand canonical equilibrium across the membrane, and their preferential interactions with the NA will lead to an imbalance in ion concentration between the two compartments. These concentrations can be read out, often using atomic emission spectroscopy, and the values of the preferential ion interaction coefficients determined [34–38].

The power of ion counting experiments is that they can directly test theoretical models of nucleic acid electrostatics, from the most simple (e.g., counterion condensation theory) all the way to the most complex (e.g., fully atomistic molecular dynamics simulations), because these theories generally make predictions about the composition of the ion atmosphere. As mentioned above, however, these theories often report full spatial distributions of ion concentration, a more information-rich description than the single number (i.e., Γ) resulting from an ion counting experiment. As such, it must be remembered that agreement of a theory with the ion counting results is a necessary but insufficient criterion for the validation of that theory. In other words, many possible ion distributions could give rise to the same value of the ion excess.

Some additional information can be obtained through experiments that are not equally sensitive to all of the ions in the ion atmosphere. Examples, which will be discussed further in Section 5.2.3, include competition experiments, in which one ion species is titrated into an NA solution containing another while the excess of each ion species is monitored, and differential ion counting experiments, in which the change in ion excess is measured as an NA transitions between conformational states. Competition experiments are most sensitive to the ions closest to the NA, where high local concentrations render the differences between ion species most pronounced, and denaturation experiments are most sensitive to those ions that stabilize the NA structure. Results from both types of experiments can further constrain theory, by providing additional information beyond absolute stoichiometry, but still do not contain all the information of a full three-dimensional distribution.

1.5 Structure of the dissertation

In the chapters that follow, I will discuss my work towards understanding the electrostatics of ssNAs using both elasticity and ion counting measurements. Chapters 2–4 deal with the elasticity work, first reviewing the existing theoretical and experimental findings (Chapter 2) before moving on to my experiments at low (Chapter 3) and intermediate (Chapter 4) forces. A similar structure is followed in the second half of the dissertation concerning measurements of the ion atmosphere: I first review the existing theory and experiment (Chapter 5) before presenting my new results (Chapter 6). There follows, in Chapter 7, a conclusion discussing further questions raised by this research.

Following the main body of the dissertation are several appendices. Appendices A–C give descriptions of materials and methods, additional control experiments, and derivations to support the results presented in the main chapters. Appendix D formalizes and extends the theory of single-molecule differential ion counting experiments. Appendix E uses theoretical and simulation approaches to assess the degree to which two-state conformational switching—as in the hairpin experiments of Chapter 6—biases magnetic tweezer force calibration. A comprehensive listing of abbreviations and mathematical symbols used is given in Appendix F.

Chapter 2

Elasticity of single-stranded nucleic acids: Theory and existing results

Elasticity measurements, when interpreted in terms of the tensile screening length, contain information about the conformation of a polymer: elastic behaviors that arise at higher forces correspond to structural features on shorter length scales. In the well-established theory of uncharged polymer physics, characteristic elastic regimes are separated by length scales corresponding to the important structural parameters of the polymer. The addition of charge, as in nucleic acids, complicates this picture through the addition of competing length scales that vary as a function of salt concentration. Experimental measurements of NA elasticity therefore provide a fertile ground for testing theories aimed at describing this richer behavior.

2.1 Scaling theory of uncharged polymers

Scaling theories—theories describing the power-law dependence of one parameter on another—are often employed in polymer physics as a way of extracting universal behavior

that does not depend on the chemical details of the particular polymer under study. These theories, valid for chains in the infinite-length limit, have a formal correspondence with the theories of critical phenomena, in the limit of small perturbations away from the critical temperature [1]. The canonical scaling theories of polymer physics describe the physical extent of the polymer (often the root-mean-square end-to-end extension, or Flory radius, R_F) as a function of the number of monomers; these are not linearly related since, to maximize entropy, the polymer will exist as a random coil, and not as a straight chain, in solution. These scaling laws are given by: [1]

$$R_F \sim N^{1/2} \text{ (ideal)} \quad (2.1)$$

$$R_F \sim N^\nu \text{ (real)}, \quad (2.2)$$

where N is the degree of polymerization (i.e., number of monomers) and $\nu \approx 3/5$ in three dimensions. In the correspondence with critical phenomena, R_F is dual with the magnetization correlation length and ν is dual with the critical exponent of ferromagnetic order [1]. An *ideal* chain refers to one that does not interact with itself and therefore exists as a simple random walk through space; a *real* chain refers to one that does interact with itself and, because of excluded volume, instead exists as a self-avoiding random walk. It is for this reason that the real result, unlike the ideal case, depends upon the dimensionality of space: the fewer dimensions there are, the more constrained is the polymer and the more likely it is to bump into itself. For example, in one dimension $\nu = 1$; i.e., the constraint that the polymer not cross itself forces it into a straight line.

Most of the studies I will discuss will be of polymers in the real regime. This corresponds to a polymer in a good solvent; that is to say, a solvent in which the polymer has stronger affinity for the solvent than for itself. The opposite is a poor solvent, in which the polymer will tend to aggregate, rendering experimental study by the methods I will

discuss impractical. Between these two regimes, the solvent is said to be at a Θ point and ideal behavior is seen even for polymers that would ordinarily be self-avoiding [3]. This ideal behavior occurs not because the polymer becomes non-interacting, but because the interactions tending towards swelling and collapse are equally balanced.

The tensile screening length, $\xi = k_B T / f$, allows a connection to be drawn from scaling laws connecting R_F and N to scaling laws connecting X and f [32]. This is possible because ξ sets the length scale over which monomers can thermally interact in the presence of force. Thus, the portion of the polymer within this region, known as a “tensile blob”, essentially acts as an independent polymer in its own right with $R_F = \xi$ and N given by the appropriate scaling law (e.g., Equations 2.1 or 2.2). Using this argument, the $X(f)$ scaling behavior expected for an uncharged polymer can be obtained; these results are shown in Figure 2.1 [9]. Note that these results account only for entropic elasticity. Chain elasticity can also arise from enthalpic (i.e., spring-like) stretching of the backbone, which is typically modeled by the addition of a stretch modulus (e.g., 500–1500 pN for dsDNA [39], 8400 pN for ssDNA [40]).

In good solvent (Figure 2.1A), five force-extension scaling regimes are expected. At the lowest forces, corresponding to $\xi > R_F$, the force slightly perturbs the otherwise random-coil conformation of the polymer, and the extension scales linearly with force: $X \sim f$ [9]. As the force increases, the tensile screening length divides the chain into tensile blobs, oriented on long length scales but exhibiting real chain statistics within the blobs. This manifests as an $X \sim f^{2/3}$ (Pincus) scaling law [9,32]. Such excluded-volume-dominated behavior persists until $\xi \approx \xi_t$, where ξ_t is the thermal blob size, at which point the linear $X \sim f$ law again occurs. This linear behavior arises because, as the force increases and the tensile screening length correspondingly decreases, the number of monomers within each tensile blob becomes so small that the excluded volume interaction energy becomes less than $k_B T$, leading to ideal chain statistics within the blobs [3,9]. The

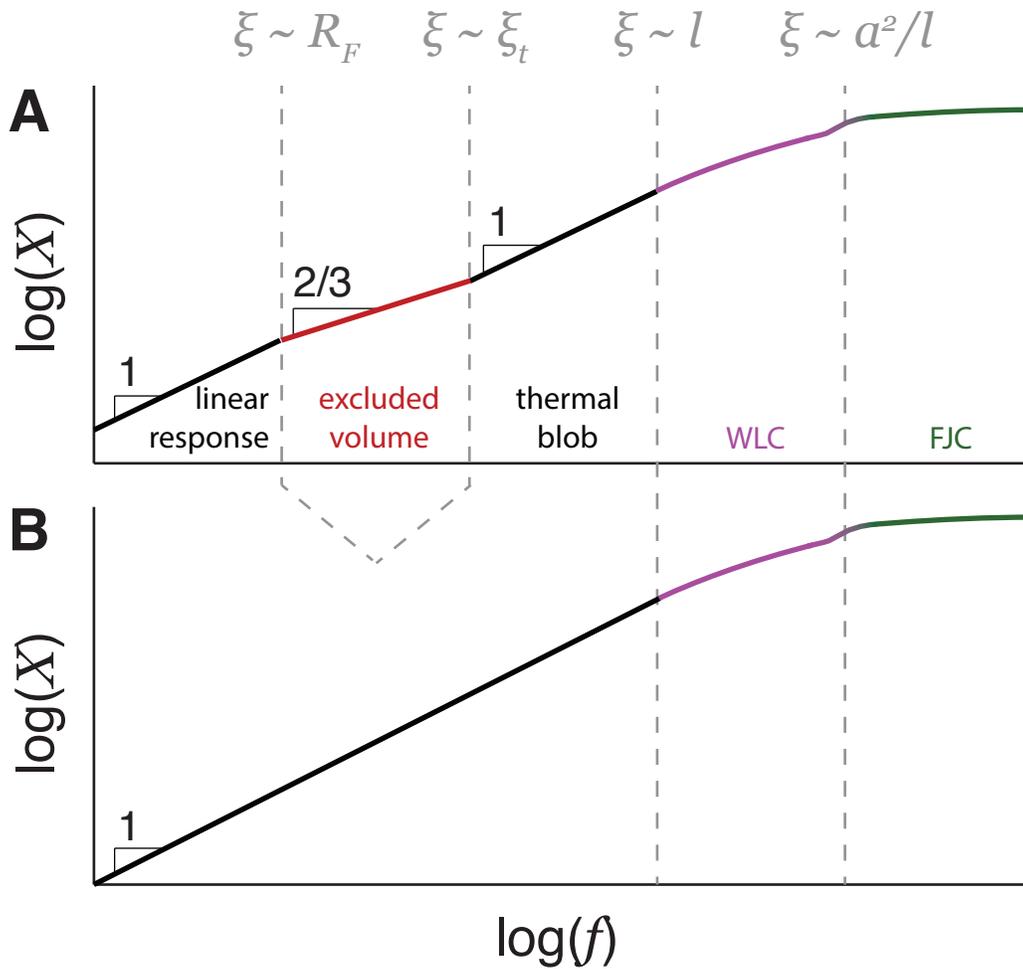


Figure 2.1: Expected entropic scaling behavior for an uncharged polymer in (A) good and (B) Θ solvent. Because the Pincus regime arises from excluded volume interactions, it is absent in the Θ case, where the excluded volume parameter is zero. Also noted are the characteristic forces—related to length scales through the tensile screening length—that correspond to the transitions between scaling regimes. Based upon Figure 1 of McIntosh *et al.* [9].

thermal blob size is given, approximately, by $\xi_t \sim l^4/v$, where l is the statistical (Kuhn) monomer size and v is the monomer excluded volume [3, 9]. Thermal blob scaling is expected to persist in the range $l < \xi < \xi_t$. At higher force, the chain enters a regime where ξ is smaller than the monomer size, aligning the chain completely with the applied force, modulo fluctuations, and leading to an asymptotic approach to the chain contour length. The exact nature of this approach depends on how the chain is modeled on short length scales. If the chain is treated as a sequence of rigid monomers linked by flexible joints, a “freely jointed chain” (FJC) model, the approach goes as $(L_c - X) \sim f^{-1}$, where L_c is the chain contour length. However, if the chain is modeled as a continuously deformable rod (i.e., allowing for deformation within the monomers—the WLC model), the approach goes as $(L_c - X) \sim f^{-1/2}$. The latter is generally found to be in better agreement with the elastic behavior of double- and single-stranded nucleic acids in the MT force range [26–28, 39, 41–43]; however, a transition from WLC to FJC behavior is expected at $\xi = a^2/l$, where a is the effective backbone bond length [44, 45].

In Θ solvent (Figure 2.1B), the scaling picture is simplified by the lack of excluded volume interactions and, therefore, the lack of the Pincus regime. Instead, the linear-response and thermal blob regimes run together into a broad, low-force $X \sim f$ regime, which gives way to the the same asymptotic high-force behavior when $\xi < l$. This means that, at Θ , the entire force-extension curve up to the WLC-FJC transition can be fit by the Marko-Siggia WLC interpolation formula: [28]

$$f = \frac{k_B T}{l_p} \left[\frac{1}{4} \left(1 - \frac{X}{L_c} \right)^{-2} - \frac{1}{4} + \frac{X}{L_c} \right], \quad (2.3)$$

which accounts for WLC behavior at high force and linear scaling at low force, with a smooth and experimentally justified interpolation in between. Here l_p is the chain persistence length discussed in Chapter 1. If the force range under study were to include

the WLC-FJC transition, the alternate interpolation formula of Dobrynin *et al.* [44] should be used instead.

2.2 Theories describing polyelectrolyte effects

In our discussion of uncharged chains, monomer-monomer interactions entered through the effects of excluded volume, a local hard-sphere interaction. Now, however, we consider charged chains that are self-repulsive on longer length scales. In particular, in the presence of a bulk salt concentration—as is the case in biology—the extent of this repulsion will be roughly the Debye screening length:

$$\kappa^{-1} = 1/\sqrt{4\pi l_B I} \quad (2.4)$$

which is the characteristic length scale of the exponential decay of the electrostatic Yukawa potential

$$\Phi(r) = \frac{Q}{4\pi\epsilon\epsilon_0} \frac{1}{r} e^{-r\kappa}, \quad (2.5)$$

valid in the Debye-Hückel (weak potential) limit of the Poisson-Boltzmann equation [46]. In these equations I is the ionic strength* and Q is the charge of the object in question. In the absence of bulk salt, the exponential of Equation 2.5 goes to unity and the electrostatic repulsion takes on the long-ranged $1/r$ dependence.

The effects of electrostatic repulsion on polyelectrolyte conformation can be first understood in the limits of bulk salt concentration, c_{bulk} . At high c_{bulk} , the salt screens away all of the electrostatic interactions (i.e., the exponential of Equation 2.5 goes to zero) and the polymer behaves as if it is uncharged. That is to say, in the limit of high salt,

* $I = \sum_{i=1}^{N_s} c_i Z_i^2$, where N_s is the number of ion species in solution, c_i is the concentration of a particular species, and Z_i is the integer valence of that species.

the same scaling laws of Equations 2.1 and 2.2 are still expected to hold. The opposite limit, of no added salt, is treated by the de Gennes-Pincus-Velasco-Brochard (dGPVB) theory [47], which finds that the unscreened electrostatic repulsion imposes a directional axis on the polyelectrolyte: the polymer is stretched into a straight chain on long length scales, but maintains random-walk statistics, albeit with additional electrostatic excluded volume, on short length scales. That is to say, the polymer adopts a “chain of blobs” configuration, just as in the case of an applied force. Because the polymer is oriented on long length scales in the dGPVB picture, its end-to-end separation scales linearly with the degree of polymerization: $R_F \sim N$.

The chain of blobs nature of the dGPVB model, in the limit of no added salt, motivates a more general understanding of the effect of charge on polymer conformation in terms of electrostatic tension at intermediate c_{bulk} . For a long straight chain with charge separation b in the Debye-Hückel limit, Netz [48] derived, by explicit summation, an expression for the electrostatic tension, f_{el} , as a function of Debye length (which encodes c_{bulk}):

$$f_{\text{el}} = \frac{k_B T l_B}{b^2} \left(\kappa b \frac{e^{-\kappa b}}{1 - e^{-\kappa b}} - \ln(1 - e^{-\kappa b}) \right). \quad (2.6)$$

A similar expression, further accounting for counterion condensation, was also derived by Manning [49]. Interpretation of polyelectrolyte elasticity in terms of this tension will be further discussed in Section 2.3.2 and Chapter 4. The expectation in this view is that the force-extension behavior of a polyelectrolyte will be the same as that of an uncharged polymer, except that the true force applied to the molecule should be reckoned as the sum of the applied force and the electrostatic tension.

The Netz electrostatic tension (Equation 2.6) has been directly incorporated into the concept of persistence length by Manning [49], who used the internal electrostatic tension and the phenomenon of counterion condensation to determine a relationship between the

persistence length of a chain at a particular salt concentration, $l_{p,\text{el}}$, as a function of the bare persistence length in the absence of charge, $l_{p,0}$. The prediction of this Manning-Netz (MN) theory is given by:

$$l_{p,\text{el}} = \left(\frac{\pi}{2}l_{p,0}\right)^{2/3} \frac{R^{4/3}}{Z^2l_B} \left[\left(\frac{2Zl_B}{b} - 1\right) \frac{\kappa b e^{-\kappa b}}{1 - e^{-\kappa b}} - 1 - \ln(1 - e^{-\kappa b}) \right], \quad (2.7)$$

where R is the radius of the cylinder approximating the chain.

A separate, but related, way of treating the elasticity of polyelectrolytes at reasonably high applied force is to think of the electrostatic repulsion not as generating an additional electrostatic tension, but as leading to an additional electrostatic stiffness. In the theories of Odijk [50] and Skolnick and Fixman [51] (OSF), later extended by Barrat and Joanny [52], the chain's persistence length is written as a sum of bare and electrostatic terms: $l_p = l_{p,0} + l_{p,\text{el}}$. The bare term, $l_{p,0}$, depends on the chemical structure of the polymer and does not vary with c_{bulk} . The electrostatic term, $l_{p,\text{el}}$ arises from the tendency of the charges to resist backbone bending, since such bending would bring the charges closer together and increase their Coulomb energy. Also in the Debye-Hückel limit—at high force and for semi-flexible chains; i.e, when the bending can be treated as a perturbation to an otherwise straight structure—OSF theory predicts that:

$$l_{p,\text{el}} = \frac{l_B}{4b^2\kappa^2}. \quad (2.8)$$

This electrostatic persistence length can be inserted into the WLC model to give predicted high-force force-extension curves for polyelectrolytes [52].

Both methods of calculating an electrostatic persistence length (Equations 2.7 and 2.8) necessarily assume that l_p is a well-defined quantity in the system under study; i.e., that there is an exponential decay of tangent vector correlations. However, non-

exponential decays have been predicted from theory and simulation studies of flexible chains [53].

To summarize, the addition of charge to a polymer introduces electrostatic repulsion that tends to elongate the chain on long length scales in a salt-dependent manner. The long-length-scale behavior can be modeled quantitatively through an electrostatic tension—a concept that will be further developed in this dissertation—or an electrostatic persistence length. On short length scales, the system is still expected to behave as a real chain, as described in the prior section, but with additional excluded volume arising from the electrostatics. The crossover between these two regimes may be complicated by the competition between the Kuhn length, which separates them for an uncharged chain, and the Debye length, which is the characteristic length scale of electrostatics and varies as a function of c_{bulk} .

2.3 Elasticity measurements

The elasticity of a polymer can be directly studied by applying a force and measuring the resulting molecular extension, via the single-molecule force spectroscopy (SMFS) techniques discussed in Section 1.3. The force-extension curves measured in these experiments can directly test the predicted elasticity scaling laws, such as those summarized in Figure 2.1, provided that the experimental technique is sensitive to the appropriate force range.

A good example of such studies, for uncharged polymers, is that of Dittmore *et al.* [54], performed on polyethylene glycol. This study successfully observed both the high-force WLC and low-force linear regimes in Θ solvent (as in Figure 2.1B) and the high-force WLC and intermediate-force thermal blob and Pincus blob regimes in good solvent (as in Figure 2.1A). The low-force linear regime in good solvent and the high-force WLC-FJC

transition went undetected, since they are expected to occur outside of the force range probed. This study concretely establishes the basic scaling behavior of uncharged chains that serves as a starting point for understanding more complex polyelectrolyte elasticity.

Below, I review that major force spectroscopy results relevant to the nucleic acids.

2.3.1 Double-stranded nucleic acids

Double-stranded DNA is semi-flexible, meaning that its bare persistence length is much larger than the monomer size, an effect arising from its rigid double-helical structure. Because of this large persistence length, the transition from thermal blob ($X \sim f$) to Pincus blob ($X \sim f^{2/3}$) scaling (Figure 2.1A) is expected to occur at inaccessibly low forces, since $\xi_t \sim l_p^4$. Thus, over the experimentally accessible force range, dsDNA is expected to behave as if it were a polymer in Θ solvent, characterized by the scaling behavior of Figure 2.1B.

This expected behavior is consistent with what is seen in the experimental force spectroscopy results of Bustamante and co-workers, first using an early magnetic tweezers setup [28, 41] and later using optical tweezers [39]. In particular, they measure elasticity over the force range 0.04–40 pN that is well-fit by the WLC interpolation formula (Equation 2.3) and, going up to 65 pN, that is well-fit by the WLC with an added stretch modulus term. Above 65 pN, dsDNA exhibits a sudden lengthening termed “overstretching”, which is a structural transition rather than an entropically driven polymer phenomenon [55–62]. Several conclusions can be drawn from the good agreement between these WLC models and the data. First, the good agreement with the WLC model and the comparatively poor agreement with the FJC model [28] indicates that the asymptotic behavior of dsDNA, up to 65 pN, is best modeled by as the straightening of a continuously deformable rod. That is to say, the dsDNA can be bent on length

scales shorter than the Kuhn length, a conclusion that makes sense given that each Kuhn length in dsDNA contains ~ 100 structural monomers (i.e., base pairs). Second, that the WLC interpolation formula, which encodes only ideal (i.e., $X \sim f$) low-force behavior, continues to fit the data down to sub-pN forces indicates that no excluded volume effects are manifesting within the probed force range, consistent with the high stiffness, and correspondingly large ξ_t , of dsDNA. Finally, the need to introduce a stretch modulus to maintain the quality of the WLC fit at the highest forces indicates that the double-helix can be deformed by sufficiently high stress. This deformability arises from the polymer's compact (i.e., contour length much shorter than length of sugar-phosphate backbone), double-helical structure, which is held together by weak base stacking and base-pairing interactions.

More recent magnetic tweezers experiments on dsRNA have confirmed essentially the same behavior: that the RNA force-extension curve is well-fit by the WLC interpolation formula over the force range of 0.01–10 pN [27]. These studies did not go to high enough forces to address the stretch modulus of dsRNA. The study did identify a slightly larger bare persistence length of dsRNA compared with dsDNA, not surprising given their different helical structures: the more compact A-form in the case of RNA, the more elongated B-form in the case of DNA.

The polyelectrolyte character of dsDNA is seen in the salt concentration-dependence of its persistence length, which was found by Baumann *et al.* [39] to be in qualitative agreement with the predictions of OSF theory, in that the persistence length decreases as a function of c_{bulk} and asymptotes to a salt-independent, bare value at high salt. This bare value does, however, vary with the ion valence: $l_{p,0} \rightarrow 50$ nm in monovalent salt, but $l_{p,0} \rightarrow 25$ nm in trivalent cobalt hexamine [39]. This likely arises from the Debye-Hückel assumption underlying OSF, an assumption that breaks down for strongly interacting multivalent ions. Furthermore, I say that the Baumann *et al.* agreement with OSF is

only qualitative, because, as Manning observes, a numerical pre-factor must be used that is not physically justified [49].

The situation is further confused when one considers a broader range of experiments, including studies other than force spectroscopy, as surveyed by Savelyev [24]. A comparison with this large dataset reveals that the characteristic OSF behavior (i.e., asymptotic approach to bare persistence length in the limit of large c_{bulk}) is only clearly seen in a few of the surveyed studies; in many others, the total persistence length continues to fall as the salt concentration increases well above 1 M, in better agreement with MN theory (Equation 2.7).

Thus, while the general elastic behavior of dsDNA is well-settled experimentally (i.e., ideal polymer behavior at low force, WLC behavior at high force giving way to an overstretching transition), the detailed salt-dependence of the elasticity is still at topic of both experimental and theoretical debate.

2.3.2 Single-stranded nucleic acids

Because it does not occur in a double-helical structure, single-stranded DNA is much less stiff than dsDNA and is considered to be a flexible polymer on the basis that its persistence length is comparable to the structural monomer size. Since the Kuhn length of ssDNA is smaller (~ 1 nm), the characteristic force separating the thermal blob and Pincus regimes is expected to be higher, and the $X \sim f^{2/3}$ scaling behavior becomes accessible at forces that can be attained using magnetic tweezers instruments.

Elasticity measurements on samples that were either natively single-stranded—due to an appropriately chosen sequence; e.g., poly(U)—or chemically denatured by treatment with glyoxal, formamide, or formaldehyde were first reported, for ssDNA, by Smith *et al.* [55]. Later studies were performed on ssDNA by Rief *et al.* [63], Dessinges *et al.* [64],

and Saleh and co-workers [25, 43] and on ssRNA by Seol *et al.* [42]. These studies span a range of applied forces from roughly 0.04 pN, using MT, to 1000 pN, using AFM. Only in the MT range were comprehensive measurements reported as a function of ionic strength [25, 43].

At the lowest forces probed in these experiments, the $X \sim f^{2/3}$ Pincus behavior is clearly seen. The presence of this regime means that it is not generally appropriate to fit the data with the WLC interpolation formula (Equation 2.3), since that model does not account for this excluded volume effect. Persistence lengths can be obtained under solution Θ conditions, where the ideal behavior of Figure 2.1B is observed. This Θ condition is satisfied by raising the bulk salt concentration surrounding the ssDNA to about 3.5 M, meaning that the persistence length obtained under these conditions can be thought of as the bare persistence length in the absence of electrostatic effects that have been screened away by the salt. The persistence length at Θ is found to be 0.60 ± 0.02 nm for ssDNA in NaCl solution, with very similar values found in MgCl₂ and CaCl₂ [9]. As I will show in Chapter 3, this is lower than the 0.83 nm persistence length of ssRNA at Θ [26].

As the bulk salt concentration is decreased away from Θ , Pincus scaling begins to appear at low force. Interestingly, no intermediate domain of linear scaling is preserved. That is to say, in good solvent conditions, the ssDNA force-extension curve transitions directly from the Pincus regime to high force behavior, without passing through the thermal blob regime. This effect was explained by McIntosh *et al.* [9] in terms of the interplay between electrostatics and the characteristic length scales separating the regimes. Recall from Section 2.1 that the characteristic tensile screening length separating the Pincus and thermal blob regimes is the thermal blob size, $\xi_t \sim l^4/\nu$, and that the length scale separating the thermal blob and high-force regimes is the Kuhn length, l . For an uncharged chain, the monomers are essentially cylindrical, meaning that the excluded volume goes

as the square of the Kuhn length and $\xi_t^{\text{uncharged}} \sim l^2$. This means that, for an uncharged chain, $\xi_t > l$ and the thermal blob regime will have non-negligible width. However, in the case of a polyelectrolyte the excluded volume is dominated by electrostatic repulsion, which is essentially spherical (i.e., isotropic). This means that, for charged chains the excluded volume goes as the Kuhn length cubed and $\xi_t^{\text{charge}} \sim l$. Thus, to a scaling factor, $\xi_t \approx l$ and no thermal blob regime will occur, as seen in experiment.

Disagreement with WLC predictions are also seen, in the good solvent regime, at high force. In particular, in monovalent salt high-force $X \sim \log(f)$ scaling is observed, corresponding to greater elasticity than that expected for a WLC. In divalent salt, even greater compliance manifested. This behavior, which was not seen in the flexible-but-uncharged PEG system [54], was assumed to be an effect unique to flexible polyelectrolytes and was reproduced in simulations of such systems [53, 65]. Recent simulations have also indicated that such scaling can arise for non-electrostatic excluded volume effects [66]. It has been postulated the this logarithmic behavior may arise from the polymer having power-law, rather than exponential, decay of its tangent vector correlations on short length scales [53]. An alternative, but not necessarily contradictory, explanation of this effect in terms of the internal electrostatic tension of the polyelectrolyte will be given, along with higher-force data probing further into this regime, in Chapter 4.

Even though the ssNA persistence length becomes ill-defined away from Θ due to the inapplicability of the WLC interpolation formula, its scaling with salt concentration can be inferred from the dependence of the crossover force, f_c (separating the $X \sim f^{2/3}$ and $X \sim \log(f)$ regimes), on c_{bulk} . Since this transition is expected to occur at forces corresponding to $\xi \sim l$, where $l = 2l_p$, the persistence length is related to the crossover force scaling through $l \sim f_c^{-1}$. Performing such an analysis revealed that, in monovalent salt, $l \sim c_{\text{bulk}}^2 \sim \kappa^{-1}$, in contrast with the $l \sim \kappa^{-2}$ scaling predicted by OSF theory (Equation 2.8) but in agreement with Barrat and Joanny's modification to OSF theory

to account for conformational fluctuations in flexible chains [52]. Much stronger scaling of l with c_{bulk} in divalent salt was also observed, but has not been quantitatively explained [25, 26].

Thus, to summarize, elasticity measurements of unstructured ssNAs [25, 42, 43, 64] reproduce the expected polymer scaling behavior in good solvent (Figure 2.1A), except for several perturbations owing to electrostatics. First, the thermal blob regime is not seen. Second, the high-force behavior in monovalent salt shows $X \sim \log(f)$ scaling and in divalent salt shows even greater compliance. Third, the Kuhn length is salt-concentration-dependent and shows $l \sim \kappa^{-1}$ scaling.

The above results hold for the unstructured ssNAs: those that behave as non-interacting random walks in solution. Certain ssNA sequences, such as poly(A) in particular, instead exhibit base stacking and will form ordered helical structures that are locally stiffer. SMFS studies of such sequences in DNA [67, 68] and RNA [69] reveal a structural transition—akin to the overstretching transition in dsDNA—corresponding to the unstacking of these bases at a salt-dependent force around 23 pN. McIntosh *et al.* [68] posited that the total intramolecular tension required to trigger the unstacking is actually salt-independent and that this total tension is a sum of the applied force, from the magnetic bead, and a mean-field electrostatic tension arising from charge-charge repulsion. By modeling f_{el} using Equation 2.6, they were able to reproduce the experimental salt-dependence of the unstacking force using a model with just two free parameters: the total intramolecular tension of unstacking and the effective charge spacing. The fitted value of b needed to give good agreement with the data was 0.9 nm, larger than the physical charge spacing of ssDNA; this discrepancy may arise from the Debye-Hückel approximation underlying Equation 2.6.

Chapter 3

Low-force elasticity of single-stranded RNA

Portions of this chapter are adapted with permission from Ref. [26]. Copyright 2013, The Biophysical Society.

Once thought to act as a mere messenger within the central dogma of molecular biology [12], RNA is now known to fill a diversity of roles, including as an enzyme [18] and genetic regulator [19]. These functions depend on the folded, three-dimensional structure of the RNA [70]. In addition to the effects of secondary (base pairing) and tertiary structural elements, folding is strongly ion-dependent [71, 72]. Ions, which can interact specifically by binding to the RNA or nonspecifically through condensation [21] and the formation of an ion “atmosphere”, serve to neutralize the strong, negative charge of the backbone. It is this screening that allows charged monomers to come into close contact, facilitating stable folding. Thus, to fully understand the folding behavior of RNA, we must understand the coupling of electrostatics and conformation.

In the absence of base-pairing and tertiary structure, RNA assumes a random-walk

conformation, modulated by the electrostatic self-repulsion of the charged, flexible backbone. This is essentially the same behavior seen in ssDNA (Section 2.3.2). Such behavior has traditionally been modeled as a WLC with an electrostatics-dependent persistence length. Studies of ssDNA have shown, however, the inadequacy of such an approach. In particular, ssDNA was seen to exhibit a self-avoiding excluded volume scaling regime ($X \sim f^{2/3}$) at low force and an unexpected logarithmic scaling regime at higher forces. Through measurements of the end-to-end extension of poly(U) RNA under 0.1–10 pN applied force (detailed materials and methods information in Appendix A), we extend these prior ssDNA studies to consider the structurally similar ssRNA. We observe the same general scaling behavior: a low-force, power-law regime characteristic of a chain of swollen blobs on long length scales and a high-force, salt valence-dependent regime consistent with ion-stabilized crumpling on short length scales. This short-scale structure is additionally supported by force- and salt-dependent quantification of the RNA ion atmosphere composition, which shows that ions are liberated under stretching; the number of ions liberated increases with increasing bulk salt concentration. Both this result and the observation of two elastic-response regimes directly contradict the WLC model, which predicts a single elastic regime across all forces and, when accounting for a standard model of the electrostatics-dependent persistence length, the opposite trend in ion release with salt concentration. We conclude that ssRNA, like ssDNA, is better described as a snakelike chain, characterized by smooth bending on long length scales and ion-stabilized crumpling on short length scales. In monovalent salt, these two regimes are separated by a characteristic length that scales with the Debye screening length, highlighting the determining importance of electrostatics in RNA conformation.

3.1 Results

3.1.1 Elasticity measurements

Since we sought to probe the elastic response of RNA in the absence of higher-order structure, we studied the polyuridine homopolymer (poly(U)), which does not undergo base-pairing or tertiary structure formation, which possesses an unfavorable free energy for base-stacking interactions [73], and which showed no signature of base-stacking in earlier elastic measurements [69]. Nonetheless, poly(U) regions do occur in biologically significant RNAs, such as regulatory sRNA [74].

Enzymatically synthesized poly(U) was end-labeled and tethered, at one end, to a functionalized glass coverslip and, at the other end, to the paramagnetic bead (see Appendix A.1.1) required for the magnetic tweezers experiment.

We measured the elasticity of poly(U) RNA as a function of the concentration of three salt species: NaCl, MgCl₂, CaCl₂. A representative subset of these data, rescaled by the extension at 10 pN applied force, is shown in Figure 3.1 (see further discussion of this rescaling in Appendix A.2.6). All of these curves exhibit the same qualitative behavior (Figure 3.2A): two elastic-response regimes separated by a salt-dependent crossover force, f_c . Over a range of salt concentrations (20 to 2000 mM NaCl, 0.2 to 20 mM MgCl₂, 0.2 to 10 mM CaCl₂), and for $f < f_c$, the extension obeys $X \sim f^{0.6}$, approximately the $f^{2/3}$ Pincus power-law elasticity expected of a self-avoiding chain [32]; these salt concentrations thus correspond to the good-solvent regime for the RNA. For $f > f_c$, the response depends on the screening ion valence, a phenomenon that is further described below.

For each salt species, there is a concentration for which the low-force elasticity exhibits the linear behavior expected of an ideal chain ($X \sim f$), corresponding to the solution Θ condition. At Θ , the RNA elasticity is well fit by the WLC interpolation formula

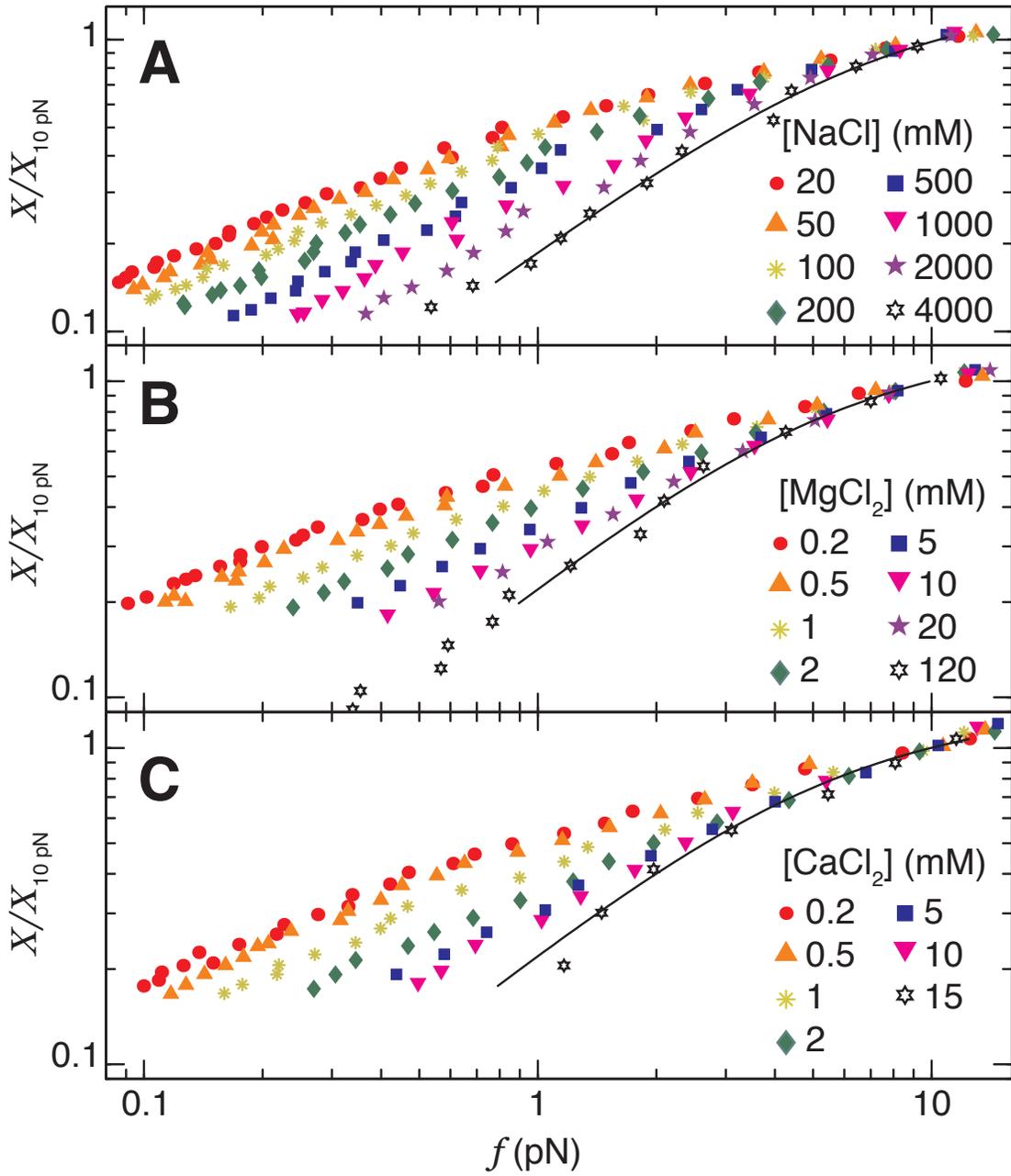


Figure 3.1: Elasticity measurements of poly(U) RNA in various concentration solutions of (A) NaCl, (B) MgCl₂, and (C) CaCl₂. Lines indicate WLC fits at the solution Θ points.

Table 3.1: Persistence length of poly(U) RNA at Θ in various salt solutions, compared with ssDNA results of McIntosh and Saleh [25]

salt	RNA		ssDNA [25]	
	c_Θ (mM)	l_p (nm)	c_Θ (mM)	l_p (nm)
NaCl	$\approx 4,000$	0.83 ± 0.05	$\approx 3,500$	0.60 ± 0.02
MgCl ₂	≈ 120	0.78 ± 0.08	≈ 50	0.64 ± 0.03
CaCl ₂	≈ 15	0.83 ± 0.07	≈ 20	0.61 ± 0.02

(Equation 2.3) [28]. WLC fits at Θ are also shown in Figure 3.1 and the fitted values of persistence length and corresponding Θ concentrations, c_Θ , are listed in Table 3.1. We also list, for comparison, c_Θ and l_p for ssDNA [25]. Values of l_p listed were obtained by fitting all curves at the Θ -concentration, not just the representative curve plotted in Figure 3.1. The values of l_p are in general agreement across all salt species tested, consistent with the marginalization of electrostatic effects at Θ .

3.1.2 Regime crossover analysis

From each force-extension curve well below the Θ concentration (Figure 3.1), we extract the inter-regime crossover force, f_c , and extension, X_c , as described in Appendix A.2.6. We then use those values to rescale each curve and find that they collapse to a single, universal curve for a given salt species (Figure 3.2A). Although poly(U) in all three species conforms to an $X \sim f^{2/3}$ dependence for $f < f_c$, for $f > f_c$ the behavior is valence-dependent: in monovalent NaCl, $X \sim \log(f)$, whereas in the divalent salts we observe additional compliance.

Since an applied force corresponds to a tensile screening length (Equation 1.2), the crossover force f_c corresponds to a crossover length scale, $\xi_c = k_B T / f_c$, separating conformational regimes [43]. To investigate the salt-dependence of that length scale, we plot in Figure 3.2B the relation between f_c and the solution ionic strength, I . Ionic strength is used because it is the relevant parameter in the Debye-Hückel treatment of solution elec-

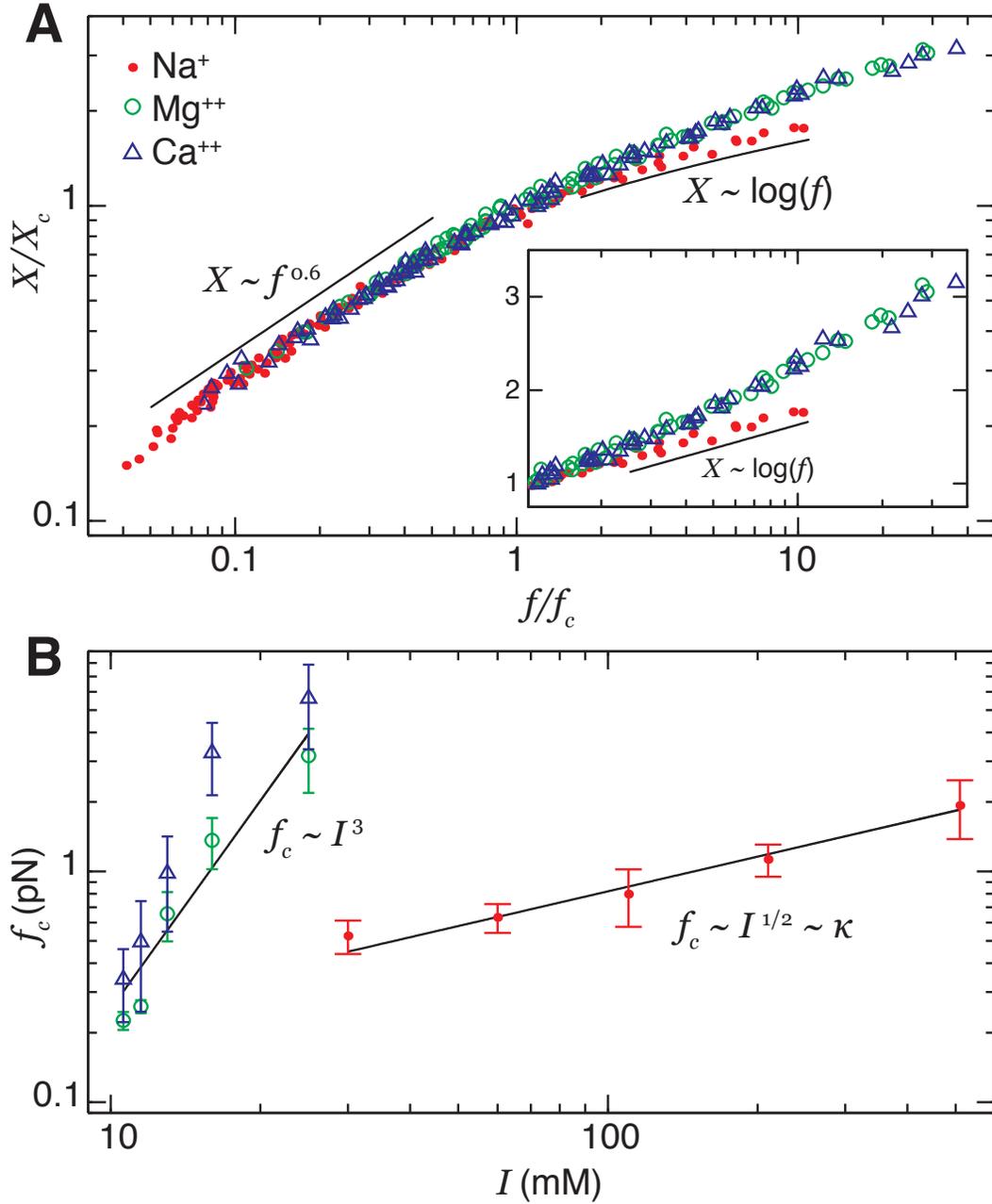


Figure 3.2: Analysis of elastic regime crossover force, f_c . (A) Force-extension curves, in good solvent, rescaled by f_c and L_c . Lines are drawn to highlight the low-force power-law behavior common to both valences and the high-force logarithmic behavior in monovalent salt. (B) f_c plotted as a function of ionic strength (same symbols as in (A)). Lines corresponding to a power-law dependence of f_c on I are shown for both valences; in monovalent salt this relationship is consistent with inverse κ^{-1} scaling.

trostatics and provides a framework to compare mono- and divalent salt. For monovalent salt, we find $f_c \sim I^{0.58 \pm 0.09}$, consistent with crossover length scaling as $\xi_c \sim I^{-1/2} \sim \kappa^{-1}$. We thus identify κ^{-1} as an important conformational length scale of unstructured RNA in monovalent salt, corresponding to the SLC blob size separating short- and long-scale conformations. In divalent salt, $f_c \sim I^3$ scaling is observed, which does not correspond to an obvious physical length scale and remains unexplained.

3.1.3 Quantification of associated ions

Force-extension data permit quantification of the ion atmosphere surrounding the molecule as a function of both f and c_{bulk} . The concept of the ion atmosphere will be further discussed in Chapter 5 and the theory used to extract ion atmosphere measurements from SMFS data is discussed in much greater depth in Appendix D. Briefly, since NAs are negatively charged, they tend to attract cations and repel anions. This leads to a region surrounding the NA of enhanced cation concentration and reduced anion concentration. This region is referred to as the ion atmosphere. The ion excess, Γ , is a stoichiometric quantification of this atmosphere: the number of ions of a particular species in the vicinity of the NA above, or below, the number that would be present due to the bulk concentration alone. Here we use measurements of Γ both to study how the number of associated screening ions is modulated by RNA conformation and to directly test a WLC-derived model of salt-dependent RNA elasticity. In our analysis, the independent experimental variables are f and μ , the chemical potential; these are related to the thermodynamically conjugate, dependent variables X and Γ . We can relate f , c_{bulk} , X , and Γ by the Maxwell relation: [75, 76]

$$\left. \frac{\partial X}{\partial \mu} \right|_f = \left. \frac{\partial \Gamma}{\partial f} \right|_\mu. \quad (3.1)$$

From this, we derive an integral expression giving the difference in ion excess, $\Delta\Gamma$, between a state with applied force f and some reference state with applied force f_o :

$$\Delta\Gamma = \frac{c_{\text{bulk}}}{2k_B T} \int_{f_o}^f \left. \frac{\partial X}{\partial c_{\text{bulk}}} \right|_{f'} df'. \quad (3.2)$$

This equation assumes an ideal solution; accounting for actual ion activities changes the results by less than 10% [75] and does not affect key conclusions. To perform the partial differentiation and integration required by Equation 3.2, we fit a smooth surface to $X(f, c_{\text{bulk}})$. Applying Equation 3.2 to this surface, we can quantify changes to the polymer-associated ion excess as the RNA is stretched.

The results of this analysis, for a representative poly(U) molecule in the presence of 50 to 1000 mM NaCl and subject to 0.1 to 10 pN applied force, are shown in Figure 3.3A. As the RNA is stretched, we observe that associated ions are driven away ($\Delta\Gamma < 0$). For sufficiently large forces, the number of ions driven away increases with increasing salt concentration, a result we explain by arguing that $\Delta\Gamma$ per SLC blob is constant and that the number of blobs per length increases with salt concentration, since blob extent scales with κ^{-1} . This phenomenon is seen directly in Figure 3.3B, where we plot the ion excess per approximate blob size, κ^{-1} , versus rescaled force, f/f_c . The excess at f_c , the force corresponding to the blob size, was chosen as a natural zero-reference. Under this rescaling, all of the ion excess curves collapse to a single, master curve, indicating that the variation between curves was attributable to the disparate number of SLC blobs per length. Because most of the ion excess change occurs when $f/f_c > 1$, corresponding to $\xi < \xi_c$, we conclude that the ions are predominately associated with the intrablob crumpling and not with the longer-length-scale chain of blobs. This analysis, which confirms the SLC picture of ion-stabilized crumpling on short length scales (see Section 3.2), was repeated for two other RNA molecules and yielded comparable results.

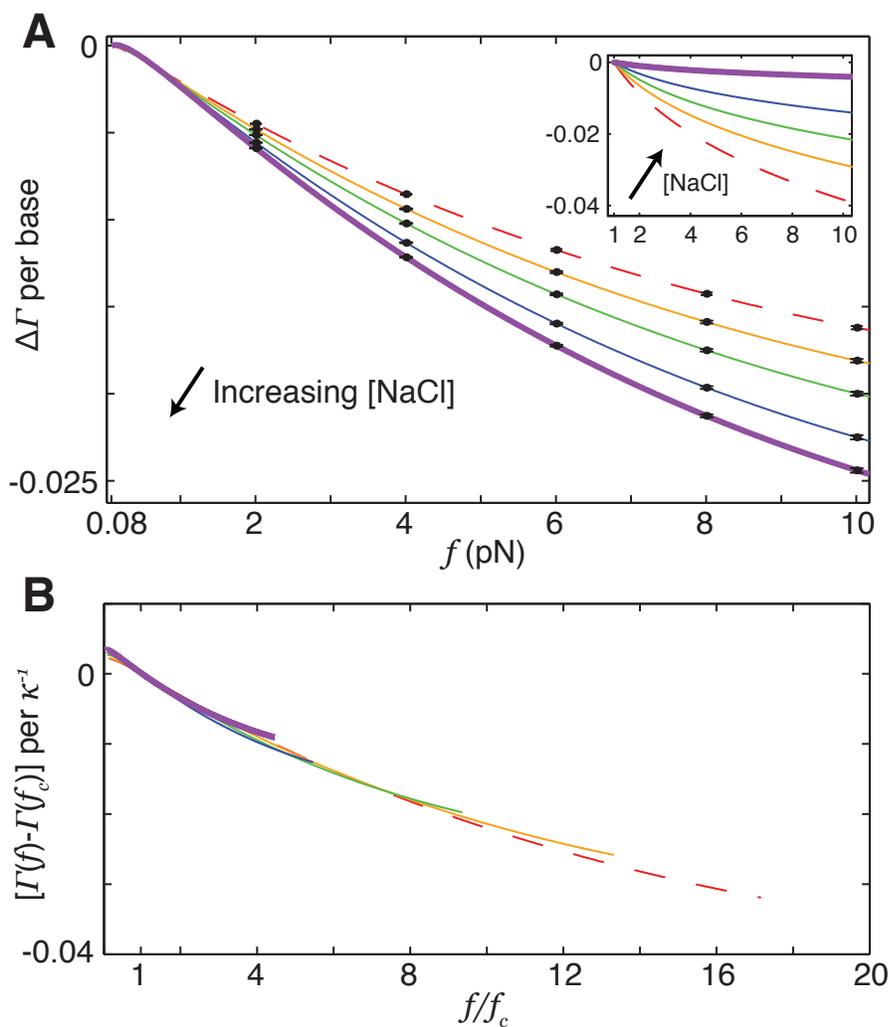


Figure 3.3: Determination of RNA-associated ion excess as a function of applied force and monovalent salt concentration. (A) Change in ion excess, per nucleotide, as a function of force, compared with a low-force reference state, for various NaCl concentrations (dashed line: lowest concentration; heavy line: highest concentration). Error bars reflect propagated uncertainty from force measurement; an additional 10% error common to all curves arises from uncertainty in polymer contour length. Inset: Expected ion excess curves for a WLC model with electrostatics-dependent persistence length [28, 52], using parameters of Ref. [42] (same ordering of line weight with concentration). (B) Change in ion excess, per Debye length, as a function of f/f_c , compared with a reference state at f_c (values of f_c from Figure 3.2B) for the same NaCl concentrations.

These data also provide a test of the modified WLC model with electrostatics-dependent persistence length from the Barrat-Joanny modification of the OSF theory [28, 52] that is sometimes used to account for the polyelectrolyte character of RNA. Using such a modified WLC model, with parameters obtained from the poly(U) elasticity data of Seol *et al.* [42], we generated force-extension curves at various salt concentrations and, from them, computed the change in ion excess. The results, shown in the inset to Figure 3.3A, show a contradictory trend in $\Delta\Gamma$ with increasing c_{bulk} compared with our experimental observations.

3.2 Discussion

3.2.1 Elastic response

The crossover force separating observed elastic-response regimes (Figure 3.2B) corresponds, via a tensile screening length argument, to a cross-over length scale, ξ_c , separating spatial conformational regimes. In monovalent salt solution, we observe that $f_c \sim I^{1/2}$, corresponding to $\xi_c \sim I^{-1/2}$. In the Debye-Hückel theory, the Debye length varies as $\kappa^{-1} \sim I^{-1/2}$; thus $\xi_c \sim \kappa^{-1}$. This indicates that, in monovalent salt solution, there are separate conformations of unstructured RNA on short and long length scales, and that they are separated by the Debye screening length, as illustrated in Figure 3.4. This result is in agreement with the finding—from the simulation studies of Toan and Thirumalai [53]—that the tangent vector correlation of a flexible polyelectrolyte transitions from power-law to exponential decay at κ^{-1} and with similar behavior seen in ssDNA (Section 2.3.2). In divalent salt solution, however, this scaling is not observed. Instead, we see $f_c \sim I^3$ for $c_{\text{bulk}} \leq 5$ mM ($I \leq 25$ mM) and non-power-law scaling for the full range of tested salt concentrations in the good solvent regime. Thus, κ^{-1} is not the

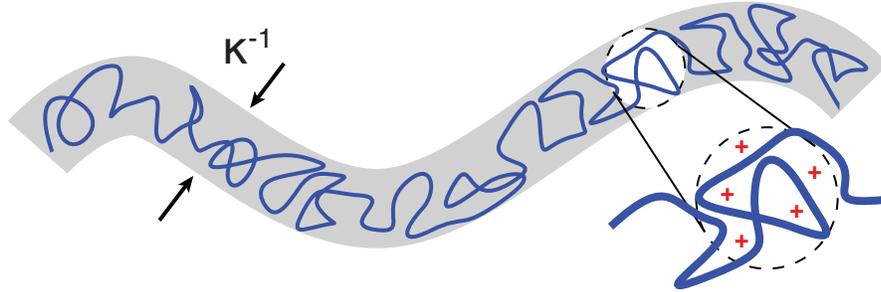


Figure 3.4: SLC conformation. In monovalent salt, the long length scale bending and short length scale, ion-stabilized crumpling regimes are separated by the Debye screening length, κ^{-1} . In divalent salt, κ^{-1} is replaced by a crossover length scale that decreases more rapidly with increasing ionic strength and the conformation on short length scales is more tightly packed than in the monovalent case.

important length scale separating conformational regimes in the presence of appreciable divalent salt, as is the case under physiological conditions. Instead, κ^{-1} of Figure 3.4 is replaced with a length scale that decreases more rapidly with increasing I , consistent with the enhanced efficiency of divalents (Figures 3.1B and 3.1C, compared with 3.1A) and the breakdown of the Debye-Hückel picture in this case [25].

Having identified the length scale that separates the two conformational regimes of unstructured RNA, we now characterize those regimes. In the limit of long length scales ($f < f_c$), in both monovalent and divalent salt, we observe, approximately, the $X \sim f^{2/3}$ scaling characteristic of a chain of blobs swollen by excluded volume interactions [32]. Unlike in uncharged polymers, where steric effects dominate, here the primary mechanism of swelling is the Coulomb repulsion between charged monomers. This is seen by noting the absence of the thermal blob regime.

That the exact numerical exponent we measure for this low-force scaling, $X \sim f^C$ with $C = 0.6$, differs in detail from the $C = 2/3$ (Pincus scaling) result for an infinite chain of swollen blobs can be attributed to the finite length of the polymers under study

(4000 to 10,000 nucleotides). Simulation studies have shown that Pincus scaling does not arise in flexible, neutral polymers with degree of polymerization less than $N \approx 10^5$ monomers [77]. For polymers shorter than this, the applied force is felt within the blobs. More recent simulations have shown that Pincus scaling occurs in flexible, charged polymers for $N = 25,000$, but not for $N = 5000$ ($C = 0.60$ scaling is seen in that case) [78].

In the limit of short length scales ($f > f_c$), we observe an elastic response that depends strongly on salt valence. In monovalent salt, we find $X \sim \log(f)$, a more compliant response than the WLC behavior expected for an uncharged polymer; in divalent salt we observe even further compliance (Figure 3.2A). These results support the SLC picture of short length scale polymer crumples stabilized by transient interactions with screening ions: contour length sequestered in the crumples is liberated by applied force and manifests as increased compliance. That additional high-force compliance, and therefore additional local crumpling, is observed for divalent ions is consistent with this picture: the more highly charged ions have stronger interactions with the RNA backbone, allowing it to adopt sharper bends and thus sequester additional contour length. This short length scale crumpling, and its variation with ion valence, is also seen in molecular dynamics simulations of a charged, flexible polymer in the presence of explicit screening ions [65]. Behavior in this higher-force regime will be more fully discussed in Chapter 4.

3.2.2 Associated ion excess

The existence of an ion-associated, crumpled conformation on short length scales is further supported by our measurements of the force- and salt-dependent ion excess (Figure 3.3). In particular, it is seen in the decrease in ion excess with increasing force since, if the contour length sequestered in crumples is stabilized by co-localization of

screening ions, the liberation of that contour length by an applied force concomitantly liberates those ions. The decrease in ion excess change with increasing salt concentration, for sufficiently high forces, is consistent with the liberation of a constant number of ions per SLC blob, but is directly at odds with the prediction of a modified WLC model [28, 42, 52] employing an electrostatics-dependent persistence length (Figure 3.3A). That all of the ion excess curves collapse to a single, master curve when plotted as $[\Gamma(f) - \Gamma(f_c)]$ per κ^{-1} vs. f/f_c (Figure 3.3B) indicates that the salt-dependence enters only through changing the size of the Debye length-scale SLC blobs. This is consistent with our measurements of elasticity, which indicate (Figure 3.3A) universal behavior across all salt-concentrations modulo a salt-dependent crossover force that corresponds to the SLC blob size.

3.2.3 Comparing ssRNA and ssDNA

The SLC conformation of unstructured RNA described above is in good qualitative agreement with findings from earlier studies of ssDNA using the same techniques [25, 75]. However, there exist quantitative differences in the persistence lengths and Θ conditions between the two polymers. Although RNA and DNA are structurally analogous, they differ in the substituent of the 2' sugar carbon. The added hydroxyl group in RNA forces the pentose ring into the $C_{3'}-endo$ conformer (DNA is in equilibrium between $C_{3'}-endo$ and $C_{2'}-endo$) [73], possibly because of the formation of a water bridge between the 2'-hydroxyl and 3'-phosphate [79]. This structural constraint is expected to increase the rigidity of RNA compared with DNA, which is indeed borne out by the longer persistence lengths we observe in RNA compared with DNA (Table 3.1). Base stacking has been proposed as an alternate origin of this rigidity [80], but we see no signature of such stacking in these elasticity measurements or those of Seol *et al.* [69].

The l_p comparison we report is especially illuminating given that the two measurements were made using the same technique (single-molecule elasticity at low force under Θ conditions); the same was true in a study of poly(U) and poly(dT), using small-angle x-ray scattering and single-molecule Förster resonance energy transfer, which found a 34% difference in l_p between ssDNA and ssRNA when extrapolated to Θ [80], in general agreement with the 38% difference we report. Away from Θ , reported measurements of l_p vary widely (Ref. [81], and citations therein); given the results reported here, we attribute this to the non-WLC nature of the single-stranded nucleic acids.

Additionally, the extra hydroxyl group of RNA, acting as a hydrogen bond donor, is expected to increase water solubility. Since the Θ condition occurs at the transition between the good solvent and poor solvent regimes [1], this increased solubility of RNA explains the higher observed Θ concentrations of NaCl and MgCl₂ compared with those of ssDNA. That the Θ concentration of CaCl₂ is not similarly higher in ssRNA than in ssDNA is not explained by this argument and could be due to some other effect, perhaps a specific chemical interaction between the salt ions and the RNA.

3.3 Conclusion

Full understanding of the diverse biological functions of RNA can only be obtained through a detailed understanding of its three-dimensional conformation and structure. Many factors contribute, including the interactions responsible for secondary and tertiary structure. However, in the absence of these phenomena, the conformation of RNA is dominated by the solution-screened Coulomb interaction between the charged phosphate groups of the backbone. From the results presented above, we conclude that unstructured RNA, like ssDNA, exhibits an SLC conformation (Figure 3.4), characterized by bending on long length scales and ion-associated crumpling on short length scales. These two

regimes are separated by the salt-dependent SLC blob size that, in monovalent solution, scales with the Debye screening length. The existence of two conformational regimes is supported by our experimental observation of two elastic-response regimes in the force-extension curves, separated, in monovalent salt, by a force scaling inversely with κ^{-1} . That the short-range conformation is stabilized by ion association is supported by both the enhanced high-force compliance observed in divalent salt solution and the observation that excess ions are liberated from association with the RNA when it is pulled apart.

Although the WLC model of polymer elasticity has long served as an empirically effective, and analytically tractable, model for unstructured RNA conformation, our results indicate that it does not accurately describe the microscopic details. In particular, it predicts $X \sim f$ scaling in the low-force limit, an $(L_c - X) \sim 1/f^{1/2}$ approach to the contour length in the high-force limit [28], and, when adjusted to account for electrostatics-dependent persistence lengths, a decrease in $\Delta\Gamma$ with increasing salt. In contrast, the SLC behavior we report in this study occurs as $X \sim f^{2/3}$ scaling at low force, $X \sim \log(f)$ at moderately high force, and increasing $\Delta\Gamma$ with increasing salt. This discrepancy likely contributes to the inconsistency across different types of experiments in measuring l_p , since in most cases WLC behavior must be assumed. Better characterization of SLCs requires analytical models; whereas Toan and Thirumalai provide an elastic prediction for the SLC [53] that replaces the Marko-Siggia WLC elastic theory [28], it is based on the phenomenological assumption of power-law correlations rather than microscopic physical details. The description and experimental validation of such a physically motivated model is the subject of Chapter 4. Additionally, we are unaware of an SLC-motivated prediction for the scattering structure factor that would replace the results of Pedersen and Schurtenberger [82].

Finally, these results leave open the question of the physical origin of some of the elastic behavior observed. While the low-force Pincus scaling is well described in terms of

the behavior of a self-avoiding chain of excluded volume blobs, the higher-force logarithmic behavior has not been explained in terms of a physical model of polymer behavior, except through numerical simulations [53,65] or the power-law decay conjecture of Toan and Thirumalai [53]. Since the logarithmic regime characterizing the SLC extends beyond the high end of the force range of this study (~ 10 pN), further insight can be gained through measurements of ssDNA and ssRNA in the 10–100 pN window. Such results are also discussed in Chapter 4.

Chapter 4

Intermediate-force elasticity:

Experiments and mean-field model

Single-stranded nucleic acid elastic behavior—and that of flexible polyelectrolytes more generally—is complicated by the strong negative charge of the molecule. This is particularly true for applied forces corresponding to nanometer-scale structure, where many characteristic length scales of the system become conflated. These include the chemical bond length (~ 0.2 nm), the statistical monomer (Kuhn) length (1.2 nm for ssDNA, 1.7 nm for ssRNA [26]), the structural charge spacing (0.7 nm for ssDNA, 0.59 nm for ssRNA [23]), the Debye screening length (0.8 nm at 150 mM ionic strength), and the Bjerrum length (0.7 nm in water).

Prior studies have elucidated ssNA elastic behavior in the low- and high-force limits. At low force, corresponding to length scales larger than κ^{-1} , electrostatic repulsion swells the random coil conformation and manifests in the elastic response as excluded volume (Pincus) scaling: $X \sim f^{2/3}$ [25, 26, 32, 43]. At high force, corresponding to length scales smaller than the chemical bond length, electrostatic effects become negligible and the same behavior is seen as in uncharged polymers: entropic wormlike chain (WLC) elas-

ticity transitioning to entropic freely jointed chain (FJC) behavior and then to enthalpic stretching of the backbone [40, 44, 45, 55, 83, 84].

There are not similar results, grounded in the microscopic physics of the system and comprehensively verified by experiments, for the intermediate force regime in which the system transitions from electrostatics- to backbone structure-dominated behavior. Models have been put forward in terms of an electrostatics-dependent persistence length [50–52, 85]. These models have the disadvantage of being couched in terms of a quantity that is ill-defined unless the chain exhibits an exponential decay in its tangent vector correlation function; simulations suggest that power-law decay may instead occur in this force range [53]. Experimental agreement has been seen with these models [42, 64], but only when a phenomenological, salt-dependent charge density is invoked. This charge density [86] is obtained from an effective charge Debye-Hückel approach [87–90] that is only expected to be valid in the far field, not for the interactions of neighboring charges on the same ssNA backbone. Furthermore, comparison to experiment has only been made at a single salt concentration, constituting a poor test of an electrostatic effect that manifests as a function of ionic strength.

In this chapter, we use MT SMFS [91] to comprehensively measure the elasticity of ssDNA and ssRNA over 5–100 pN applied force (corresponding to 50–85% relative extension) and 1–1000 mM ionic strength. We then introduce a new model—in terms of an unambiguously defined mean-field electrostatic tension—that captures the observed behavior. Finally, we combine this result with the existing knowledge at lower and higher forces to furnish a comprehensive picture of ssNA elasticity—and therefore statistical conformation—over all accessible forces.

4.1 Results

4.1.1 Elasticity measurements

The elasticity of 1–6 μm intrinsically single-stranded ssNAs was measured using magnetic tweezers [91,92], in which the molecule is attached, at one end, to the substrate and, at the other end, to a paramagnetic bead. A magnetic field gradient exerts a force on the bead, and the resulting molecular extension is measured. In contrast with Chapter 3, here 2.8 μm beads are used to allow higher forces to be probed.

The chemical means used to synthesize the ssNAs (see Appendix A.1.1) produce polydisperse samples, meaning that the chain contour length, L_c , must be left as a free parameter in subsequent fitting. By collecting data on the same molecule across a broad range of ionic strength, I , the L_c parameter can be constrained and fit globally across all of the curves. Examples of force-extension data, for representative molecules, are shown in Figures 4.2A (DNA) and 4.2B (RNA).

4.1.2 Force-extension model

We model these ionic strength-dependent force-extension data by representing the repulsive interactions between the charged monomers as a mean-field electrostatic tension, f_{el} (Figure 4.1A). In the high-force limit ($X > L_c/2$) and neglecting end effects, this tension aligns with the applied force and both act uniformly along the length of the chain. We thus make the substitution $f \rightarrow f_{\text{app}} + f_{\text{el}}$ in the high-force expression for WLC elasticity [28] to obtain

$$X = L_c \left(1 - \sqrt{\frac{k_B T}{4l_p(f_{\text{app}} + f_{\text{el}})}} \right), \quad (4.1)$$

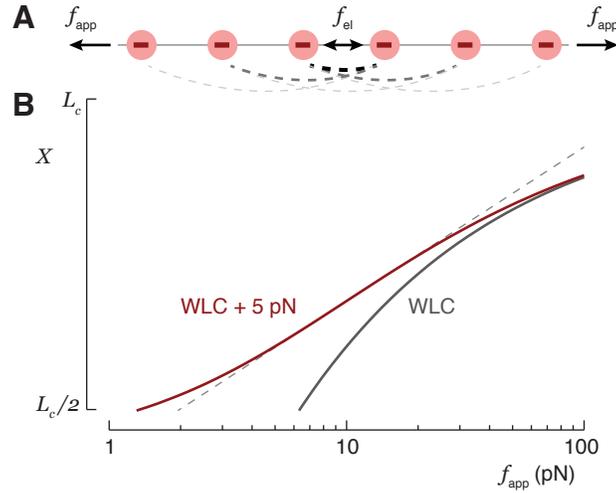


Figure 4.1: (A) When a polyelectrolyte is stretched, the applied force, f_{app} , acts uniformly along the length of the chain. In a mean-field approximation, the Coulomb repulsion between charges can be similarly treated as a uniform tension, f_{el} . (B) Introduction of 5 pN electrostatic tension to Equation 4.1 shifts the high-force WLC curve to larger extension. A domain of $X \sim \log(f)$ scaling (dashed line) emerges. Evaluated with $l_p = 0.6$ nm.

where l_p is the bare (i.e., non-electrostatic) persistence length. In other words, we subsume all electrostatics into f_{el} : a physically well-defined quantity that could, for example, be read out from the results of simulations. An example of the effect of adding $f_{el} = 5$ pN to the WLC is shown in Figure 4.1B, in which the force-extension curve shifts to higher extension and a domain of $X \sim \log(f)$ scaling is observed. Note that this plot is limited to $X > L_c/2$, the cutoff needed to ensure chain alignment discussed above. For large values of f_{el} , an additional cutoff ($f_{app} > k_B T / \kappa^{-1}$) would have to be imposed to account for the transition to the low-force excluded volume regime [93]; this is the SLC blob size of Chapter 3.

Mean-field electrostatic tension has been successfully invoked in prior studies to explain the salt-dependence of discrete nucleic acid structural transitions. These include measurements of the DNA overstretching transition [48] and, more comprehensively, the base unstacking transition of poly(dA) ssDNA [68]. In both cases, the transition was

assumed to occur at a particular total tension, the sum of internal and external components. As the salt concentration was increased, the internal contribution decreased, meaning that more external tension had to be applied to effect the transition. Here we extend this approach to consider the effect of f_{el} on a continuous measurement of elasticity, with the total force needed to bring about a particular extension replacing the total force of a transition.

Our WLC-derived model (Equation 4.1) does not account for the WLC to FJC transition known to occur at sufficiently high force [44, 45, 83, 84]. Based on the arguments of Dobrynin *et al.* [44], and on fitting the WLC-FJC interpolation function of those authors to our data, we believe that this transition occurs at a force just above those probed in this study. As such, our data do not have the power to constrain the extra fitting parameter (i.e., the effective bond length) that would be needed to implement that model.

Fits of Equation 4.1 to the data are shown in Figures 4.2A and 4.2B. Since data for each molecule were collected across a broad range of I , we were able to constrain L_c and l_p globally across all curves and fit only f_{el} on a per-curve basis. In this way, N force-extension curves—each with 11–26 data points—are fit with $N + 2$ fit parameters.

For ssDNA, there is an approximately 2% disagreement between the data and fit at the highest force value. This likely arises from both the 8.4 nN stretch modulus of ssDNA [40] (accounting for 0.9% of the 2%) and from the WLC-FJC transition [44, 45, 83, 84], neither of which are accounted for in our fitting. Note that this value of the stretch modulus, due to Hugel *et al.*, is determined from both quantum chemistry calculations and rigorous atomic force microscopy experiments at the high force values relevant to this effect. This is in contrast with the early 0.8 nN value put forward by Smith *et al.* [55], which was based on force spectra collected at much lower forces and analyzed using an FJC model that did not account for the WLC-FJC transition. Addition of the correct stretch

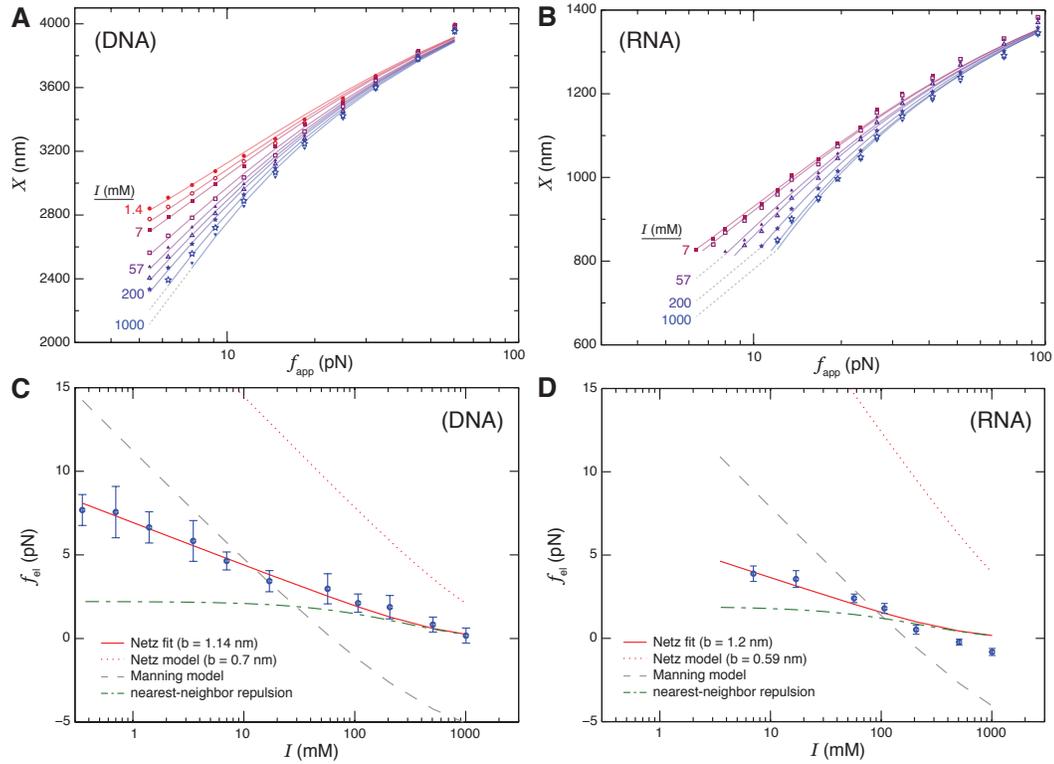


Figure 4.2: (A,B) Force-extension curves of a representative (A) ssDNA molecule over $I = 1.4$ –1000 mM and (B) ssRNA molecule over 7–1000 mM. The curves are fit to Equation 4.1, with a different f_{el} value for each curve and L_c and l_p fit globally across all the curves for a particular molecule. No particular model of f_{el} is assumed in the fitting. (C,D) Resulting values of f_{el} , combined across molecules, are plotted against I and are well fit by the Netz model of electrostatic tension (Equation 4.2) with concentration-independent b . Error bars are standard errors of the mean. (C) DNA data are best fit by $b = 1.14 \pm 0.05$ nm (solid line), compared with the result for the 0.7 nm structural charge spacing (dotted line). (D) RNA data are best fit by $b = 1.2 \pm 0.3$ nm, versus 0.59 nm structural spacing. In both cases, the dashed line is the best fit result of the Manning model [49] and the dot-dashed line is the result considering only nearest-neighbor effective charges.

modulus, which becomes more significant above about 100 pN (see Figure 4.3) does not appreciably change the results of the fit. A similar high-force disagreement is not seen in the RNA data, presumably because it is washed out by the effect of an anomalously small persistence length, discussed below.

4.1.3 Electrostatic tension

The values of $f_{\text{el}}(I)$ obtained from fitting Equation 4.1 to our data are shown, averaged across all experiments, in Figures 4.2C and 4.2D. As expected, f_{el} decreases as a function of I , consistent with greater charge screening at higher salt concentrations. In fact, at the highest values of I , f_{el} is dominated by just the interaction of nearest-neighbor charge pairs.

Two analytical models of polyelectrolyte electrostatic tension were fit to the $f_{\text{el}}(I)$ data with the effective charge spacing, b , as a free parameter. Both models consider the pairwise interactions between all charges along the chain (Figure 4.1A), as weighted by the effects of screening. Nearest-neighbor charges contribute only a fraction of the tension at low I (Figures 4.2C and 4.2D). One model we used is that of Netz [48], which considers a straight chain of charges (e.g., Figure 4.1A) subject to Debye-Hückel screening and predicts

$$f_{\text{el}} = \frac{k_B T l_B}{b^2} \left(\kappa b \frac{e^{-\kappa b}}{1 - e^{-\kappa b}} - \ln(1 - e^{-\kappa b}) \right), \quad (4.2)$$

where l_B is the Bjerrum length. The other model is that of Manning [49], which builds upon Equation 4.2 by explicitly accounting for counterion condensation. The Netz model is able to fit both our DNA and RNA data with quantitative precision, whereas the Manning model disagrees qualitatively.

As in the ssDNA unstacking experiments of McIntosh *et al.* [68], the Netz model is fit to our data with effective charge spacings (DNA: 1.14 ± 0.05 nm; RNA: 1.2 ± 0.3 nm)

Table 4.1: Parameters from global fitting of b , l_p , and L_c to each molecule using Equation 4.1 with f_{el} given by Equation 4.2. Values were averaged across all molecules studied. Uncertainties are standard errors of the mean.

	b (nm)	l_p (nm)
ssDNA	1.16 ± 0.03	0.55 ± 0.04
ssRNA	1.12 ± 0.04	0.41 ± 0.04

notably larger than the structural charge spacing of 0.7 nm for DNA and 0.59 nm for RNA [23]. For comparison, Figures 4.2C and 4.2D also show the Netz model evaluated with these structural separations; the disagreement with the data is apparent. The need to employ effective charge densities, which likely arises from the quantitative failure of the Debye-Hückel assumption underlying the model, is further discussed below.

4.1.4 Three-parameter global fit

By combining the Netz model of electrostatic tension (Equation 4.2) with our mean-field elastic model (Equation 4.1), we can describe a full, ionic strength-dependent elastic dataset using just three global fit parameters: chain contour length, bare persistence length, and effective charge spacing. Such fits (not shown) are of similar quality to those of Figures 4.2A and 4.2B. Parameter values from these fits were combined across all molecules studied and are listed in Table 4.1; values of L_c are not given, since the molecules were polydisperse.

4.2 Discussion

4.2.1 Force-independence of effective charge spacing

The good fit quality of our model, having force-independent b , is in contrast with the expectation that the effective charge spacing along the pulling axis should decrease

with decreasing force as the molecule becomes less extended. However, a charge density that does not change with extension is a prediction of counterion condensation theory [21, 22, 94, 95], which holds that sufficiently many monovalent cations will accumulate in the immediate vicinity of the negatively charged ssNA to reduce the effective charge spacing to l_B . Since the structural charge spacing of fully stretched ssDNA is 0.7 nm (equal to the Bjerrum length in room temperature water) and that of ssRNA is 0.59 nm, counterion condensation is expected to occur in the case of both polymers. In the low-salt limit, where this theory is formally valid, the change in number of condensed counterions between two extensions, $\Delta\Gamma$, is expected to scale linearly with the change in extension, ΔX .*

We can extract a measure of $\Delta\Gamma$ directly from our salt-dependent force-extension data via the thermodynamic argument introduced in Chapter 3 and further elaborated upon in Appendix D [30, 75, 76]. Starting from a Maxwell relation connecting applied force, extension, chemical potential, and number of condensed ions, we can derive the ion counting relation:

$$\Delta\Gamma = \frac{1}{k_B T A} \int_{f_1}^{f_2} \left(\frac{\partial X}{\partial \ln I} \right) df, \quad (4.3)$$

where A is an ion activity correction factor (approximately unity) [96] and f_1 and f_2 are the two forces between which $\Delta\Gamma$ is being measured.

We applied this relation directly to the ssDNA global fit results, which are in good agreement with the data and allow the partial derivative in Equation 4.3 to be easily

*Counterion condensation occurs whenever $b \leq l_B$, which is the case for both ssDNA and ssRNA regardless of extension. Consider a molecule in one of two stretched states, each with a different extension. Due to condensation, the effective charge spacing along the axis of extension is l_B in both states. The number of effective charges along the chain length at a particular extension is given by $\mathcal{N}_e = eX/l_B$. The change in number of effective charges between the states is the negative of the change in number of condensed cations. Therefore,

$$\Delta\Gamma = \frac{e}{l_B} \Delta X.$$

evaluated. At 1 mM ionic strength, we find that 0.07 ions per nucleotide are released when the ssDNA is stretched from $f_1 = k_B T / \kappa^{-1}$ (the onset of the mean-field WLC behavior) to $f_2 = 100$ pN. This is in contrast to the fractional extension change, with respect to the contour length, of 37%.

Thus, a nontrivial fraction of the charge density increase with decreasing extension is compensated for by counterion condensation, but not nearly the whole effect. One explanation for the discrepancy could be that b does actually have some degree of force-dependence, but that our model is not very sensitive to it. Electrostatic effects are most pronounced at lower forces; at high forces, all force-extension curves converge towards the same asymptotic WLC behavior (e.g., Figure 4.2A). Hence, it could be that the best-fit values of b are primarily constrained by the lower-force data.

4.2.2 Salt-independence and magnitude of effective charge spacing

The good quality of the fit of the Netz model to the $f_{\text{el}}(I)$ data implies that the effective charge density is salt-independent—note the sensitivity of this fit by comparing the same model evaluated for a different b in Figures 4.2C and 4.2D. This empirical result directly contradicts prior modeling of high-force ssNA elasticity [42,64] based upon an electrostatics-dependent persistence length [50–52,85]. These models relied upon a highly salt-dependent charge density [86] obtained from an effective-charge Debye-Hückel model [87–90] valid only in the far field and not suitable for treating interactions between neighboring charges. Furthermore, validation of these models was sought in studies that were not well-suited to critically test such a salt-dependent charge density, since the studies themselves were not conducted over a broad range of I .

The values of b obtained from our analysis (Table 4.1) are similar for ssDNA and ss-

RNA and are larger than the values expected based on the structural charge separation of the stretched chain. As mentioned above, anomalously large values of b were also seen in prior studies when Equation 4.2 was applied to the ssDNA unstacking transition [68]. In both cases, this means that the electrostatic monomer-monomer repulsion is weaker than expected, an effect that likely arises from the failure of the Debye-Hückel approximation underlying Equation 4.2. Numerical Poisson-Boltzmann calculations (not shown) predict significantly elevated ion concentrations close to the ssNA compared with the bulk, due to preferential attraction of cations. This is in contrast with Debye-Hückel, which predicts that cation attraction and anion exclusion will contribute equally to neutralizing the NA, leading to a local *total* ion concentration equal to that in the bulk. Locally elevated total ion concentrations are seen in the experiments of Chapter 6, which show roughly 0.4 more ions per nucleotide in the ssNA ion atmosphere [97] compared with the Debye-Hückel prediction, a result that decreases only very slowly with I (Figure 6.5). The locally elevated ion concentrations are consistent with an effective b value larger than the structural spacing. The weak dependence of the ion excess on I could explain why the effective b values are apparently independent of ionic strength, even though we hypothesize that they arise from an electrostatic effect.

4.2.3 Recovery of bare persistence length

The fitted l_p of ssDNA is in good agreement with the prior value (0.60 ± 0.02 nm) from measurements of single-molecule elasticity at low force (0.1–10 pN) under solution Θ conditions [25], where electrostatic interactions become negligible. This agreement validates the basic premise of our model: that the salt-dependent electrostatic interactions are contained only in f_{el} and that l_p should take on its bare, salt-independent value.

The ssRNA persistence length, however, is considerably lower than the 0.83 ± 0.05

nm value of Chapter 3. We investigated this effect by collecting ssRNA data over a wider force range (0.7–60 pN) under Θ conditions ($I = 4$ M). We find that fitting the data with the Marko-Siggia WLC interpolation formula [28] at high force (10–60 pN) gives $l_p = 0.43$ nm, similar to the result in Table 4.1. However, fitting at lower forces (0.7–10 pN) recovers a value of 0.85 nm, in agreement with the prior result [26]. Thus, we see that fitting Equation 4.1 to our data does recover the bare persistence length, in keeping with our assumption that only f_{el} is electrostatics-dependent, but that the bare persistence length itself is force-dependent in a way that that of ssDNA is not. This may arise from the C3'-*endo* to C2'-*endo* sugar pucker isomerization, or another structural transition, in ssRNA [73,98]. Regardless of the origin of the effect, this anomalous high-force compliance introduces a complication into the analysis of the ssRNA data that is not present in the ssDNA and that may explain the negative values of f_{el} seen at large I in Figure 4.2D.

4.3 Regimes of elastic behavior

We have demonstrated the validity of a new model of flexible polyelectrolyte elasticity in the intermediate force domain through its ability to reproduce experimental ssDNA and ssRNA force-extension data using only three fit parameters. This model accounts for the transition between behavior dominated by electrostatics, at the lower end of the force domain, and by the structure of the covalent backbone, at the higher end. Unlike prior efforts, this model is couched in terms of a clearly defined physical quantity, the electrostatic tension, and has been rigorously validated by comparison with experimental data over a broad range of ionic strength.

Our model bridges behaviors that were already known to occur at lower and higher forces to provide a comprehensive view of flexible polyelectrolyte elasticity over all con-

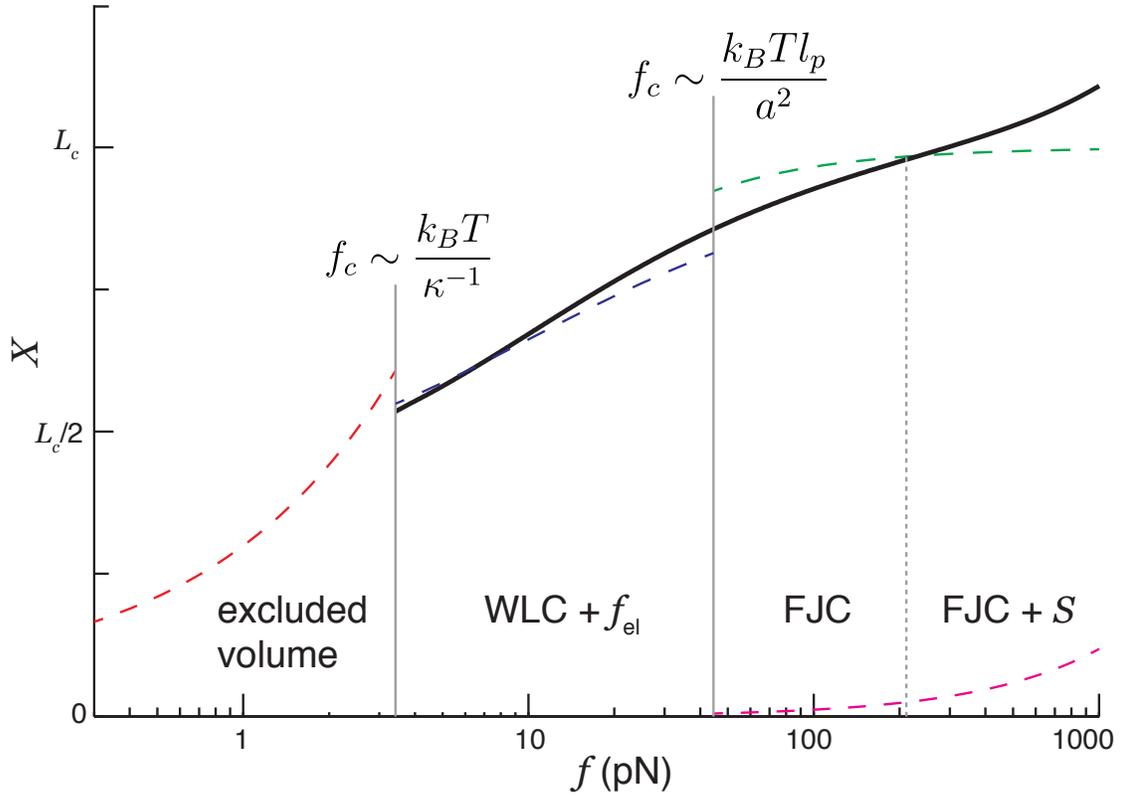


Figure 4.3: The mean-field WLC model (WLC + f_{el}) connects lower-force elastic regimes dominated by electrostatics and higher-force regimes dominated by the structure of the chemical backbone. Dashed lines denote the expected scaling in the different force regimes: $X \sim f^{2/3}$ excluded volume scaling, the WLC with mean-field electrostatic tension (Equation 4.1), the FJC, and the additional contribution from stretch modulus $S = 8.4$ nN. The lowest-force linear response regime is not shown. The solid line is the elastic model of Dobrynin *et al.* [44], modified to account for mean-field electrostatic tension, as in Equation 4.1, and for the stretch modulus. A broad domain of quasi-logarithmic scaling arises from the interpolation between the excluded volume, mean-field WLC, FJC, and adiabatic stretching regimes.

formational length scales (Figure 4.3). In particular, we now expect, starting from the lowest forces, regimes of (1) $X \sim f$ scaling arising from linear response, (2) $X \sim f^{2/3}$ scaling arising from electrostatic excluded volume interactions, (3) mean-field WLC behavior corresponding to Equation 4.1, (4) FJC behavior arising from the fundamentally discrete nature of the chemical backbone, and (5) enthalpic stretching of the backbone itself.

Combining our model with the WLC-FJC interpolation formula of Dobrynin *et al.* [44] and the stretch modulus results of Hugel *et al.* [40], it is now possible to evaluate the elasticity in the three highest-force regimes in a completely quantitative manner using only four fit parameters. These are the three parameters of our model, plus the effective chemical bond length, a , needed in the Dobrynin expression; enthalpic stretching, at least for ssDNA, is completely constrained with no added parameters. This model is plotted in Figure 4.3, with $l_p = 0.60$ nm, $f_{el} = 5$ pN, and $a = 0.34$ nm. Similarly quantitative descriptions of the linear response and excluded volume regimes [99], going beyond the level of a scaling theory, require a model of the salt-dependence of the ssNA excluded volume; we are unaware of such a model.

The Dobrynin model with added electrostatic tension and stretch modulus shows that the quasi-logarithmic scaling behavior of our mean field WLC model (e.g., Figure 4.1B) actually extends, through interpolation with the higher-force regimes, for several decades in force. A further extension into the excluded volume regime is also expected, based upon the simulations of Stevens and Saleh [66]. This result is consistent with the magnetic tweezers study of Dessinges *et al.* [64] and the atomic force microscopy study of Rief *et al.* [63], which together showed a very broad logarithmic scaling regime at a single salt concentration. Here we have come to understand this scaling in terms of a series of specific effects: chain stiffening due to electrostatic self-repulsion, the discrete nature of the chemical backbone, and the enthalpic deformability of that backbone. An alternative

explanation of this behavior, not grounded in terms of such microscopic details, was put forward by Toan and Thirumalai [53], who showed that logarithmic elastic scaling can arise from a power-law (as opposed to exponential) decay of the chain's tangent vector correlation function. These approaches are not mutually exclusive, and it may be that chain-straightening effects, deriving from the electrostatic tension outside of the well-aligned regime, manifest as such a power-law decay.

Chapter 5

The nucleic acid ion atmosphere: Theory and existing results

Double- and single-stranded DNA and RNA are all strongly negatively charged, with sub-nanometer inter-phosphate charge separations. This high charge density means that the free energy of NA interactions will be strongly influenced by electrostatics. Under physiological conditions, salt ions in solution mitigate electrostatic effects by screening the potential of the NA, reducing it from a long-range $1/r$ decay to an exponential decay in the far-field. Because the screening ions are free to move in solution, positive cations will accumulate near the negatively charged NA while negative anions will be repelled from its vicinity; this region of perturbed ion concentration is known as the *ion atmosphere* [100] and can differ significantly from the bulk concentration [101].

We begin this chapter by defining the parameters used to quantify the NA ion atmosphere and discussing the methods used to experimentally measure these parameters. We then briefly review theoretical models of the ion atmosphere before surveying the existing experimental results. We will critically analyze the extent to which the experimental findings can be used to constrain the theoretical models. Throughout, we will focus on

diffusely associated ions; ions that bind to specific structural sites on the NA [102] are beyond our scope.

5.1 Quantifying the ion atmosphere

5.1.1 Three-dimensional distribution

The ion atmosphere of an NA can be fully described, under given conditions, by specifying the local concentration of each ion species as a function of position. Such distributions are often the output of theoretical models of the ion atmosphere, and an experimentally measured spatial distribution could directly test these theories. However, to our knowledge no experimental study has measured the full distribution of the ion atmosphere of any NA. Partial spatial information has been obtained from anomalous small-angle x-ray scattering (ASAXS) experiments [103–106], in which x-ray scattering data are recorded at several wavelengths near an absorption transition of the ions. This allows the scattering pattern of the ions (typically Rb^+ and Sr^{2+}) to be isolated from that of the solvent and the NA. The spatial information in these studies is contained in the q -dependence of the scattering intensity, where q is the scattering wave vector. However, constant signal-to-noise is not seen across all values of q ; in particular, robust signal is seen at small q (corresponding to large spatial length scales), whereas poor signal is seen at large q (corresponding to short length scales). This disparity occurs because smaller features in the ion distribution necessarily present smaller scattering cross-sections and means that ASAXS, while an extremely powerful technique, is limited in the information it provides about short-range interactions of ions with the NA.

5.1.2 Ion counting

Absent the ability to measure the full distribution of the ion atmosphere across all length scales, knowledge of its composition must be pieced together from other experimental data, including both ASAXS and *ion counting* studies. Ion counting experiments measure how many of each ion species (i.e., cation, anion) are present in the ion atmosphere. These quantities are constrained since, taken together, the NA and its ion atmosphere must be charge neutral. Measurements are typically made as a function of bulk salt concentration, ion identity, and NA structure.

The canonical ion counting experiment is Donnan equilibrium dialysis [33], sketched in its idealized form in Figure 5.1A. In such an experiment, two compartments—a sample compartment containing an NA solution and a reservoir with a particular bulk salt concentration—are separated by a membrane that is impermeable to the NA under study but permeable to the ions. In the absence of NA, each ion species has the same concentration on both sides of the membrane. When NA is added, electrostatic interactions with the ions lead to a higher cation concentration, and lower anion concentration, in the sample compartment compared with the bulk reservoir. This imbalance contains information about the net ion atmosphere of the NA under study, which is quantified by the *preferential ion interaction coefficient*, Γ , defined formally by a limiting derivative and read out experimentally as a ratio of concentrations: [35, 36, 107, 108]

$$\Gamma_i^{\text{molal}} \equiv \lim_{m_{\text{NA}} \rightarrow 0} \left(\frac{\partial m_i}{\partial m_{\text{NA}}} \right)_{\mu_i} = \frac{m_i^{\text{sample}} - m_i^{\text{reservoir}}}{m_{\text{NA}}^{\text{sample}}}, \quad (5.1)$$

where m_i and m_{NA} are the molal concentrations of a particular ion species (indexed by i) and of the NA, respectively; the derivative is taken at constant μ_i , the bulk chemical potential of the ion. We will often normalize this quantity by the number of phosphates to obtain the per-nucleotide interaction coefficient, $\bar{\Gamma}_i$; this allows for generalization and

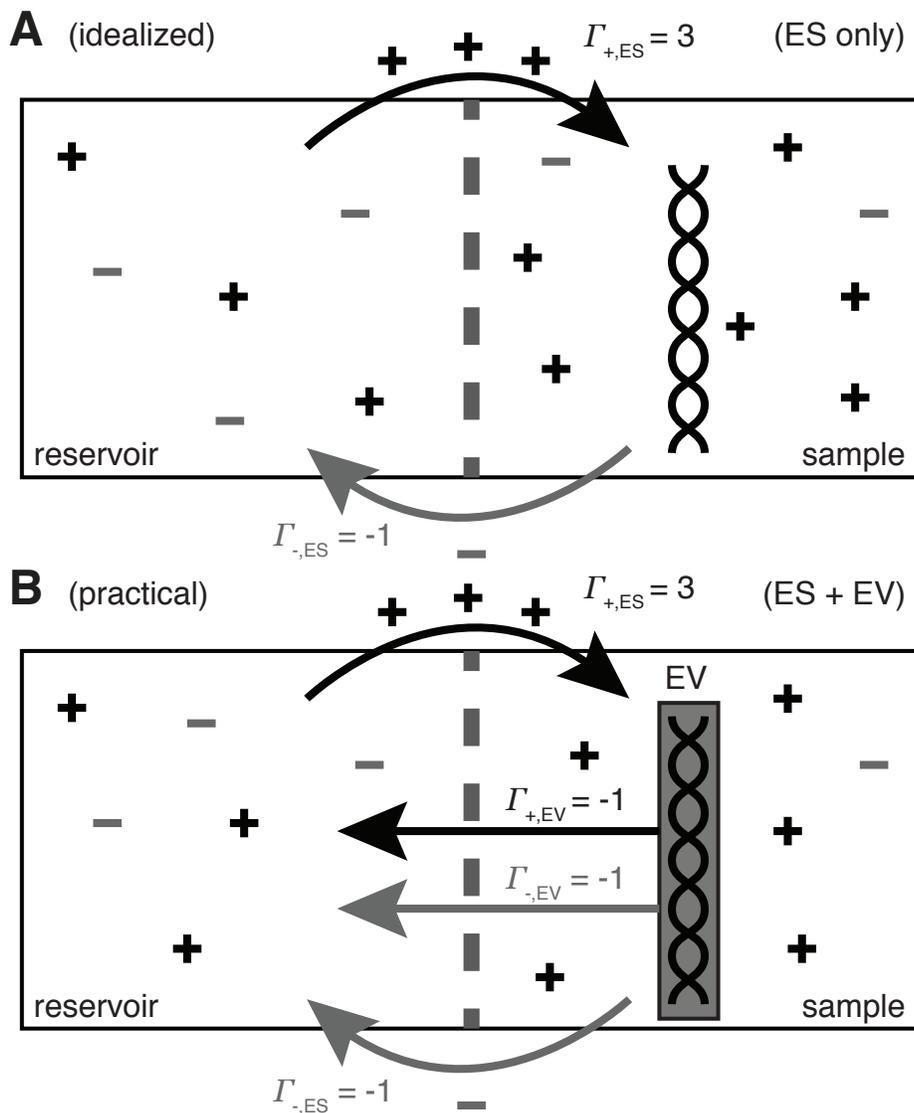


Figure 5.1: Depictions of Donnan equilibrium dialysis experiments under (A) idealized (molal) and (B) practical (molar) conditions. In both cases, an NA-containing solution is equilibrated with a bulk salt reservoir across a membrane permeable to ions but impermeable to the NA, which here has -4 charge. Preferential electrostatic (ES) interactions between the ions and the NA lead to an imbalance in ion concentration on each side of the membrane. Under practical conditions, excluded volume (EV) interactions also contribute to this imbalance and must be subtracted to reveal purely electrostatic effects.

comparison of results between NAs of different lengths. Practically, Equation 5.1 says that, in the dilute NA limit, Γ_i^{molal} is equal to the change in the number of ions in the sample compartment with respect to changes in the number of NA molecules. Since it represents the number of ions in the vicinity of the NA in excess of that expected due to the bulk salt concentration alone, Γ is also referred to as the *ion excess*. Typically, values of the cation excess, Γ_+ , will be positive and values of the anion excess, Γ_- , will be negative, as in Figure 5.1A, which corresponds to an experiment with NA charge of -4 . In this example, the NA charge is offset by the migration of three cations into, and one anion out of, the sample compartment, meaning that $\Gamma_+^{\text{molal}} = 3$ and $\Gamma_-^{\text{molal}} = -1$.

It is not generally possible, in experimental practice, to measure ion concentrations in molal units (i.e., number of ions per *solvent* volume); rather, molar concentrations, c , are typically reported (i.e., number of ions per *solution* volume, including the volume occupied by the NA). The molar analog to Equation 5.1 is written:

$$\Gamma_i^{\text{molar}} \equiv \lim_{c_{\text{NA}} \rightarrow 0} \left(\frac{\partial c_i}{\partial c_{\text{NA}}} \right)_{\mu_i} = \frac{c_i^{\text{sample}} - c_i^{\text{reservoir}}}{c_{\text{NA}}^{\text{sample}}}. \quad (5.2)$$

When the NA excluded volume (EV) is non-negligible, Γ_i^{molar} exhibits the behavior sketched in Figure 5.1B, in which the total ion excess is the sum of the electrostatic contribution, $\Gamma_{i,\text{ES}}$, and a contribution from the ions forced across the membrane due to exclusion from the occupied volume of the NA, $\Gamma_{i,\text{EV}}$. In the example of Figure 5.1B, the NA charge is again -4 and is again electrostatically neutralized by association of three cations and exclusion of one anion. However, in this case the non-zero EV of the NA also displaces one cation and one additional anion, meaning that $\Gamma_+^{\text{molar}} = 3 - 1 = 2$ and $\Gamma_-^{\text{molar}} = -1 - 1 = -2$. The EV effect—especially at high salt concentration where it is strongest—thereby obscures the electrostatic effects. Hence, it is often desirable to correct the experimental data into molal units by estimating $\Gamma_{i,\text{EV}}$ and subtracting it

from Γ_i^{molar} [38]. This correction preserves the simple interpretation of the ion excess as the number of ions surrounding the NA beyond the number expected due to the bulk concentration and allows for direct comparison with electrostatics-derived theories.

Ion excesses defined in terms of molal and molar units are related differently to the three-dimensional ion distribution, $c(\vec{r})$. In the molar picture (i.e., including EV effects), Γ_i^{molar} is equal to the volume integral of the excess ion concentration above, or below, the bulk concentration, c_{bulk} : [37, 109]

$$\Gamma_i^{\text{molar}} = \int_V dV [c(\vec{r}) - c_{\text{bulk}}]. \quad (5.3)$$

This integration is over the entire volume of the system, V . EV effects are included in this integral because $c(\vec{r}) = 0$ within the spatial extent of the NA. In the molal picture, however, EV effects are explicitly excised by restricting the integration to the solvent-accessible volume; i.e., $V \rightarrow V_{\text{solvent}}$ in Equation 5.3. These integral relations are important in that they connect the experimental output, Γ , with the typical output of most theories, $c(\vec{r})$.

5.2 Theoretical models

We now review theoretical models of the ion atmosphere at several levels of detail and computational sophistication. We first consider simple, analytically tractable theories to predict Γ in the low- and high-salt limits. We then consider numerical theories that give the full ion distribution at all intermediate salt concentrations.

5.2.1 Limiting behavior: Counterion condensation and Debye-Hückel theories

The ion excess in the low-salt limit can be estimated using Oosawa-Manning counterion condensation theory, which holds that ions will tend to condense near to the surface of a cylindrical polyelectrolyte when its linear charge density exceeds a certain threshold [21, 22, 94, 95]. This phenomenon arises from the competition between the favorable electrostatic enthalpy for ions to bind to the NA and the unfavorable entropy change associated with the loss of their translational freedom in solution. When combined with Poisson-Boltzmann theory (to be further discussed in the next section), this theory gives the following expected low- c_{bulk} limiting behavior in monovalent salt [21, 110]:

$$\lim_{c_{\text{bulk}} \rightarrow 0} \bar{\Gamma}_+ = \begin{cases} 1/2 + \zeta/4 & \zeta < 1 \\ 1 - 1/(4\zeta) & \zeta > 1 \end{cases}, \quad (5.4)$$

where $\zeta = l_B/b$ is the Manning parameter. The charge spacing of the dsNAs is estimated by projecting the phosphate positions onto the helical axis (dsDNA: 0.17 nm; dsRNA: 0.12 nm) and that of the ssNAs is estimated from the phosphate spacing of a fully elongated chain (ssDNA: 0.7 nm; ssRNA: 0.59 nm) [23]. The assumptions underlying Equation 5.4 render it inappropriate for treating higher valence ions. In these cases, we can instead consider the strong coupling limit [111], in which the entire NA charge is offset by cation association (i.e., complete condensation), leading to $\bar{\Gamma}_+ \rightarrow 1/Z$, where Z is the ion valence.

In the high-salt limit, charge screening reduces the electrostatic potential to the extent that the full Poisson-Boltzmann equation can be linearized—also known as the Debye-Hückel approximation. In this linear regime, there is no preference for cation association

versus anion exclusion and the ion excess is predicted to take on the limiting value [112, 113]

$$\lim_{c_{\text{bulk}} \rightarrow \infty} \bar{\Gamma}_+ = \frac{1}{1 + Z}. \quad (5.5)$$

Note that this high-salt limit is derived only on the basis of electrostatics, and is thus a prediction of Γ^{molal} , not of Γ^{molar} . This distinction was unimportant for Equation 5.4, since EV effects are negligible at low c_{bulk} .

The degree to which the high- and low-salt limiting laws correctly reflect experimental results is assessed in Figure 5.2, in which experimental data, corrected for excluded volume effects, are compared with the relevant limits, plotted as dotted lines. We see that Equation 5.4, or $1/Z$ in the divalent case, reasonably describes the low-salt behavior; some disagreement is expected due to the assumption of cylindrical geometry. The high-salt limit is less successful. Both dsDNA and ssDNA monovalent data (Figures 5.2A and 5.2C) do decrease towards the expected $\bar{\Gamma}_+ \rightarrow 1/2$ limit, although only weakly. In contrast, the divalent data (Figure 5.2B) increase with c_{bulk} , manifestly inconsistent with the $\bar{\Gamma}_+ \rightarrow 1/3$ limit. This disagreement indicates that the Debye-Hückel approximation does a poor job of describing the NA-ion system, even at high salt concentration. For divalents, this is a well-known result owing to the small size and high valence of the ions [114]. For monovalents, the cause of the disagreement is less obvious and may arise from finite size effects: Debye-Hückel theory is formulated for point charges, and does not, for example, consider the finite volume of an ion and its hydration shell.

5.2.2 Poisson-Boltzmann theory

More comprehensive—and complicated—theoretical models are needed to achieve quantitative precision and to map out the full salt-dependence and spatial distribution of the ion atmosphere. The first of these is *Poisson-Boltzmann* (PB) theory, which has

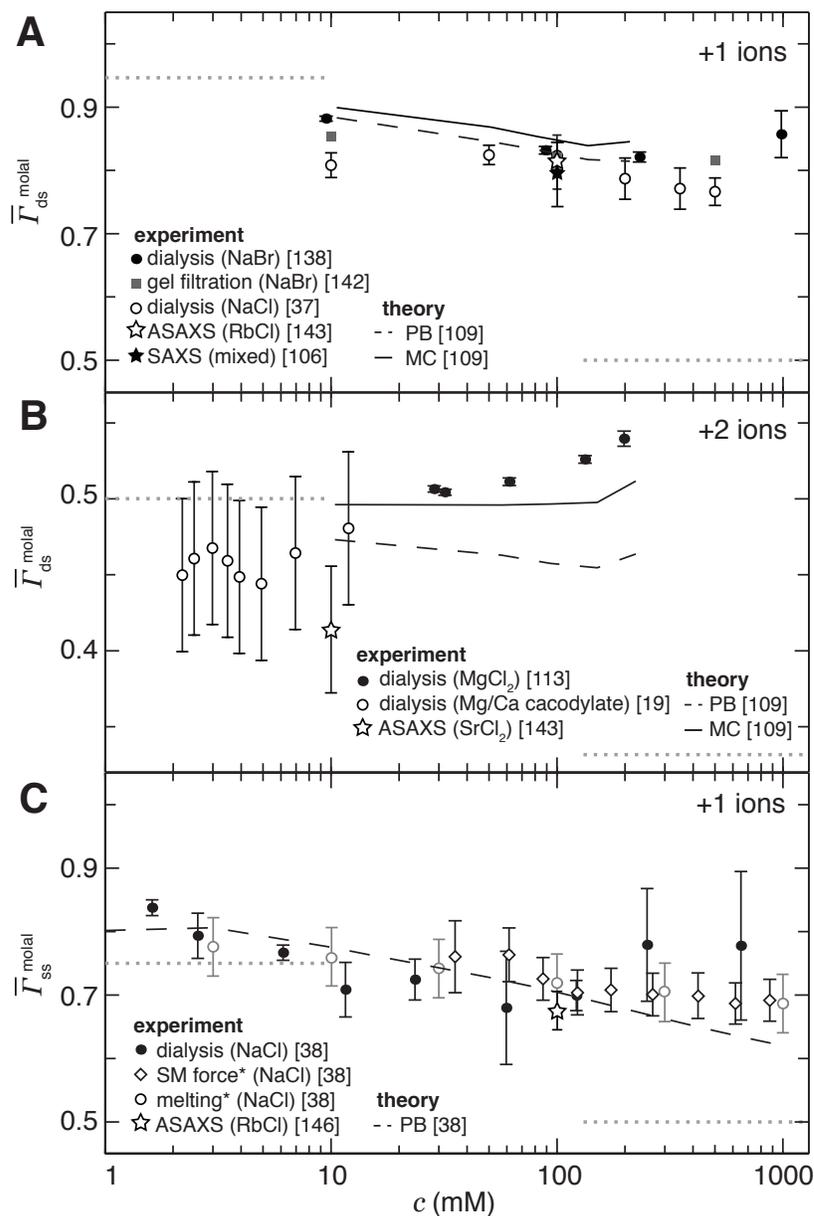


Figure 5.2: Molal cation excesses of dsDNA in (A) NaCl and (B) MgCl_2 and of (C) ssDNA in NaCl, compared with PB theory (dashed line) and the MC simulations of Ni *et al.* [109] (solid line). Molar literature values have been corrected into molal units using the method of Ref. [38] to isolate electrostatic effects. Horizontal dotted lines represent the expected limiting behavior at low- and high-salt (Equations 5.4 and 5.5); the low-salt limit assumes a cylindrical geometry. The SAXS data of Ref. [106] uses the heavy ion replacement method and is not sensitive to differences between monovalent ion species. The divalent data of Ref. [35] are extracted from a mixed $\text{Mg}^{2+}/\text{Ca}^{2+}$ competition experiment. The asterisked ssDNA studies are indirect techniques based on a thermodynamic cycle analysis.

at its core the mean-field approximation that the salt ions can be treated not as discrete particles but as a continuous charge density. Under this approximation, $c_i(\vec{r})$ is found via a Boltzmann factor relation:

$$c_i(\vec{r}) = c_{i,\text{bulk}} \exp\left(\frac{-Q_i\Phi(\vec{r})}{k_B T}\right). \quad (5.6)$$

The electrostatic potential, $\Phi(\vec{r})$, is obtained by solving the PB equation, a second-order differential equation that is analytically soluble only in a few special cases [115] and, more generally, can be handled using easy-to-use numerical solvers [116]. The PB equation is given by

$$\nabla^2\Phi(\vec{r}) = \frac{-1}{\epsilon\epsilon_0} \left[\rho_{\text{NA}}(\vec{r}) + \sum_i c_{i,\text{bulk}} q_i \exp\left(\frac{-q_i\Phi(\vec{r})}{k_B T}\right) \right], \quad (5.7)$$

where ρ_{NA} is the charge density of the NA and the summation is over all species of mobile ion in solution. Note that the Debye-Hückel approximation, used to obtain Equation 5.5, arises from linearizing the exponential in Equation 5.7.

PB calculations have been used in many studies to model the ion atmosphere of nucleic acids [101, 117–124]. The level of agreement between such studies and experimental results will be assessed below. In general, PB does a good job of describing the interactions of NAs with monovalent cations (e.g., Figures 5.2A and 5.2C), but breaks down for higher valence ions where strong ion-ion interactions render the mean-field approximation inappropriate (e.g., Figure 5.2B). Some enhancements to PB results can be obtained by accounting for the sizes of the mobile ions or for dielectric saturation effects [117, 122, 123]. The fundamental mean-field limitation is addressed in the tightly bound ion (TBI) model by treating the higher-valence ions near the NA surface using a cell-binding approach [125–128]. While the TBI model does a good job of reproducing experimental free energies of secondary structure formation [126] and ion binding com-

petition curves [127], it does not allow for the full distribution of the ion atmosphere to be obtained.

5.2.3 Simulations

Today the benchmark method for obtaining the full distribution of the ion atmosphere is atomistic molecular dynamics (MD) simulation [129]; in the past, grand-canonical Monte Carlo (MC) simulations were preferred [130]. These simulations go beyond PB by explicitly accounting for ion-ion interactions (i.e., there is no mean-field approximation) as well as other non-electrostatic effects, such as interplay with the structure of water. Unfortunately, these improvements come at a practical cost: MD simulations are vastly more computationally intensive when compared with PB, rendering impractical studies that span a wide parameter space or that probe large molecules. For this reason other more efficient models, relying on varying degrees of simplification, have been developed [131–134]. MD simulations are also susceptible to vagaries in the choice of force fields, the detailed parameters describing the microscopic molecular interactions. Different force fields have been shown to yield different ion atmosphere distributions [135]. Nonetheless, MD/MC studies are able to reproduce all existing ion counting data for which a direct comparison is possible (Figure 5.2).

5.3 Ion counting experiments

An overview of many ion counting experiments in the literature is given in Table 5.1. While the results of these studies are of interest in their own right, our focus below is on a critical analysis of the degree to which they can validate or constrain the theoretical models. In limiting our discussion to ion counting methods, we leave out other techniques that also bear on understanding the ion atmosphere. Besides ASAXS, these

Table 5.1: Overview of ion counting experiments in the literature. All studies report data on DNA samples unless denoted (R) for RNA samples or (D/R) for both DNA and RNA samples.

	ion valence		
	+1	+2	+3/higher
Γ_{ds}	Shack <i>et al.</i> (1952) [136] Strauss <i>et al.</i> (1967) [138] Shapiro <i>et al.</i> (1969) [140] Lindström <i>et al.</i> (1977) [142] Bai <i>et al.</i> (2007) [35] Pabit <i>et al.</i> (2010) [141] Kirmizialtin <i>et al.</i> (2012) (R) [105] Nguyen <i>et al.</i> (2014) [143, 144] Meisburger <i>et al.</i> (2015) [106] Gebala <i>et al.</i> (2015) [37] Gebala <i>et al.</i> (2016) [145]	Skerjanc and Strauss (1968) [113] Bai <i>et al.</i> (2007) [35] Pabit <i>et al.</i> (2010) [141] Nguyen <i>et al.</i> (2014) [143, 144]	Rubin (1977) [137] Braunlin <i>et al.</i> (1982) [139] Plum and Bloomfield (1988) [34]
Γ_{ss}	Meisburger <i>et al.</i> (2013) [146] Jacobson and Saleh (2016) (D/R) [38]	Holland and Geiger (2012) [147] Walter <i>et al.</i> (2013) [148]	
$\Delta\Gamma_{ds-ss}$	Record (1967) [149] Record (1975) [150] Bond <i>et al.</i> (1994) (D/R) [152] Williams and Hall (1996) (R) [153] Owczarzy <i>et al.</i> (2004) [154] Stellwagen <i>et al.</i> (2011) [155] Reiling <i>et al.</i> (2015) [156]	Record (1975) [150] Serra <i>et al.</i> (2002) [151]	
$\Delta\Gamma_{ds-stretch}$	Dittmore <i>et al.</i> (2014) [96] Jacobson and Saleh (2016) (D/R) [38]		Todd and Rau ^a (2008) [157]
$\Delta\Gamma_{stretch-ss}$	Landy <i>et al.</i> (2012) [75] Jacobson <i>et al.</i> (2013) (R) [26]		
competition	Bai <i>et al.</i> (2007) [35] Andresen <i>et al.</i> (2008) [158] Gebala <i>et al.</i> (2016) [145]	Bai <i>et al.</i> (2007) [35] Gebala <i>et al.</i> (2016) [145]	Andresen <i>et al.</i> (2008) [158]

^a measured $\Delta\Gamma$ between condensed and uncondensed states dsDNA

include NMR [159], gel electrophoresis [160], conductivity [161], and osmotic pressure [162] measurements.

5.3.1 Absolute ion atmosphere stoichiometry

Ion atmosphere stoichiometry experiments, like the hypothetical study of Figure 5.1B, measure the ion excess in molar units, Γ_i^{molar} , incorporating *both electrostatic and excluded volume effects*. These experiments are practically realized either as equilibrium dialyses [38, 113, 136, 138] or as forced dialyses across a centrifugal filter [35, 37]. Concentrations are typically read out by atomic emission spectroscopy [34, 35] or, alternatively, by liquid scintillation counting [140]; Γ_i^{molar} is then found by evaluation of Equation 5.2. The experimental literature describing these measurements of Γ is summarized in rows 1 and 2 of Table 5.1.

Whereas these experiments measure Γ_i^{molar} , it is actually Γ_i^{molal} that contains the purely electrostatic information we are most interested in. Therefore, in Figure 5.2, we have taken experimental literature values of $\bar{\Gamma}_+^{\text{molar}}$ for dsDNA and ssDNA in monovalent and divalent salt and applied an excluded volume correction to couch them in terms of $\bar{\Gamma}_+^{\text{molal}}$. We made this correction, as further discussed in Chapter 6, by estimating the EV of dsDNA and ssDNA using structural models from the Nucleic Acid Builder package [163] and then subtracting from $\bar{\Gamma}_+^{\text{molar}}$ the number of bulk ions contained within the NA volume, as a function of c_{bulk} .

The data of Figure 5.2 are from experiments carried out using several different species of salt. Comparative studies have shown that the identities of both the cation and anion have an effect on Γ [37]. Gebala *et al.* have recently proposed that the ion excess, especially at high c_{bulk} , depends strongly on the mean ion activity coefficient, γ_{\pm} , of the salt, especially at higher values of c_{bulk} [37].

Theoretical predictions, where available, are co-plotted in Figure 5.2. Under monovalent salt conditions, for both dsDNA (Figure 5.2A) and ssDNA (Figure 5.2C), PB calculations [38,109] do a good job of reproducing the experimental results. This is expected, as the mean-field approximation of PB is most applicable to weakly interacting monovalent ions. Following from the same argument, it is not surprising that the PB agreement breaks down for the divalent data of Figure 5.2B [109], in which ion-ion interactions are expected to become significant. In fact, in divalent salt Γ_+ increases with the bulk salt concentration, an effect at odds with the expected limiting behavior but further confirmed by studies that have observed strongly favorable Mg^{2+} binding [164,165]. Although the results of both PB and the non-atomistic MC simulations of Ni *et al.* [109] (ions modeled as atoms, but NA as a cylinder and water as a continuum) do increase at high salt, only the MC does a reasonable job of reproducing the divalent data, while retaining agreement with the monovalent.

5.3.2 Competition between ions

An extension of the absolute ion stoichiometry experiments discussed above are competition experiments, in which Γ_i values of two different ion species are measured as the relative bulk concentrations of those species are varied. Competition studies can directly probe the subtle effects arising from differences between ion species; these differences can include valence, hydrated radius, ion activity, and degree of dissociation. Many of these effects are expected to become most pronounced in the crowded, high-electrostatic-potential region directly adjacent to the NA. Thus, whereas absolute ion counting reports equally on all excess ions, competition experiments are expected to be most sensitive to the numerically fewer ions close to the NA.

A representative competition experiment is that of Bai *et al.* [35]. Many ion pairs were

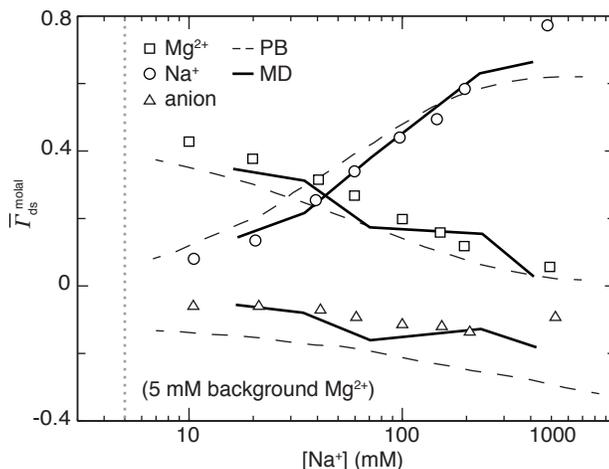


Figure 5.3: Competition experiment measuring Na^+ and Mg^{2+} ion excess as Na^+ is titrated into a background of 5 mM MgCl_2^{2+} . Compared are the predictions of PB theory (dashed lines) and the MD simulations of Yoo and Aksimentiev (solid lines). The anion used in the experimental study is cacodylate and in the MD is Cl^- ; this discrepancy likely explains the anomalous behavior at the highest $[\text{Na}^+]$. The dashed vertical line denotes the condition $[\text{Na}^+] = [\text{Mg}^{2+}]$. We have EV-corrected the experimental and MD data into molal units. Based upon Figure 7A of Ref. [35] and Figure 4C of Ref. [166].

probed in this study; in Figure 5.3 we show the results of the $\text{Na}^+/\text{Mg}^{2+}$ competition as an example. Also shown in Figure 5.3 are the predictions of two theoretical models. One is PB theory [35], which fails to capture the detailed shape of the curves across the full range of salt concentrations. The agreement is especially poor at low $[\text{Na}^+]$, where the system is dominated by the divalent Mg^{2+} ions. By contrast, the MD results of Yoo and Aksimentiev [166] (corrected for EV into $\bar{\Gamma}^{\text{molal}}$) show generally good agreement, except at the highest value of $[\text{Na}^+]$. This high- $[\text{Na}^+]$ disagreement likely arises because of anion effects: for technical reasons related to atomic emission spectroscopy, the experimental data were collected with cacodylate as the anion, whereas the simulations were based on Cl^- . Significant differences in the high-salt behavior of these two anions have been described by Gebala *et al.* [37]. Notwithstanding this complication, we see here another case in which PB theory breaks down in the presence of divalent ions, but a series of MD

simulations successfully reproduces the ion counting results. The reference interaction site model, of intermediate computational complexity between PB and MD, has also been tested by competition studies, which were unable to reproduce that theory's predicted cation size effect [145].

5.3.3 Changes with denaturation

In addition to measuring the total ion excess, Γ_i , it is also possible to directly measure differences in ion excess, $\Delta\Gamma_i$, between NA conformations. For example, one can measure the change in ion excess when a dsNA is denatured into its single-stranded state, when a dsNA is mechanically unfolded into a stretched, single-stranded state, or when a relaxed ssNA is stretched; the last of these was discussed in Chapters 3 and 4. The available transitions are summarized in the thermodynamic cycle of Figure 5.4A, and the existing experimental results are listed in rows 3–5 of Table 5.1. Like the competition experiments, $\Delta\Gamma$ measurements do not report equally on all atmospheric ions, but are most sensitive to those ions associated with structural or conformational stabilization. Beyond the information provided about the ion atmosphere stoichiometry, these measurements can also be analyzed to obtain the differential free energy of NA stabilization due to ion association, as discussed in Appendix D.

In one differential technique, NA duplex melting temperatures, T_m , are analyzed to obtain the change in ion excess between double- and single-stranded NAs, $\Delta\Gamma_{\text{ds-ss}}$ using: [167]

$$\Delta\Gamma_{\text{ds-ss}} = \frac{1}{2A\beta} \frac{N-2}{N} \frac{\partial T_m}{\partial \ln c_{\text{bulk}}}, \quad (5.8)$$

where A is an ion activity correction factor [96], N is the number of base pairs, and $\beta = R_0 T_m^2 / \Delta H^\circ$ is a compound constant measured by calorimetry of NA conformational transitions with value $\beta = 55 \pm 10\%$ [168]. Thus, concentration-dependent $\Delta\Gamma_{\text{ds-ss}}$ values

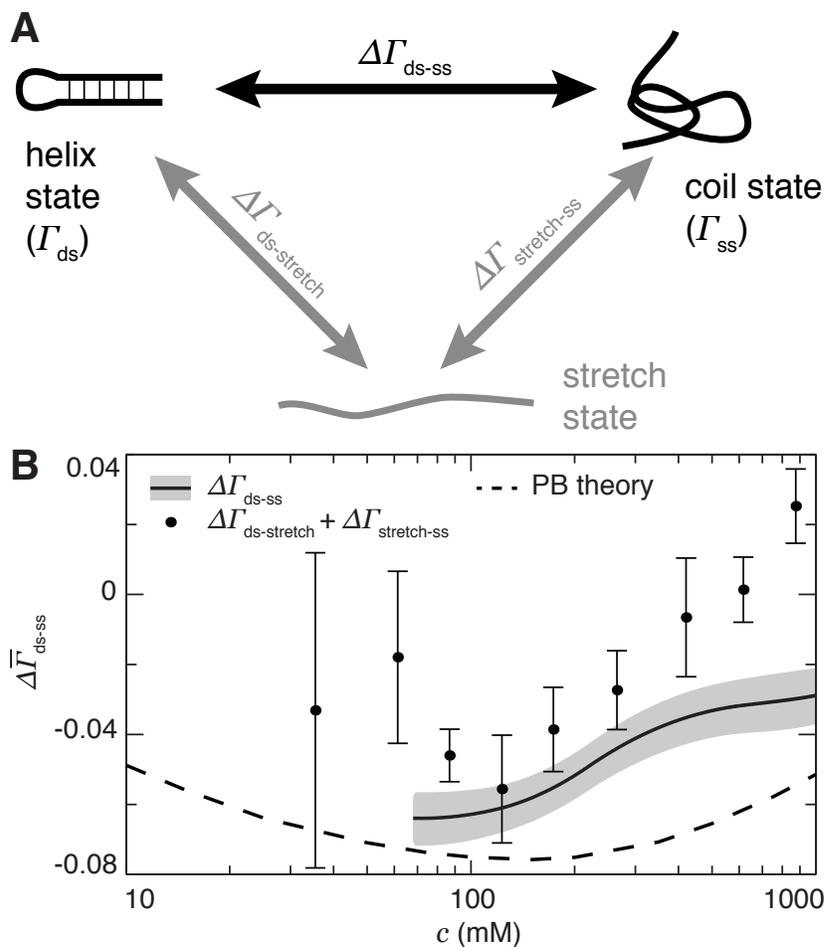


Figure 5.4: (A) Thermodynamic cycle of nucleic acid secondary structure formation. Ion counting measurements have been made of all indicated quantities: of both the folded, double-stranded and unfolded, single-stranded species (Γ_{ds} and Γ_{ss} , respectively) and of changes in ion atmosphere between the states, denoted $\Delta\Gamma_{i,j}$. (B) Comparison of $\Delta\Gamma_{ds-ss}$ measurements of DNA made directly by melting experiments (line) [38, 154] and indirectly by single-molecule force spectroscopy (points) [38, 75, 96] with a PB prediction based on the difference of the dsDNA and ssDNA curves from Ref. [38]. Adapted from Figures 1 and S7 of Ref. [38].

can be obtained in an experiment that, unlike dialysis, does not involve a membrane and can be performed at arbitrarily dilute NA concentrations. Such measurements have been made for DNA in both monovalent [149, 150, 152, 153, 155, 156] and divalent [150, 151] salt as a function of structural parameters and ion species. In the case of NaCl, a rich dataset amenable to this analysis is that of Owczarzy *et al.* [154].

More recently, a second related technique has been developed that measures the change in ion excess associated with mechanically stretching an NA, $\Delta\Gamma_{\text{ds-stretch}}$ or $\Delta\Gamma_{\text{stretch-ss}}$, by analyzing the salt dependence of the equilibrium unfolding force from single-molecule force spectroscopy experiments [30]. Similar arguments have also been applied to the analysis of simulated RNA pseudoknot unfolding data by Hori *et al.* [169]. Several thermodynamic relations can be derived to analyze the data (see Appendix D), one of which, applicable to an NA hairpin system, is given by [30, 96, 157]

$$\Delta\Gamma_{\text{ds-stretch}} = \frac{\Delta X}{2k_B T A} \left(\frac{\partial f_{1/2}}{\partial \ln c_{\text{bulk}}} \right), \quad (5.9)$$

where ΔX is the change in molecular extension as the hairpin construct is unfolded and $f_{1/2}$ is the equilibrium mechanical force of unfolding. This equation is valid only in monovalent salt, but similar relations can be derived for other valences [30]. Note that this relation has a similar form to Equation 5.8, but does not require an independently measured quantity analogous to β . Also, the single-molecule nature of these studies allows the realization of the $c_{\text{NA}} \rightarrow 0$ limit in the formal definition of the ion excess (Equation 5.2); i.e., there is not a concern about NA-NA interactions. $\Delta\Gamma_{\text{stretch-ss}}$ data can likewise be obtained by applying another ion counting relation to force spectra of intrinsically single-stranded samples [26, 75]. These data can be combined with $\Delta\Gamma_{\text{ds-stretch}}$ to complete the thermodynamic cycle of Figure 5.4A and obtain a measure of $\Delta\Gamma_{\text{ds-ss}}$ equivalent to that from the melting studies.

Results for both melting and mechanical stretching measurements of $\Delta\Gamma_{\text{ds-ss}}$ for 25 base pair DNA helices are shown in Figure 5.4B [38]. We see that the two methods give salt-dependent values with the same general shape and trend, but that they do not agree quantitatively within uncertainty. Also plotted is a PB estimate of $\Delta\Gamma_{\text{ds-ss}}$ found by taking the difference between the PB results for dsDNA and ssDNA in Ref. [38]. Even though those two PB results are, separately, in very good agreement with the absolute ion excess data, their difference is in poor agreement with the results of Figure 5.4B. Thus, we see that $\Delta\Gamma$ measurements provide an additional, and perhaps more stringent, test of the theories. Unfortunately, a lack of simulation-based studies of single-stranded NAs precludes comparing these differential data with MD or MC predictions.

5.4 Discussion

We have described several types of ion counting experiments and how they can be used to critically, but not comprehensively, test theoretical models of the ion atmosphere. We have seen that PB theory reproduces most of the monovalent ion counting data (except for the differential data of Figure 5.4B), but disagrees notably with the divalent $\Gamma_+(c)$ data of Figure 5.2B and the mixed monovalent/divalent competition data of Figure 5.3, consistent with the limitations of the mean-field approximation underlying that theory.

In contrast, MD and MC calculations are in agreement with all of the ion counting results we have surveyed, both monovalent and divalent (Figures 5.2 and 5.3). For this reason, MD simulations are generally considered to be a “gold standard” in the field and their predictions of the three-dimensional ion distribution are used as benchmarks against which to compare similar results from other, simpler models [101,124]. However, there are two caveats to the general acceptance of MD results. First, while MD and ion counting experiments are in agreement in all cases where both simulation and experimental data

are available, nothing approaching all possible pairings (Table 5.1) of the two have been explored. As one example, to our knowledge no ssDNA simulations have been performed that would allow a comparison with the differential data of Figure 5.4. Second, even if MD simulations were fully validated against all possible ion counting experiments, that is not a guarantee of the accuracy of the full three-dimensional distributions. This is illustrated by comparing the MD and reference interaction site model simulations of Giambaşu *et al.* [124]: while both simulation schemes give similar values of Γ_+ , and are thus equally in agreement with experiment, their predicted radial ion distributions disagree markedly (compare Figures 5 and 6 of that paper). It may be that MD force fields, in the current state of the art, are adequate for reproducing the ion excess but fail to capture certain subtleties of the full ion distribution.

Thus we see that ion counting experiments have the power to validate and constrain theoretical models of the ion atmosphere, but that, at present, they cannot fully test the most detailed predictions of the most sophisticated models. Some progress can be made, as mentioned above, by filling in the holes in the existing ion counting corpus—the gaps in Table 5.1—and by conducting matching theoretical analyses. Beyond that, a great leap in understanding would accompany the development of an experimental technique capable of measuring the full, three-dimensional distribution of the ion atmosphere. ASAXS experiments [104] have come the closest to this goal, in that they give information about the distribution over moderate to long length scales. New or improved techniques will be needed, however, to probe the short length scales where subtle perturbations in the ion atmosphere could have large impacts on biomolecular interactions.

Chapter 6

Measurements of the ssNA ion atmosphere

Portions of this chapter are adapted from Ref. [38], which was released under the terms of the Creative Commons Attribution License. Copyright 2016, David R. Jacobson and Omar A. Saleh.

To form secondary structure, nucleic acids must overcome electrostatic strand-strand repulsion, which is moderated by the surrounding atmosphere of screening ions. The free energy of NA folding therefore depends on the interactions of this ion atmosphere with both the folded and unfolded states. We quantify such interactions using the preferential ion interaction coefficient, or ion excess: the number of ions present near the NA in excess of the bulk concentration. The ion excess of the folded, double-helical state has been extensively studied (as reviewed in Chapter 5); however, much less is known about the salt-dependent ion excess of the unfolded, single-stranded state. We measure this quantity using three complementary approaches: a direct approach of Donnan equilibrium dialysis read out by atomic emission spectroscopy and two indirect approaches involving either

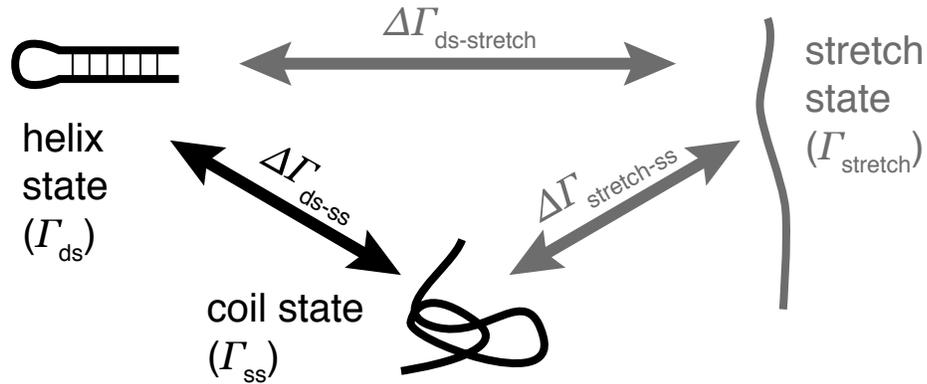


Figure 6.1: Thermodynamic cycle diagram for mechanical unfolding of an NA hairpin. The helix state refers to the intact hairpin, the stretch state to the hairpin after being mechanically unfolded by an applied force, and the coil state to the unfolded hairpin in the absence of force. Each state is characterized by a salt-dependent preferential cation interaction coefficient (ion excess), Γ_i , and each transition is associated with a change in ion excess, $\Delta\Gamma_{i-j}$.

single-molecule force spectroscopy or existing thermal denaturation data. The pathways that these different approaches use to arrive at the final result for single-stranded NAs, Γ_{ss} , are summarized in the thermodynamic cycle of Figure 5.4. In particular, the dialysis method constitutes a direct measurement of Γ_{ss} , whereas the other, differential measurements begin with the known results for the double-stranded state, Γ_{ds} and then subtract changes between states. The results of these three independent approaches are in good agreement. Even though the single-stranded NAs are flexible polymers that are expected to adopt random-coil configurations, we find that their ion atmosphere is quantitatively described by rod-like models that neglect large-scale conformational freedom, an effect that we explain in terms of the competition between the relevant structural and electrostatic length scales.

6.1 Results

6.1.1 Coil ion excess measured by D-AES

We measured the ion excess of the unfolded coil state, $\Gamma_{\text{ss}}^{\text{molar}}$, directly using dialysis and atomic emission spectroscopy (D-AES). This method relies on the phenomenon of Gibbs-Donnan equilibrium [33], which was discussed in Section 5.1.2. Attractive interactions with cations, repulsive interactions with anions and excluded volume effects all lead to perturbations in the molar ion concentrations across the dialysis membrane. These perturbations, which we measure using the atomic emission spectroscopy methods described in Appendix A.3, yield $\Gamma_{\text{ss}}^{\text{molar}}$ through Equation A.6. In these experiments we used 50 nucleotide NA homopolymers that do not exhibit appreciable secondary structure formation or base stacking [69, 170]. D-AES results for $\bar{\Gamma}_{\text{ss}}^{\text{molar}}$ of both dT₅₀ DNA and rU₅₀ RNA are shown in Figure 6.2. There is known to be a length-dependence to the NA ion excess [167, 171]; as such, the results we present can only be extrapolated to NAs of other lengths through appropriate correction.

As discussed in Chapter 5, different values of Γ are obtained when the ion concentrations are measured in either molal or molar units. In particular, ion excess values derived from molal units reflect only the electrostatic (ES) interactions, which are of principal interest, while values derived from molar units also include a contribution from the NA excluded volume (EV). Dialysis experiments naturally give concentrations measured in molar units, so the raw data plotted in Figure 6.2 include both ES and EV contributions. The nature of the EV contribution to the high-salt behavior—that it asymptotes to a linear decrease in Γ^{molar} —is most clearly seen in the linear-linear plot of Appendix Figure B.4. Later, in Figure 6.5, we estimate and subtract the EV contribution to give ion excesses that only incorporate ES effects; this same correction was applied to the reviewed literature results of Chapter 5.

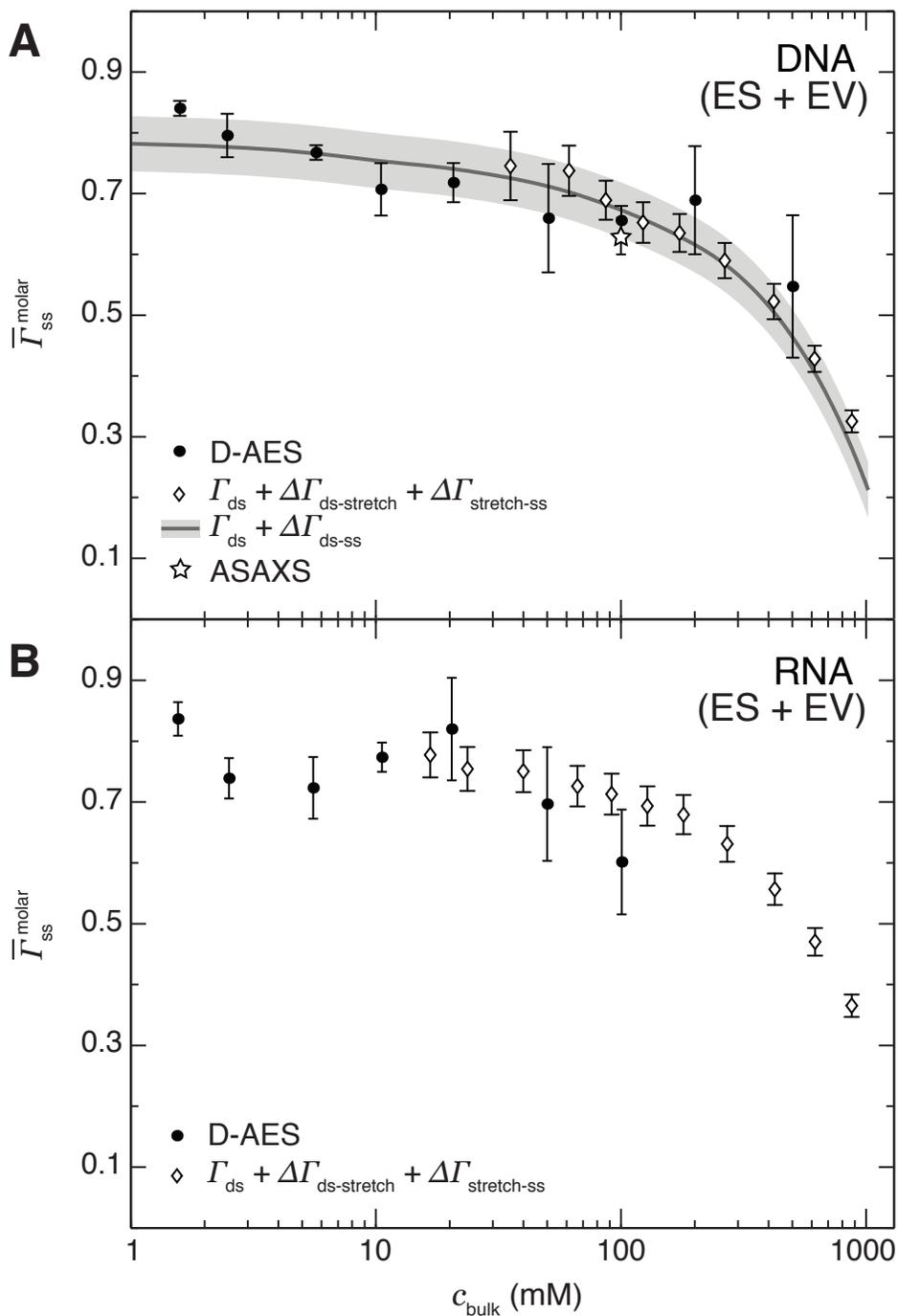


Figure 6.2: Values of $\bar{\Gamma}_{ss}^{\text{molar}}$ (i.e., including both ES and EV effects) for (A) DNA and (B) RNA, determined directly by D-AES (solid points) and indirectly by thermodynamic cycle analysis using PB calculations and either single-molecule experiments [26, 75, 96] (open diamonds) or melting experiments [154] (line/shaded region). Also plotted is the DNA ASAXS result of Meisburger *et al.* [146] in RbCl (star).

6.1.2 PB calculations

We obtain Γ_{ds} , the ion excess of the folded, helical state, using numerical solutions to the Poisson-Boltzmann equation on model A-form (RNA) and B-form (DNA, Figure 6.3A) double helices under various bulk salt concentrations. In using the canonical double-helix geometries, we assume that the helices do not undergo significant conformational fluctuations over time or significant conformational changes between salt concentrations. Molecular dynamics simulations of dsDNA, which show no clear trend in various structural parameters with salt [172], support this assumption. In using the PB equation, we assume that ion-ion effects are sufficiently weak that the ion density can be treated as a continuous mean field, which is well-established in the case of monovalent ions [173]. Agreement with experimental data further validates the PB. Figure 6.3C shows the results of our PB calculations, in molar units, for dsDNA and dsRNA as solid lines; co-plotted are the results of DNA buffer equilibration/atomic emission spectroscopy studies that show reasonable agreement with the PB [37]. That there is some disagreement, especially at high-salt, could arise from the non-ideal solution behavior of NaCl; Gebala *et al.* have shown that PB calculations best reproduce experimental ion counting data for ion species with the highest mean ion activity coefficients [37]. We use the average extent of this disagreement as an estimate of the PB error.

The dashed lines in Figure 6.3C show results of similar PB calculations for simple models of the ssNAs (i.e., the dsNA models with one strand removed, Figure 6.3B). We do not expect such models to truly capture the physics of the single-stranded system, as the flexible ssNAs exist in a conformational ensemble that is unsampled by these models. Shkel and Record have successfully used such modeling to explain some ssDNA properties [174]. These naive models give a handle on certain local aspects of the problem: the effects of close co-localization of strands and of the charge spacing along the backbone.

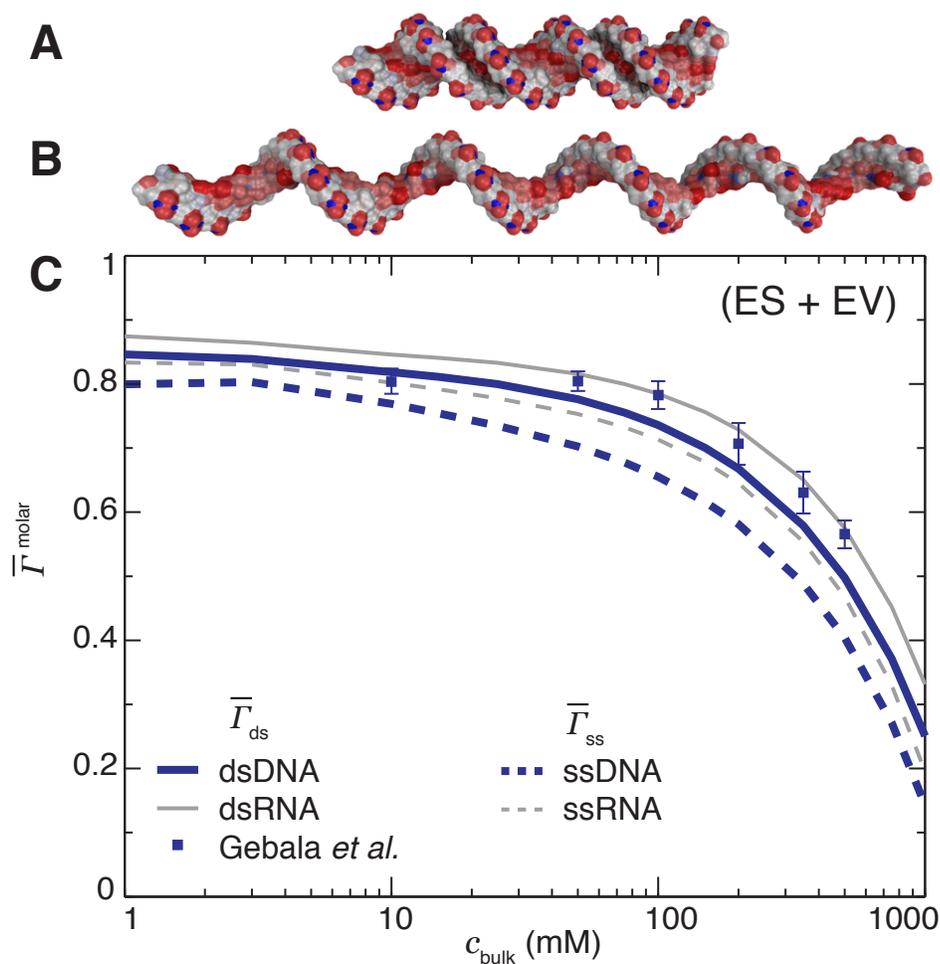


Figure 6.3: Structural model of dsDNA (A) and naive model of ssDNA (B) used in PB calculations, each having 50 nucleotides. Equivalent models for RNA were also used. (C) Ion excess of dsDNA (solid blue) and dsRNA (solid gray) in the molar ensemble, calculated by numerical evaluation of the PB equation as a function of bulk salt concentration. The use of PB to estimate $\bar{\Gamma}_{\text{ds}}$ is validated by comparison with the experimental results of Gebala *et al.* [37] (blue squares). Similar PB results for models of the ssNA structure neglecting any conformational freedom are plotted for DNA (dashed blue) and RNA (dashed gray).

We see the first effect by comparing the dsNA and ssNA curves in Figure 6.3C: for a particular species and at a particular salt concentration, the ion excess of the ssNA is slightly lower than that of the dsNA (approximately 3-6% for DNA, 3-10% for RNA). This reduction is expected based on the reduced charge density of the single-stranded state, which puts the system closer to the Debye-Hückel limit. We see the second effect, that of the charge spacing along the backbone, by comparing the DNA and RNA curves. Because DNA has a greater charge spacing along the sugar-phosphate backbone than RNA (0.70 nm versus 0.59 nm) [23], it has a smaller local charge density and is thus also closer to this limit.

6.1.3 Single-molecule measurement of $\Delta\Gamma_{\text{ds-stretch}}$

We use single-molecule magnetic tweezer experiments [92] to measure the mechanical force needed to unfold an RNA helix as a function of bulk salt concentration. The experimental geometry is sketched in Figure 6.4A and the method used to determine the equilibrium unfolding force, $f_{1/2}$, is illustrated in Figure 6.4B. We then apply a thermodynamic relation [30] to these $f_{1/2}(c_{\text{bulk}})$ data to extract the change in ion excess associated with this unfolding, $\Delta\Gamma_{\text{ds-stretch}}$. This analysis is more completely described in Appendix A.2.6 and the thermodynamic arguments described in Appendix D and Dittmore *et al.* [96].

Figure 6.4C plots the equilibrium unfolding force, $f_{1/2}$, for the RNA helix. We see that $f_{1/2}$ increases linearly with $\ln c_{\text{bulk}}$ over low-to-moderate salt concentrations and flattens off as c_{bulk} approaches 1 M. Equation 5.9 is applied to these data to obtain $\Delta\bar{\Gamma}_{\text{ds-stretch}}$, which is non-constant across most of the salt range studied (Figure 6.4D); similar results for DNA [96] are co-plotted. In both cases ions are released when the NA helix unfolds and, at sufficiently high salt, the number released decreases with increasing

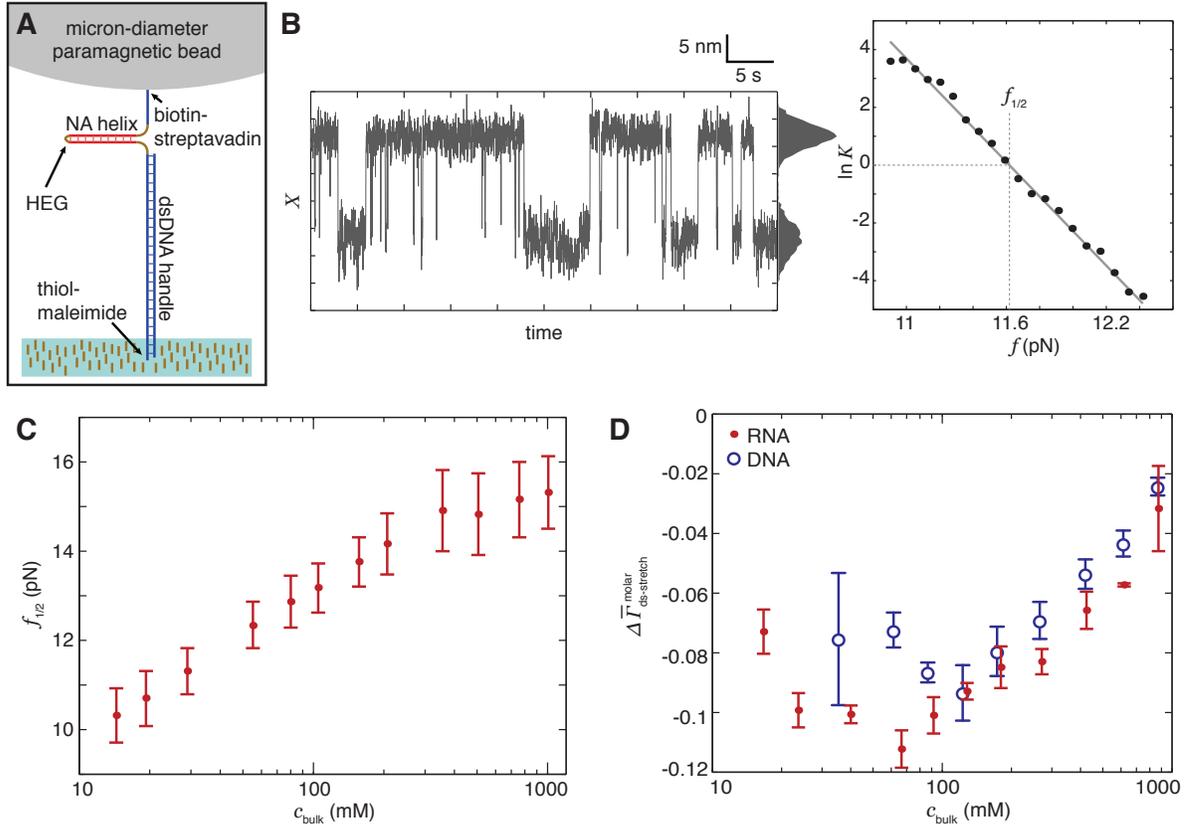


Figure 6.4: Single-molecule magnetic tweezers experiment to measure the change in ion excess, $\Delta\bar{\Gamma}_{\text{ds-stretch}}$, when an RNA hairpin is mechanically unfolded. (A) Diagram of the molecular construct. (B) Sample trace showing hopping between two conformational states, corresponding to two molecular extensions, X , and extraction of the corresponding equilibrium constant, K . The unfolding force, $f_{1/2}$, is then obtained from a plot of $\ln K$ versus applied force. (C) Plot of $f_{1/2}$ as a function of bulk NaCl concentration for the RNA hairpin. (D) $\Delta\bar{\Gamma}_{\text{ds-stretch}}^{\text{molar}}$ for an RNA hairpin (red points) and previously reported [96] values for a DNA hairpin (blue circles) using the generalized Clausius-Clapeyron method. Error bars reflect standard errors of the mean.

in c_{bulk} . This is consistent with a screening length argument. At high salt the Debye length becomes short enough that all charges on the macromolecule are electrostatically isolated; thus, the differences in charge arrangement between the helix and stretch states become unimportant and the difference in ion excess between the states, $\Delta\bar{\Gamma}_{H-S}$, goes to zero.

6.1.4 Completing the thermodynamic cycle

Starting from the Γ_{ds} results of Figure 6.3C, an indirect measurement of Γ_{ss} can be reached using either of two routes through the thermodynamic scheme of Figure 6.1: via single-molecule stretching experiments, in which

$$\Gamma_{\text{ss}} = \Gamma_{\text{ds}} + \Delta\Gamma_{\text{ds-stretch}} + \Delta\Gamma_{\text{stretch-ss}}, \quad (6.1)$$

or via oligonucleotide melting studies, in which

$$\Gamma_{\text{ss}} = \Gamma_{\text{ds}} + \Delta\Gamma_{\text{ds-ss}}. \quad (6.2)$$

Results can be obtained in either of the molar or molal ensembles by using Γ_{ds} values integrated in the appropriate way from the PB calculations, as in Equation 5.3.

Evaluation of the mechanical stretching method (Equation 6.1) requires Γ_{ds} and $\Delta\Gamma_{\text{ds-stretch}}$, obtained as discussed above, and also $\Delta\Gamma_{\text{stretch-coil}}$, the change in the ion excess as the stretched state is relaxed to a random coil without refolding. Prior studies have reported $\Delta\Gamma_{\text{stretch-coil}}$ data for homopolymers of DNA [75] and RNA [26] that do not form secondary structure; the RNA study is discussed in Section 3.1.3. These homopolymers were on the order of 1000 nucleotides in length and may exhibit length-dependent effects when compared with the 50-mers we use in the present studies. An

exact correction for such effects is not known to us, but would have only a minor effect: $\Delta\Gamma_{\text{stretch-ss}}$ contributes only a few percent to the value of Γ_{ss} . Also, whereas the previously reported $\Delta\Gamma_{\text{stretch-coil}}$ values (Figure 3.3) were obtained by fitting the data with a surface having a particular functional form, analyzing the data in a way that does not depend on the choice of model function leads us to conclude that, within uncertainty, $\Delta\Gamma_{\text{stretch-ss}}$ is salt-independent. As such, we arrive at per-nucleotide $\Delta\bar{\Gamma}_{\text{stretch-ss}}$ values of 0.027 ± 0.003 for RNA and 0.043 ± 0.005 for DNA. Using these results, we can then evaluate Equation 6.1 to obtain the first indirect evaluation of $\bar{\Gamma}_{\text{ss}}$, which is plotted in molar units alongside the D-AES results in Figure 6.2.

Alternatively, we can complete the thermodynamic cycle using Equation 6.2, which relies on the PB results for Γ_{ds} and also $\Delta\Gamma_{\text{ds-ss}}$ data from oligonucleotide melting experiments. We are unaware of suitable data for RNA melting, so we confine our analysis by this method to DNA. Owczarzy *et al.* [154] have reported melting temperatures for DNA oligonucleotides as a function of length, base composition and salt concentration. We analyzed all of the 25 base pair data in their dataset using Equation 5.8. Performing this analysis, the details of which are given in Appendix A.5, gives $\Delta\Gamma_{\text{ds-ss}}$ over $[\text{NaCl}] = 69\text{-}1025$ mM. We extrapolate to lower salt concentrations by assuming constant $\Delta\Gamma_{\text{ds-ss}}$, an assumption that is supported by other melting experiments [168]. Finally, we evaluate Equation 6.2 to obtain the second indirect measure of $\bar{\Gamma}_{\text{ss}}$, which is also plotted, in molar units, in Figure 6.2A.

6.2 Discussion

Despite being based on different experimental techniques, the various methods discussed above agree on a consistent trend in the salt dependence of the ion excess of random-coil, single-stranded DNA and RNA (Figure 6.2). The linear decrease in molar

ion excess at high salt (see plot with linear axes, Appendix Figure B.4) is due to excluded volume effects; electrostatic effects, which are of greatest interest to understanding nucleic acid-ion interactions, manifest as perturbations to this behavior. We can isolate the electrostatic effects by correcting the data into molal units. This is done, for the thermodynamic cycle measurements, by using the molal values of Γ_{ds} from the PB and, for the D-AES measurements, by adding an EV correction. This correction is obtained by multiplying the excluded volume of the NA—estimated from the structural model of Figure 6.3B incorporating a hydration layer—by the bulk salt concentration. Such EV-corrected, molal data are plotted in Figure 6.5, where we now see that the data lie, roughly, between the expected low-salt Oosawa-Manning ($\bar{\Gamma} \rightarrow 0.75\text{--}0.79$, assuming thin-cylinder geometry) and high-salt Debye-Hückel ($\bar{\Gamma} \rightarrow 0.5$) limits. Below, we explore electrostatics models to quantitatively reproduce the $\Gamma_{\text{ss}}^{\text{molal}}(c_{\text{bulk}})$ curves between these limits.

6.2.1 Uniform mean-field model

Because the single-stranded nucleic acids are so flexible and exist in an ensemble of random-coil conformations, we hypothesized that the chain would act as a uniform charge distribution, with the charge density given by the average over all configurations. In a simple model of this picture, we treat the system as consisting of two compartments in Gibbs-Donnan equilibrium with each other: one a bulk reservoir and the other a region of uniform charge density, ρ . This charged region represents the NA coil, ignoring any effects that may arise from the shape of the NA or the discreteness of its charge. Following the derivation given in Appendix C, we solve for the per-nucleotide cation excess of this

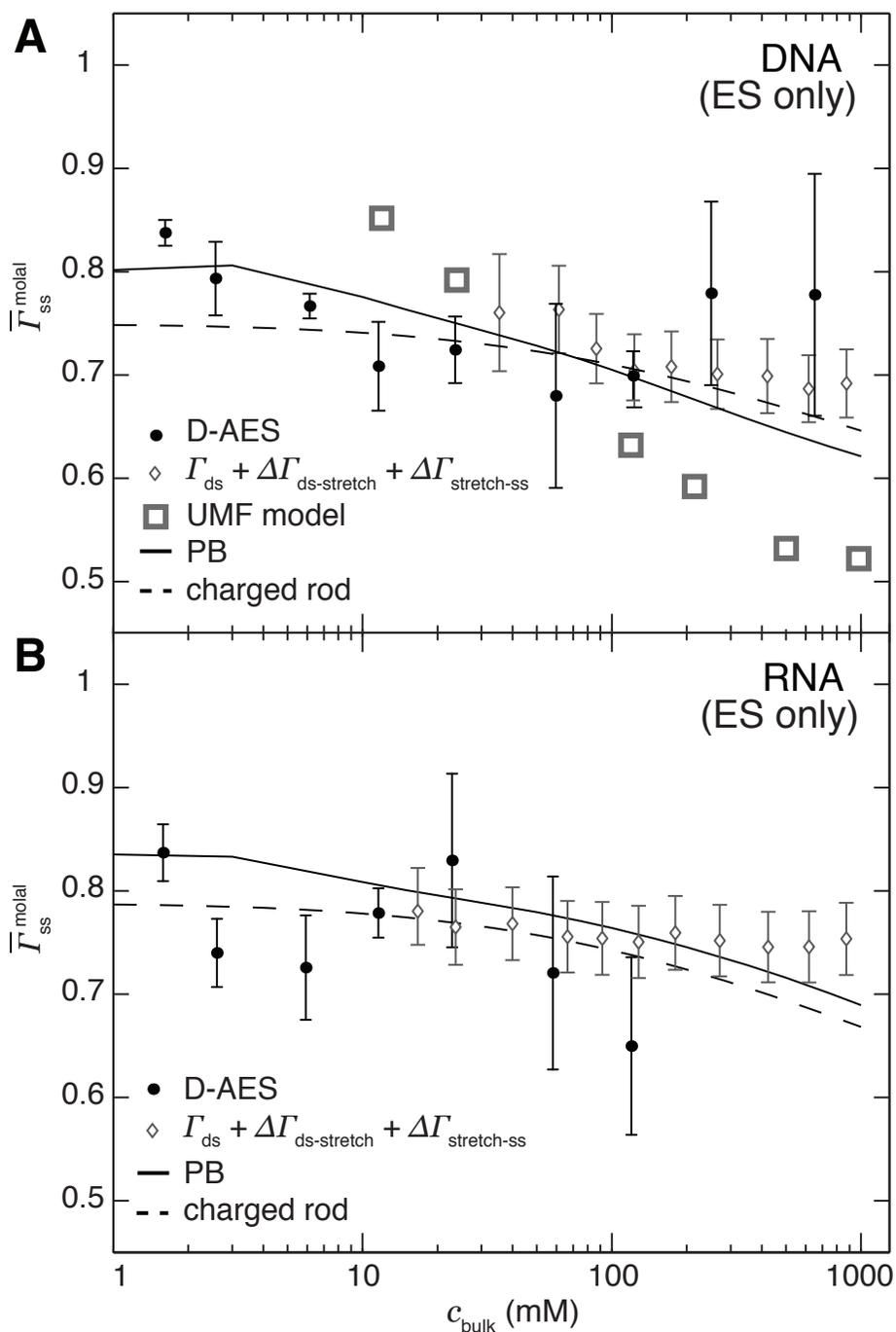


Figure 6.5: Comparison of uniform mean-field model (open squares), naive PB calculations (solid line), and charged rod model (dashed line) to D-AES (points) and single-molecule (open diamonds) data for (A) DNA and (B) RNA. Shown are values of $\bar{\Gamma}_{ss}^{\text{molal}}$, i.e., corrected for the excluded volume effect to emphasize electrostatics.

model of the coil state:

$$\bar{\Gamma}_{\text{ss}} = \frac{1}{2} - \sqrt{\frac{1}{4} + \left(\frac{c_{\text{bulk}}e}{\rho}\right)^2} - \frac{c_{\text{bulk}}e}{\rho}, \quad (6.3)$$

where ρ is the absolute value of the uniform volume charge density.

It is possible to connect this model to physical NA coils via the charge density parameter. If we treat the NA coil as a sphere, with charge Q and radius R , then $\rho = \frac{Q}{(4/3)\pi R^3}$ or, using the definition of the radius of gyration, R_g ,

$$\rho = \frac{Q}{\frac{4}{3} \left(\frac{5}{3}\right)^{\frac{3}{2}} \pi R_g^3}. \quad (6.4)$$

Salt-dependent R_g values for dT₅₀ have been measured using small-angle x-ray scattering by Sim *et al.* [175]; we are not aware of similar measurements for RNA. We can insert these R_g values into Equation 6.4 and, in turn, into Equation 6.3 to obtain semi-theoretical predictions of $\bar{\Gamma}_{\text{ss}}$ based on the uniform mean-field (UMF) model and the experimental R_g values. These predictions are co-plotted with the molal data in Figure 6.5A.

This simple, uniform mean-field model exhibits $\bar{\Gamma}_{\text{ss}} \rightarrow 0.5$ high salt limiting behavior (Appendix Figure C.2); however, its quantitative agreement with the experimental data is poor. Our hypothesis of dominant conformational flexibility, leading the system to behave as if it is uniformly charged, is thus not borne out. We require an alternative model, including some non-uniformity in the charge distribution, to explain the data.

6.2.2 Non-uniform mean-field models

Because all of our experiments are carried out in the presence of added salt, electrostatic interactions will be appreciable only on length scales shorter than the Debye screening length. The conformational flexibility that is central to the UMF model occurs

only on long length scales compared with κ^{-1} . As such, we now hypothesize that it is the short-length-scale (i.e., sub-Debye length) charge distribution of the system, which is rod-like, that dominates the interactions with the ion atmosphere.

One way to account for the local charge distribution in the ssNA coil is through the single-stranded PB calculations of Figure 6.3C, which are based on a naive structural model (Figure 6.3B) that does not sample the available conformations of the chain but does account for the discreteness of the source charge and its spacing along the backbone. Because this naive model is based on the canonical double-helical structure, it incorporates some degree of base-stacking, not present in the random-coil dT₅₀ or rU₅₀, that may introduce error into the modeling. Figure 6.5 compares these PB results with the EV-corrected DNA and RNA data, showing good agreement across the full range of salt concentration in both cases.

An alternative, structurally simpler model lacking discrete charges is that of Landy *et al.* [176], in which the NA is treated as a charged rod with radius R . In the low ionic strength, excess salt limit, this model predicts that the cation excess of the negatively charged rod is given by

$$\bar{\Gamma} \approx \frac{\zeta - 1/2}{2\zeta} \left[1 - \left(\frac{\mathcal{K}_0(\kappa R)}{\mathcal{K}_1(\kappa R)} \right)^2 \right] + \frac{1}{2}, \quad (6.5)$$

where ζ is the Manning parameter [21] ($\zeta = 1$ for ssDNA, $\zeta = 1.19$ for ssRNA) and the \mathcal{K}_i are the modified Bessel functions of the second kind; this equation is valid for $\zeta \geq 1$. For the ssNAs, the choice of R is non-obvious. Here, for both RNA and DNA, we set $R = 0.24$ nm: the distance of closest approach of Na⁺ ions to the backbone phosphates in molecular dynamics simulations of dsRNA [101]. As shown in Figure 6.5, the charged rod model also agrees with the data except at the very lowest salt concentrations.

That both the single-stranded PB and charged rod models agree with the EV-

corrected $\bar{\Gamma}_{\text{ss}}^{\text{molal}}$ data indicates that the monovalent ion excess of single-stranded nucleic acids is insensitive to their large-length-scale conformation. Rather, it is the spacing of the charges along the strand that dominates. This effect is seen by comparing the $\bar{\Gamma}_{\text{ss}}$ results for RNA (0.59 nm spacing along sugar-phosphate backbone) and DNA (0.70 nm spacing) in Figures 6.2 and 6.5. That the ion excess is insensitive to the conformation of the NA is consistent with the small magnitude of $\Delta\bar{\Gamma}_{\text{stretch-ss}}$, the change in ion excess between two states that differ, primarily, in their conformation (i.e., extended versus coiled). It is also consistent with our hypothesis that, due to electrostatic screening on length scales larger than the Debye length, the interactions between the ion atmosphere and the NA are dominated by the rod-like charge distribution on short length scales.

6.3 Conclusion

We have presented new D-AES measurements of the absolute number of ions associated with unfolded DNA and RNA oligonucleotides, single-molecule measurements of the change in number of ions associated with an RNA hairpin as it is mechanically unfolded, and PB calculations of the absolute number of ions associated with double-helical and single-stranded DNA and RNA. Using these results, along with existing data on a number of nucleic acid systems, we have obtained three different measurements of the monovalent ion atmosphere, Γ_{ss} , of unfolded, single-stranded DNA and two different measurements of unfolded RNA. We have also reported measurements of the change in ion excess as an unfolded NA folds to form secondary structure. These $\Delta\Gamma_{\text{ds-ss}}$ results, which show non-trivial salt dependence (Figure 5.4B), contribute to understanding the free energy change associated with secondary structure formation.

We interpret the observed salt-dependent Γ_{ss} behavior as an interpolation between two limiting regimes: a low-salt limit described by Oosawa-Manning counterion conden-

sation theory and a high-salt, Debye-Hückel limit in which the charge is equally offset by cation association and anion depletion. The detailed behavior at intermediate salt concentrations depends on the strandedness and charge spacing of the NA, but does not depend appreciably on its large-scale conformation, presumably due to the short-range nature of the screened electrostatics. This indicates that, when considering free energies of ion interactions in the context of nucleic acid structure formation, much of the molecular complexity on large length scales can likely be ignored. In other words, our finding of a predominately local origin of the ion excess confirms the picture from Chapter 5: that the ion excess, while often the best experimental metric that we have, is not optimally sensitive to the detailed interactions between ions and nucleic acids.

Chapter 7

Outlook

In this dissertation, I have discussed the application of two principal experimental techniques—single-molecule force spectroscopy and ion counting—to the problem of understanding nucleic acid-ion interactions. I have shown that the elastic behavior of single-stranded nucleic acids over 0.1–100 pN applied force in monovalent salt occurs in two regimes: a low-force Pincus regime characteristic of excluded volume interactions and a higher-force, quasi-logarithmic regime characteristic of the interplay between intrinsic electrostatic tension and the applied force. The transition between these two regimes varies as a function of salt concentration and scales with the Debye screening length. Using ion counting experiments, I have shown that the quantitative composition of the ion atmosphere, again in monovalent salt, varies only gradually with bulk salt concentration, and is consistent with numerical Poisson-Boltzmann models incorporating only local structure. This is consistent with electrostatic interactions that, due to screening, are short-ranged (sub- κ^{-1}) in character.

While these findings enhance the understanding of the electrostatics and ion interactions of single-stranded nucleic acids, they also provoke further questions, some of which are discussed below.

First, the origin of the non-logarithmic intermediate-force behavior of ssNAs in divalent salt (Figure 3.2) remains unknown. In Chapter 3, I presented experimental results that showed high-force logarithmic elasticity for ssNAs in monovalent salt and a more compliant, but non-logarithmic, elasticity in divalent salt. I went on, in Chapter 4, to explain the monovalent logarithmic behavior in terms of a modified WLC model, incorporating electrostatic tension, and subsequent transitions to FJC and enthalpic stretching behavior. However, a quantitative explanation of the divalent non-logarithmic behavior does not yet exist. In fact, I am not aware of force spectroscopy studies (akin to those of Chapter 4) that probe this regime as a function of divalent salt concentration. Such experimental studies are certainly technically possible, and could—along with the already measured but unexplained divalent salt-dependence of the crossover force (Figure 3.2)—form the basis for interesting new theoretical investigations. Central to this problem is the realization that basic assumptions underlying our analysis in the monovalent case will not hold for divalents. In particular, not only will the Debye-Hückel approximation used to derive Equation 4.2 be invalid, but so will the full Poisson-Boltzmann equation.

Second, there is room for further development and experimental testing of models of the ion atmosphere. As discussed in Chapter 5, one of the main goals of nucleic acid ion counting experiments is to obtain data that can be used to constrain theoretical models of the ion atmosphere—albeit incompletely, as the theories give more information-rich results than do the experiments. There is still significant work to be done in this direction, both in terms of conducting theoretical investigations that accurately mimic existing experiments (e.g., salt species used) and in terms of performing more experiments under a variety of conditions (e.g., note the paucity of divalent and trivalent data in Table 5.1). Of particular interest would be modeling of those experiments that are most sensitive to the precise effects of ion identity and NA conformation: ion competition and differential ion counting studies. A hint of this sensitivity is seen in the data of Figure 5.4B, in

which Poisson-Boltzmann theory, otherwise able to reproduce all existing ion counting results for monovalents, fails to reproduce the measured $\Delta\Gamma$ data. This may indicate that such data constitute a more stringent test of theoretical predictions than do other experiments, and would make more sophisticated computational modeling of this system, either at the level of fully atomistic molecular dynamics or of a simplified approach, very interesting.

Third, single-molecule differential ion counting experiments have yet to be performed in divalent salt. The RNA hairpin ion counting experiments of Figure 6.4D, and similar results reported previously for DNA by Dittmore *et al.* [96], contributed to our estimation of the single-stranded ion excess in Chapter 6 and are also interesting in their own right in understanding how ions interact with structured NAs. As will be further discussed in Appendix D, such data can, for example, be interpreted to give differential free energies of ion binding. All of the experiments of this type conducted thus far have been in the presence only of monovalent ions, which are believed to interact with NAs nonspecifically by mediating charge screening (this is why such interactions are well-modeled, in most cases, by the mean-field Poisson-Boltzmann equation). Divalent ions, however, are believed to exhibit more specific interactions with both single- and double-stranded nucleic acids [164, 165] and with more complicated folded structures [102, 177]. Differential ion counting experiments involving divalent ions, while rendered more difficult due to coupling between monovalents and divalents in the formalism (see Appendix D), would provide a new experimental modality through which to explore these effects.

Fourth, the apparent force-dependence of the ssRNA persistence length remains unexplained. As mentioned in Chapter 4, our ability to interpret the high-force elasticity of ssRNA—but not of ssDNA—is obscured by this apparent force-dependence. In particular, when the WLC interpolation formula is fit to Θ solution data over 0.7–10 pN, a value of $l_p = 0.85$ nm is obtained, versus $l_p = 0.43$ nm when fitting data over 10–60 pN. We

hypothesize that this additional high-force compliance could be due to the ribose sugar pucker isomerization, transitioning from the C3'-*endo* pucker, with 0.59 nm extension per nucleotide, to the C2'-*endo* pucker, with 0.70 nm extension per nucleotide. It is reasonable to expect that this transition could occur in the vicinity of the force range under study: quantum chemistry calculations estimate the difference in free energy between the two conformations to be 2 kcal/mol [98], which corresponds to a mechanical unfolding force of ≈ 100 pN. Nevertheless, to my knowledge this transition has never been directly measured. To do so, using SMFS at higher forces (100–1000 pN), would both clarify the analysis of our existing data and would provide new details about this important aspect of RNA structure.

Finally, a Netz-like model of polyelectrolyte electrostatic tension could be evaluated using the full Poisson-Boltzmann equation. In Chapter 4, I showed that force-extension data of ssNAs in the 10–100 pN domain are well-fit by a WLC model with added electrostatic tension, using the Netz model of f_{el} with an effective charge spacing that differs from the structural charge separation along the chain. We explained the discrepancy between the effective and structural charge spacing by noting that the Netz model is derived in the Debye-Hückel limit and, consequently, ignores the non-linear screening due to the preferential association of cations in the ion atmosphere (i.e., due to the realization that $\bar{\Gamma} \neq 0.5$ in Figure 6.5). This argument could be made quantitative by evaluating a chain-of-beads model, like the Netz model, using the full Poisson-Boltzmann equation to calculate the potential between the beads. This approach would capture the non-linearity of the ion atmosphere since, as seen in Figure 6.5, these effects are well-described by PB.

In summary, we as a community are building a quantitative understanding of the electrostatic interactions between ions and single-stranded nucleic acids (or, more generally, flexible polyelectrolytes). This understanding is especially well established for monovalent salt solutions, where the mean-field Poisson-Boltzmann approach forms the success-

ful basis of most arguments. From the results we have thus far, however, it is clear that this approach must generally be discarded when considering solutions of higher-valence ions, and it is here that basic work in polyelectrolyte physics, both experimental and theoretical, remains to be done.

Appendix A

Materials and methods

A.1 Sample preparation

A.1.1 Long homopolymeric ssNAs

Polydisperse poly(U), ranging from roughly 4000 to 10,000 nucleotides in length, was synthesized by elongation of 20-mer poly(U) oligonucleotides, 5'-labeled with a protected thiol group (Integrated DNA Technologies), using polynucleotide phosphorylase from *E. coli* (Sigma-Aldrich). The polymerization reaction conditions were those of van den Hout *et al.* [178]. The polymerization product was 3'-labeled with biotin by incorporation of biotin-dUTP using terminal deoxynucleotidyl transferase (Life Technologies). Purification by ethanol precipitation [179] was performed following each reaction. Product concentration was quantified by absorption spectroscopy at 260 nm. Before surface coupling, the thiol group was deprotected by treatment with tris(2-carboxyethyl)phosphine (TCEP).

Polydisperse ssDNA (eighth polypyrimidine sequence of Brockman *et al.* [180]) was synthesized, as in McIntosh *et al.* [68], by rolling-circle amplification [181] with a 5'

biotin-labeled primer for bead attachment; glass substrate attachment was by non-specific adhesion.

A.1.2 Short ssRNA/ssDNA oligonucleotides for D-AES

Samples of 50 nucleotide poly(U) ssRNA and poly(dT) ssDNA were obtained from Integrated DNA Technologies. Oligonucleotides were synthesized by that company at 1 μmol scale, typically resulting in 250–550 nmol of purified product. As above, homopolymeric sequences were chosen to prevent Watson-Crick base pairing and the U and dT monomers, in particular, were chosen because they have been shown in prior studies not to exhibit appreciable base-stacking interactions [69, 170]. Samples were initially diluted in 100–200 μL Tris-EDTA buffer and stored at -80°C . Samples were then further diluted in the desired experimental buffer prior to the start of each dialysis. ssNA concentrations were initially determined by UV/vis spectrophotometry at 260 nm; although, the concentrations used to calculate the ion excess (Equation A.6) were obtained directly from AES.

A.1.3 RNA hairpin molecular construct

The molecular construct used in the single-molecule RNA hairpin unfolding experiments is sketched in Figure 6.4A of the main text and is shown with greater sequence detail in Figure A.1. The construct consists of the RNA helix of interest, closed by a hexaethylene glycol (HEG) spacer to form a hairpin and connected, at the 3' end, to the magnetic bead via a 20 nucleotide thymidine spacer and, at the 5' end, to the surface via an 848 bp dsDNA handle. In the control experiments described in Appendix B, the HEG spacer is replaced by RNA and shown to have no discernible effect on the ion excess results.

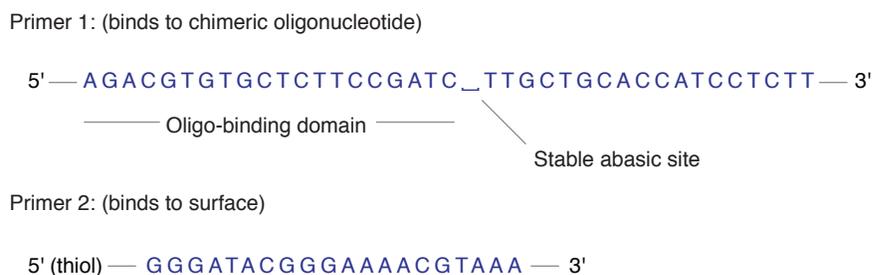


Figure A.2: Sequences of the primers used to produce the double-stranded DNA handle via PCR amplification of a region of the lambda phage genome.

The RNA helix, the single-stranded polythymidine spacer, and a 20 nucleotide domain for binding to the dsDNA handle (i.e., the sequence up to the “ligation point” in Figure A.1) were obtained as a single chimeric DNA/RNA oligomer from Integrated DNA Technologies with a 3' biotin moiety to facilitate binding to streptavidin-functionalized magnetic beads.

The dsDNA handle was prepared by PCR amplification of a section of the DNA genome of bacteriophage lambda. Both PCR primers bore 5' functionality: one had a 5' thiol moiety to facilitate surface coupling via thiol-maleimide chemistry and the other had a 20 nucleotide 5' overhang to facilitate hybridization with the handle binding domain of the chimeric oligonucleotide. The sequences of the primers are given in Figure A.2. Using these primers to perform PCR amplification on lambda DNA gives a dsDNA handle 828 bp in length, in addition to the 20 nucleotide binding region. The handle was ligated to the chimeric oligonucleotide using T4 DNA ligase (New England Biolabs). Samples were purified following enzymatic reactions by phenol-chloroform extraction.

A.2 Magnetic tweezers experiments

The basic setup of an MT experiment is shown in Figure 1.2A. The MT instrument is essentially a conventional microscope with the addition of rare-earth magnets above the stage and a piezoelectric device to precisely control the focus.

The MT experiment begins with the preparation of the sample flow cell (see Section A.2.1), consisting of a fluid channel between two glass coverslips. The bottom surface of the flow cell is chemically functionalized to allow binding of the subject molecule. The molecule is functionalized at both ends, to allow binding to the surface as well as to the magnetic beads, which are flowed into the cell subsequent to surface binding. In all of the experimental work reported in this dissertation (except the ssDNA data of Chapter 4), the surface was functionalized with maleimide, which bound to reduced thiol groups on the NA at pH 7.5; the bead was functionalized with streptavidin, which bound to biotin moieties on the NA.

Once the flow cell is prepared, it is placed on the microscope stage and the experimenter can begin looking for an appropriately tethered molecule to study. Since many tethers form in each flow cell, it is possible to translate the stage to look for a tether that (1) has the appropriate expected length, (2) does not consist of a bead tethered to the surface by more than one molecule, and (3) has its single tether centered roughly at the center of the bead. That there is only one tether can be verified by rotating the magnets above the stage, concomitantly rotating the bead. If there were more than one NA molecule tethering the bead to the surface, this rotation would braid those strands around each other and the observed tether height would decrease. That the tether is centered on the bead can also be verified by such rotation: if the tether were off-center, the rotating bead would be seen to wobble in the video microscopy.

Sometimes, as in the poly(U) experiments of Chapters 3 and 4, the chemical prepa-

ration of the sample is such that the NA length is not uniform. In such cases, longer tethers are usually preferred for two reasons. First, as will be seen in Section A.2.3, the range of forces over which direct calibration is possible extends to higher values for longer tethers, since they will exhibit greater Brownian fluctuations about their mean position. Second, at a particular force a bead attached to a longer tether will tend to be further from the surface, meaning that lower force values can be explored before the bead begins to interact directly with the substrate.

Tracking of the bead position in the imaging plane is easily accomplished using centroid tracking; tracking in the \hat{z} direction, however, requires a calibration between bead height and the diffraction ring pattern that arises from interference between incident light from the LED light source and light scattered off of the bead [182]. Such a calibration is made for a particular bead once a suitable tether has been identified. This is accomplished by moving the magnets close to the sample—thus suppressing most of the Brownian fluctuations—and then using the piezoelectric device—to which the objective lens is mounted—to step the focal plane in 100 nm increments. At each step, the radial distribution profile of the diffraction pattern is fit to a model function [92, 183] and is recorded. During this calibration, it is the bead that is held fixed while the focal plane is moved; in the actual experiment, the bead moves in response to the changing magnetic force, but the focal plane is held fixed. The two motions are related by the index correction factor: the ratio of index of refraction in experimental buffer to that in the objective immersion oil.

With this calibration accomplished, it is now possible to monitor the vertical position of the bead with respect to the reference position at which this calibration was made. However, the required experimental quantity is the height with respect to the bottom of the flow cell (i.e., the extension of the tether, X in Figure 1.2A). To accomplish this translation, a zero-force reference calibration is made by tracking the bead position in

the absence of applied magnetic force. Since the magnetic bead is slightly more dense than water, it will sediment to the bottom of the flow cell in the absence of applied force. By tracking the sedimented bead, it is possible to identify the flow cell surface ($X = 0$) in terms of the diffraction pattern calibration.

Having made both of these calibrations, the force spectroscopy experiment itself can be begun. To make a force-extension curve (e.g., Figure 1.2B), the magnets are held at a particular distance from the sample and the bead height above the surface—and thus the molecular extension, X —is measured as a function of time. The time average of X is computed and used as the extension at that magnet position, and the force is determined by direct calibration based on monitoring the horizontal bead fluctuations (see Section A.2.3). The effects of instrumental drift are minimized by subtracting from the measured extension signal the position of a reference bead adhered directly to the flow cell surface. The magnets are then moved to their next position and the process repeated. For a standard force-extension curve (e.g., Figure 3.1) spanning 0.1–10 pN, 31 separate magnet positions are typically used, separated by 0.25 mm. Because thermal fluctuations become more pronounced as the force becomes smaller, adding noise to the measurement of the average extension and shifting the characteristic frequency of the force calibration to lower values, data at each diminishing force are integrated for progressively longer lengths of time (1.175 times longer for each force in a standard experiment; this factor is referred to as the “scaling of acquisition”). All of these parameter values must be modified as needed for a particular experiment. For example, in a folding-unfolding experiment (e.g., the RNA hairpin studies of Chapter 6) only a very narrow range of forces is probed (~ 12 pN); in higher-force experiments with larger beads (e.g., Chapter 4), thermal fluctuations are less dominant and a smaller scaling of acquisition can be used.

Once a single force-extension curve has been collected, additional curves are typi-

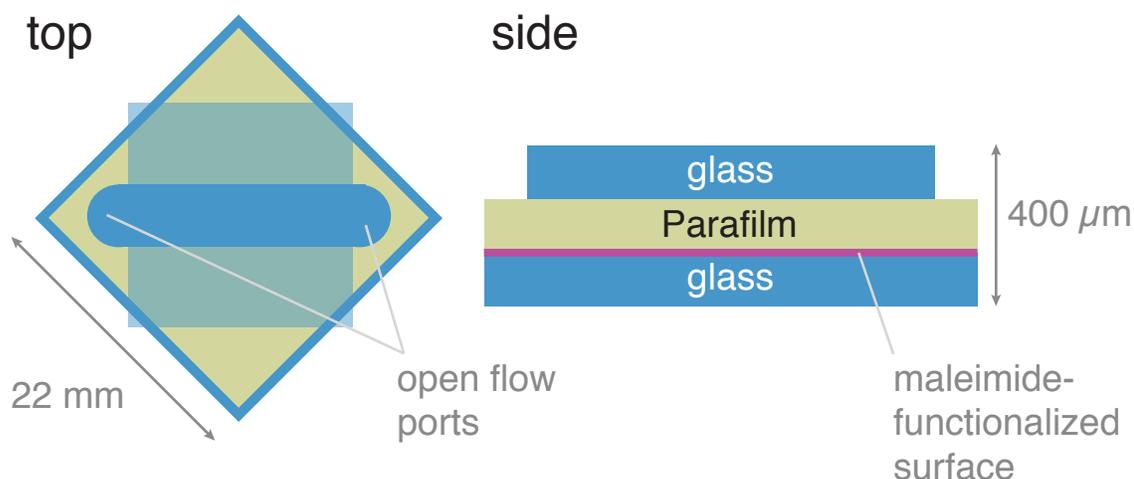


Figure A.3: Diagram of flow cell used in MT experiments, consisting of a laser-cut Parafilm paraffin wax film sandwiched between two no. 1 cover slips; the bottom coverslip is functionalized with PEG and sparsely labeled with maleimide groups. Gaps are left at the end of the channel to allow solution exchange.

cally taken as a function of solution conditions; e.g., salt concentration. The flow cell, which is open at each end, facilitates solution exchange by pipetting into one end while withdrawing solution, typically by capillary action, at the other. This flow creates a hydrodynamic force on the bead that can sometimes shear it from the surface or, under very low or very high ionic strength conditions when the bead is prone to sticking to the glass surface, cause it to adhere to the substrate. The latter issue can sometimes be avoided by flowing the new solution into the flow cell with the magnets very close to the sample, which will act to prevent the bead from contacting the surface. Both issues can also be avoided using more sophisticated permeable-membrane flow cells that separate the solution flow from the experimental chamber; although, such cells introduce other experimental complications [184].

A.2.1 Flow cell preparation

The maleimide-functionalized flow cell used in all MT experiments discussed in this dissertation was assembled from the following supplies: 22 mm square no. 1 cover glass with PEG brush coating and sparse maleimide functionalization (Microsurfaces, Inc.) (bottom glass), 18 mm square no. 1 cover glass (top glass), Parafilm Type M paraffin wax film cut into 22 mm square pattern with central channel (Figure A.3) using laser cutter.

The flow cell is assembled as follows:

1. Top glass (18 mm square no. 1 cover glass) is cleaned by dipping in 1 M NaOH followed by water and then plasma cleaned in a Harrick Plasma PDC-32G for 5 minutes on the high setting.
2. One surface of the top glass—i.e., the surface that will ultimately constitute the top surface of the flow cell—is treated with a hydrophobic coating (Rain-X, ITW Global Brands) to prevent solution from adhering to it during buffer exchange.
3. One maleimide-functionalized slide is removed from storage in a vacuum desiccator. Reference beads (i.e., beads specifically adhered to the surface) are added by treating the slide with 10 μL of isopropyl alcohol solution containing 1 μL of polystyrene bead suspension (Bangs Laboratories) per mL. Once this solution has dried on the slide, it is heated on a hot plate at 140°C for 150 s to melt the beads to the surface.
4. The flow cell is assembled by placing, atop the functionalized slide, the laser-cut parafilm and then the top glass, as shown in Figure A.3. The assembly is bonded by melting the parafilm through the glass using a soldering iron.

At this point, the flow cell is ready for introduction of the sample. Typically, the thiol-labeled NA sample has been incubated for 5 minutes in a 50 mM solution of TCEP, which

reduces the thiol groups to prepare them for surface binding. The sample is then diluted to 50 μL in a coupling buffer consisting of 50 mM sodium phosphate buffer (pH 7.2), 50 mM NaCl, and 10 mM EDTA. The sample is then incubated for at least 20 minutes to allow binding of the NA to the surface. After this time, the channel is flushed with Tris buffer, the amine groups of which will also bind to the surface maleimide to deactivate any remaining sites. Then, the magnetic beads (Dynabeads MyOne or Dynabeads M280, ThermoFisher Scientific) are flowed in at 10% stock concentration in a buffer containing 250 μM Tween-20 surfactant (this surfactant is preserved in all subsequent buffers to prevent beads from adhering to the surface). After the beads have incubated in the flow cell for 20 minutes, the excess is flushed out and the experiment is ready to begin.

A.2.2 Instrumentation

All MT experiments discussed in this dissertation were carried out on an instrument first described by Ribeck and Saleh [92] and further developed subsequently. Photographs of the instrument with major components labeled are given in Figure A.4; compare with the schematic diagram of Figure 1.2A.

The light source is a red resonant cavity LED with peak emission at 650 nm (RC-LED-650-02, Roithner Lasertechnik). The motors that translate and rotate the magnets are M-126 units from Physik Instrumente (PI), controlled by a PI C-843 unit. The stage is also motorized to allow precise movement of the sample: a PI M-545.2P microscopy stage controlled by a PI C-867.260 unit, which is operated manually using the joystick input. The Nikon planar fluorescence 60x oil immersion objective rests atop a piezoelectric device, which is controlled by a PI E-615.CR unit. The objective is operated at its specified 60x magnification by using a 200 mm focal-length tube lens placed 200 mm from the camera sensor. The camera is a 782x582 pixel progressive scan monochrome

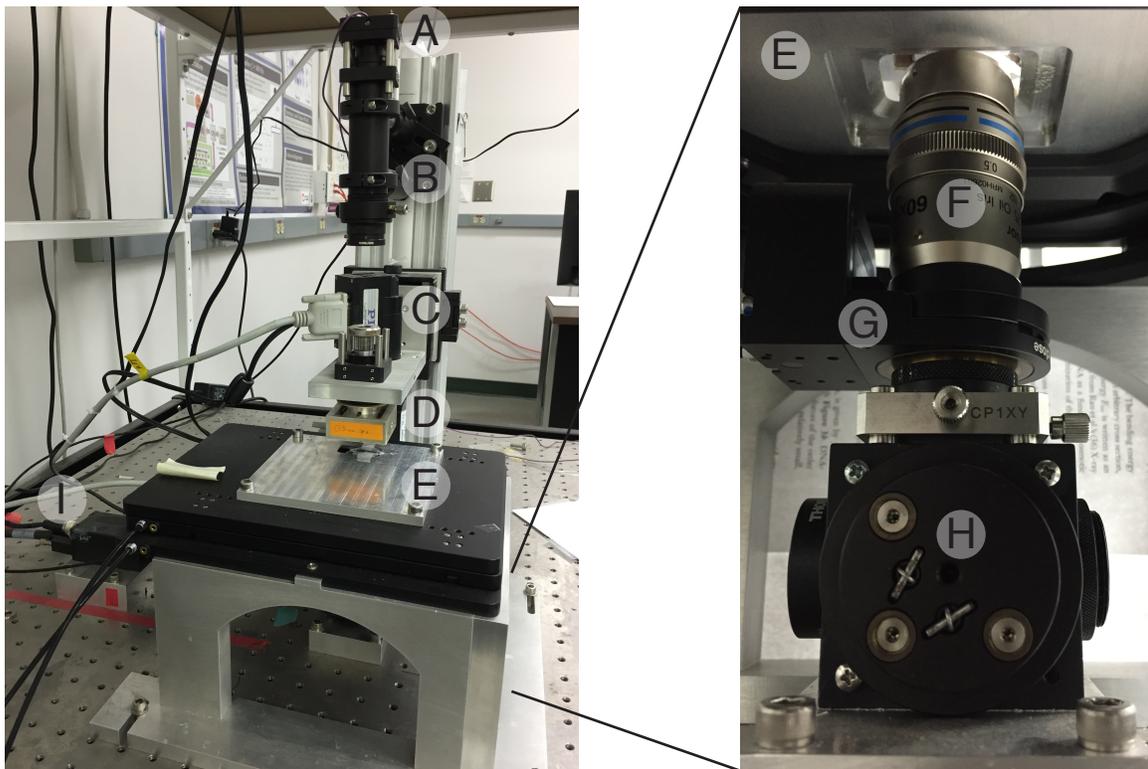


Figure A.4: Photographs of the MT instrument used in these studies, with major components labeled. (A) Light source, (B) focusing optics, (C) magnet translation/rotation motors, (D) magnet assembly, (E) stage, (F) objective lens, (G) piezo-electric device, (H) 45° slanted mirror and tube lens, and (I) camera.

CCD camera operating at a 60 Hz frame rate (CV-A10 CL, Jai Electronics) and connected to a National Instruments PCIe-1429 capture card utilizing the CameraLink standard.

Aside from the stage, all elements of the instrument are controlled by a PC using software written in the LabVIEW development environment. There is a separate LabVIEW application for each stage of the experimental workflow: a “basic microscope” program for monitoring the flow cell and identifying target molecules, a program for making the calibration between bead diffraction pattern and bead height, a tracking program to monitor the extension as a function of magnet position, a force-extension program to automatically move the magnets at timed intervals to collect force vs. extension data, and a basic analysis program to perform the direct force calibration. However, more sophisticated force calibration and other analysis routines were written in MATLAB.

A.2.3 Force calibration

In a constant-force force spectroscopy experiment, such as MT, molecular extension is measured as a function of applied force. As discussed above, the extension is measured by observing the diffraction pattern cast by the magnetic beads and then relating this pattern to a bead-height calibration. The applied force is controlled by the height of the magnets above the stage. The calibration between magnet position, z_{mag} , and force, using several different approaches, is discussed below.

Direct (thermal) calibration

The most direct force calibration approach relies on measuring the Brownian fluctuations of the magnetic bead. Since the bead is tethered to the end of an NA molecule, it can be thought of as an inverted pendulum where the magnetic force is analogous to gravity. Because the system is in equilibrium with a thermal bath, the average energy

in each of the pendulum’s two degrees of freedom is given by the equipartition theorem. From this argument, one can derive an expression for the calibrated force: [185]

$$f = \frac{\langle X \rangle k_B T}{\langle \delta_x^2 \rangle}, \quad (\text{A.1})$$

where δ_x^2 is the variance of the bead position in the direction orthogonal to the pulling axis and the braces denote time-average values. Since this calibration method relies on a single value, the variance, computed from the data, it is susceptible to bias arising from anomalous instrumental noise—leading to calibrated force values that are too low—and under-sampled fluctuations due to the frame rate of the camera—leading to force values that are too large.

More robust force calibration approaches involve calculating spectra that more completely utilize the data: either the power spectral density (PSD) [186–190] or the Allan variance [186]. In the work described in this dissertation, an Allan variance-based approach was used, following Lansdorp and Saleh [186]. In particular, for each bead trace at a certain force, the AV was computed and then fit, using a maximum likelihood approach, to the model function

$$\sigma^2(\tau) = \frac{2k_B T \alpha}{k^2 \tau} \left(1 + \frac{2\alpha}{k\tau} e^{-\frac{k\tau}{\alpha}} - \frac{\alpha}{2k\tau} e^{-\frac{2k\tau}{\alpha}} - \frac{3\alpha}{2k\tau} \right), \quad (\text{A.2})$$

where σ^2 is the Allan variance, τ is the separation time between samples, and k and α are the fit parameters: the pendulum spring constant and the dissipation due to viscous drag on the bead, respectively.

Example AV spectra for an ssRNA molecule at two different magnitudes of applied force are shown in Figure A.5. Figure A.5A shows the generally expected behavior: a region at low- τ corresponding to underdamped motion, giving way to a broad, decreasing-

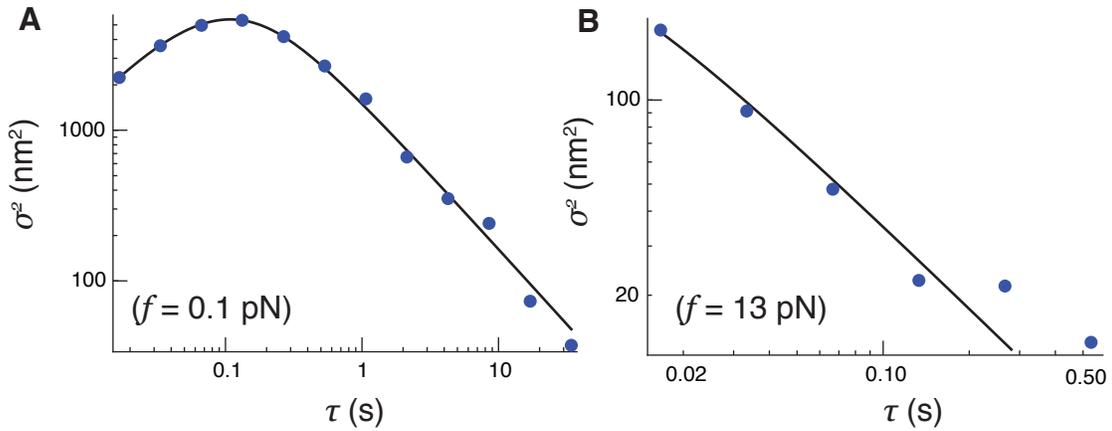


Figure A.5: Sample AV of ssRNA at two different applied forces, with best fits of Equation A.2. (A) Lower force, where the characteristic turnover in the AV is well sampled and both principle regimes are visible (there is no significant contribution from large- τ drift; such drift can be seen in Figure E.4A). (B) Higher force, where the turnover occurs at too low of τ to be fully sampled, leading to poor constraint of the model fit parameters. Constraint of α based on other data can minimize this effect.

AV regime in which thermal noise is being averaged away by longer integration times, and then, finally, a high- τ regime dominated by instrumental drift. Because it is dominated by drag, the low- τ region predominately determines the fit of α , whereas the fit of k is predominately set by the intermediate- τ region. Thus, a good fit of both parameters requires that the turnover between the two regimes be well-sampled. This is the case in Figure A.5A, but not in A.5B, where this turnover occurs at lower τ (higher frequency) than is sampled by the camera frame rate, leading to a low-quality fit.

If the maximum AV turnover value is never sampled throughout a force extension curve, direct force calibration from the bead fluctuations is not possible and longer tethers or a faster camera frame rate must be used, or an indirect calibration curve approach must be employed (see below). However, if the transition is adequately sampled over most of the force range, and it is only at the highest-force values that it shifts to too low of τ , fits can be extended into the under-sampled region by fixing the value of α . Values of α from the two-parameter fit of Equation A.2 to ssRNA data are shown in Figure

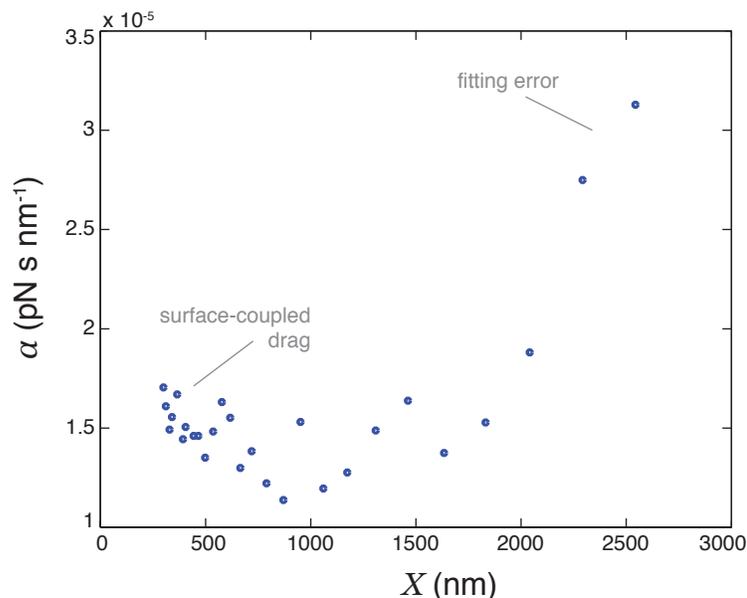


Figure A.6: Values of α from fitting Equation A.2 to ssRNA force spectroscopy data, compared with height of the bead above the substrate at that force. The increase in α at low X is due to hydrodynamic coupling with the surface, whereas the increase at high X is an anomaly of the poor fit quality. The fit quality improved at high z_{mag} by fixing the value of α in Equation A.2 to that in the bulk.

A.6. Over most of the range of X explored, the value of α is constant, as it is expected to be in bulk solution. At low force, when the bead is near to the surface, α increases due to hydrodynamic coupling [191]; the apparent increase in α at high force is due to diminution of the fit quality. Since the high-force points remain far from the surface, the bulk value of α is expected to apply there as well. Therefore, by extending the values of α from intermediate z_{mag} , and repeating the fitting of Equation A.2, now with α fixed, it is possible to accurately establish k , even from cut-off data as in Figure A.5B.

Indirect calibration curve

For molecules that are sufficiently short or are measured at sufficiently high forces, a direct force calibration cannot be made, even by fixing α as described above. The lengths and forces that cannot be directly calibrated are a function of the camera frame rate: the

faster the frame rate, the stiffer the tether can be and still be calibrated. One can tell when the force calibration will break down either by estimating the Lorentzian corner frequency ($f_{\text{corner}} = k/2\pi\alpha$) [186] or by performing Brownian dynamics simulations like those discussed in Appendix E. In the context of this dissertation, such a breakdown in the ability to perform direct force calibration occurred in the case of the hairpin folding/unfolding studies discussed in Chapter 6. In that case, the tethers were only about 300 nm long, corresponding, at 12 pN, to $f_{\text{corner}} \approx 400$ Hz, much faster than the 30 Hz Nyquist frequency of the camera.

Because of this fast characteristic frequency, an indirect force calibration procedure was instead used in these studies. In a separate flow cell, long ($\sim 2\text{--}5 \mu\text{m}$) polyethylene glycol (PEG) molecules were prepared for force spectroscopy by attachment to the substrate via the thiol-maleimide linkage and to magnetic beads via the biotin-streptavidin linkage. Force spectra were then recorded over the force range of interest for the hairpin experiments, roughly 10–20 pN. This process was repeated for approximately ten molecules. During the measurement of each molecule, the magnet position corresponding to the tips of the magnets just touching the top of the flow cell, z_{max} , was recorded. For each molecule, we could then prepare a force calibration curve as a function of magnet height above the top of the flow cell: $h_{\text{mag}} = z_{\text{max}} - z_{\text{mag}}$. This subtraction was done to account for the possibility of the magnet assembly shifting slowly over time with respect to the stage, an effect that was in fact observed to occur on a time scale of months.

Because there is a roughly 10% variation in the magnetic moment between beads, the calibration curves from the various PEG molecules were not in agreement. This is the principal disadvantage of using an indirect force calibration approach. To make our best estimate of the true force calibration, all of the PEG curves were averaged to form a single, master curve, which was used in the analysis of the hairpin data, introducing approximately 10% error into the forces in the process.

A.2.4 Elasticity measurements

The elasticity experiments of Chapters 3 and 4 were performed in a flow-cell assembled as discussed above. In the lower-force studies of Chapter 3, the Parafilm in Figure A.3 was replaced by double-sided adhesive tape (3M Company). Reference beads were attached to the surface, in the work of Chapter 3, by coupling amine-labeled latex spheres to the maleimide-functionalized surface and, in the work of Chapter 4, by melting latex beads directly to the surface as explained in Section A.2.1. Subsequent experimental steps were performed in 10 mM Tris-HCl (pH 7.5) buffer with 250 mM Tween 20 surfactant added to prevent adhesion of magnetic beads to the flow-cell surface [192]. For $I > 7$ mM, solutions were prepared by adding NaCl to a background of 10 mM Tris buffer at pH 7.5 (10 mM of this buffer corresponds to a 7 mM ionic strength). For $I < 7$ mM, the ionic strength was set by reducing the buffer concentration.

Due to alignment with the applied magnetic field, the beads have only one rotational degree of freedom. Rotational fluctuations can lead to underestimation of the true end-to-end extension of the RNA [69]. We correct for this effect by employing an equipartition argument to find the expected underestimate, δ , given by:

$$\delta = \frac{k_B T}{2f}. \quad (\text{A.3})$$

We then add δ to the measured bead height to give the true polymer extension, X . This expression for δ is valid in the $fR \gg k_B T$ limit, where R is the bead radius. We note that this amounts to a fairly minor correction, affecting the lowest-force data by 5% and the high-force data by much less. As such, the correction is applied only to the low-force data of Chapter 3.

A.2.5 Equilibrium folding/unfolding studies

In an equilibrium folding/unfolding experiment, like the hairpin study of Chapter 6, a constant force is applied to the folded molecule. If the force applied is near $f_{1/2}$, the molecule will exhibit telegraphic conformational switching, as in Figure 6.4B. As the force is raised or lowered slightly, the equilibrium between the two states will shift. For each force probed, the equilibrium constant, K , can be computed from the ratio of times spent in each of the two states. A plot of K vs. f can then be interpolated to find the exact value of $f_{1/2}$, corresponding to $K = 1$, which is used in the ion counting analysis.

From a practical standpoint, these studies were carried out, for a particular molecule and at a particular salt concentration, by first moving the magnets under manual control to find the approximate value of $f_{1/2}$. Then, a force-extension curve was collected over a very narrow range of forces bracketing this value (the width of the two-state transition, governed by Boltzmann statistics, is about 1 pN). Typically, data were collected over a magnet position range of 20 μm , with a step size of 2 μm .

A.2.6 Data analysis

Rescaling

The long single-stranded nucleic acids used in the elasticity measurements of Chapters 3 and 4 were prepared—either by polynucleotide phosphorylase synthesis in the case of RNA or by rolling circle amplification in the case of DNA—in such a way that the overall chain length varies greatly from molecule to molecule within a sample. A strength of single-molecule force spectroscopy experiments is that, since they allow individual molecules to be addressed one at a time, molecules with the desired length (i.e., long enough for direct force calibration to be effective) can be selected and studied. Nonetheless, the specific contour lengths of the molecules will still vary. To allow analysis of

the scaling behavior of molecules of differing lengths, the data are sometimes rescaled by their extension at a particular force, as in the ssRNA data of Figure 3.1, where the force-extension curves for the different concentrations of a particular salt species were, for technical reasons, not all collected on the same molecule. In the case of Figure 3.1, the data were all rescaled by the extension at 10 pN applied force, a force value at the high end of the range studied and above the Pincus-to-logarithmic crossover force of greatest interest in that study. However, the data of Chapter 4 clearly indicate that the salt-dependence of ssRNA elasticity does extend beyond 10 pN, and so the agreement of the curves at 10 pN in Figure 3.1 should be viewed as an artifact of the rescaling and not as an actual physical concurrence. For this reason, when doing this sort of rescaling, it is important to ensure that any subsequent analysis performed on the rescaled data is insensitive to the rescaling having been performed; i.e., only drawing conclusions about scaling behavior, and not about the quantitative separation between the curves.

Determination of crossover force

The low-force ssRNA elasticity measurements of Chapter 3 showed two scaling regimes, separated by a salt-dependent crossover force, f_c . These crossover forces were quantified by least-squares fitting of each force-extension curve to a continuous, piecewise function with different functional forms above and below f_c , as in Saleh *et al.* [43]. In monovalent salt, a power-law was used for $f < f_c$ and a logarithm for $f > f_c$. In divalent salt, a power-law was used in both regimes. Application of this method led force-extension curves in the good-solvent regime to collapse to a single master curve for each valence; however, a constant multiplicative factor had to be introduced to achieve good low-force agreement between monovalent and divalent salt. Since we are interested only in the scaling of f_c with ionic strength, and not its absolute value, this offset is unimportant.

ssRNA $\Delta\Gamma$ measurement

In Section 3.1.3, a previously described method [75] was used to extract the ion excess (number of ions associated with the RNA compared with an equivalent volume of bulk solution) from the applied force, f , extension, X , and bulk salt concentration, c_{bulk} . Force-extension curves of a particular poly(U) molecule were collected for $f \approx 0.1$ to 10 pN and $c_{\text{bulk}} = 20, 50, 100, 200, 500, 1000, 2000$ mM. For this analysis, the forces measured at a particular magnet position were averaged across all salt concentrations (weighted by confidence). A force-extension curve was also collected at the Θ condition and fit to a WLC model to give the overall contour length. It was previously established by Landy *et al.* [75] that $X(f, c_{\text{bulk}})$ data of this form are well fit by a second-order polynomial in $\log(c_{\text{bulk}})$ and a third-order polynomial in $\log(f)$; such a surface was fit to these data. Equation 3.2 was then applied to this surface to give the change in ion excess, $\Delta\Gamma$, as a function of force. The excess per base was found by dividing $\Delta\Gamma$ by the number of bases: L_c from the WLC fit at Θ divided by 0.59 nm per base [193]. Due to potential edge effects in the fitting, ion excess curves for the highest and lowest c_{bulk} are ignored. Error in the reported values of $\Delta\Gamma$ arises from two sources: uncertainty in the measured values of f , which is reflected in the error bars we report, and uncertainty in the fitted value of L_c at Θ , which introduces a 10% error common to all of the ion excess curves.

Hairpin $\Delta\Gamma$ measurement

The change in ion excess, $\Delta\Gamma_{\text{ds-stretch}}$, as an RNA hairpin is mechanically unfolded (Chapter 6) was determined by measuring the salt-dependent unfolding force using magnetic tweezers [92] and then applying a thermodynamic argument [30, 75, 76, 96, 157]. The molecular construct (Figure 6.4A, Figure A.1) is described in Section A.1.3. Experiments

were carried out in 10 mM MOPS buffer, adjusted to pH 7.5, and in various concentrations of NaCl determined by weight. Control experiments were performed to show that the $\Delta\Gamma$ results are insensitive to the nature of the hairpin-closing loop and to the cationic versus anionic nature of the buffer [194] (except at very low salt, $c < 25$ mM). The results of these control experiments are given in Appendix B.

The thermodynamic argument used to extract $\Delta\Gamma_{\text{ds-stretch}}$ is summarized below and fully explained elsewhere in Appendix D. We begin with a thermodynamic identity connecting changes in the relevant free energy for magnetic tweezers experiments, the grand canonical (Landau) potential of mean extension, $d\Omega$, and changes in the experimental parameters X , f , Γ_i , and μ_i :

$$d\Omega = -Xd f - \sum_{\text{species}} \Gamma_i d\mu_i. \quad (\text{A.4})$$

In an experiment, f and μ are directly controlled and X is measured; therefore, Equation A.4 can be used to derive several relations to extract $\Delta\Gamma$, the change in ion excess between conformational states, from the experimental data [30]. In Chapter 6, we rely upon one of these methods, a generalized Clausius-Clapeyron relation, although we show in Figure B.3 that the results of the other methods are in agreement. This Clausius-Clapeyron relation is given by

$$\Delta\Gamma = \frac{\Delta X}{2k_B T A} \left(\frac{\partial f_{1/2}}{\partial \ln c_{\text{bulk}}} \right). \quad (\text{A.5})$$

The salt-dependent values of $f_{1/2}$ were extracted from two-state folding/unfolding trajectories as shown in Figure 6.4B. For each trace, collected at a particular c_{bulk} and constant f , a histogram was made and the equilibrium constant calculated as the ratio of the occupancy time in the folded state to that in the unfolded state. For each salt concentration, $\ln K$ was plotted versus f and fit with a two-state Boltzmann model to

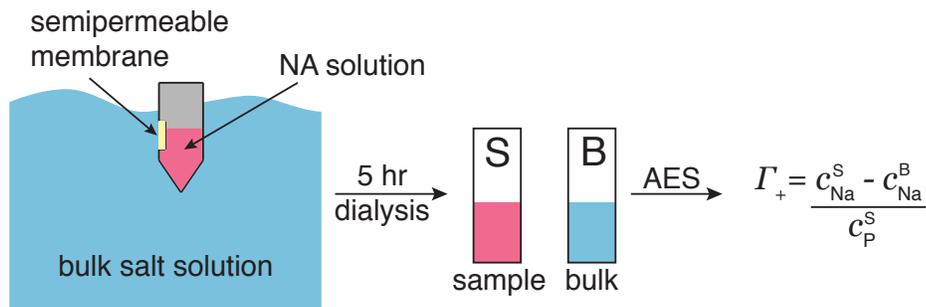


Figure A.7: Overview of D-AES experiment. A solution containing the NA sample is dialyzed, across a semi-permeable membrane, against a bulk solution with a particular salt concentration. Once the system has had time to equilibrate, aliquots are taken from both the NA solution and the bulk and the concentrations of Na and P (as a proxy for NA concentration) measured by AES. The ion excess is obtained from these concentrations.

obtain $f_{1/2}$. The derivative $(\partial f_{1/2} / \partial \ln c_{\text{bulk}})$ in Equation A.5 was evaluated separately for the $f_{1/2}$ data of each trial, yielding $\Delta\Gamma_{\text{ds-stretch}}$.

A.3 Dialysis/atomic emission spectroscopy

A.3.1 Procedure

We measured the ion excess of 50-mer homopolymeric oligonucleotides (Integrated DNA Technologies) of thymidine (DNA) and uridine (RNA); the samples are further described in Section A.1.2. The general scheme of the D-AES experiment is shown in Figure A.7. NA solutions (100 μL , 0.07-0.7 mM) were dialyzed against a 200 mL bulk reservoir of buffer containing 1 mM MOPS, adjusted to pH 7 with NaOH, and the desired concentration of NaCl. A dialysis membrane with a 6-8 kDa molecular weight cutoff was used (D-Tube Dialyzer Mini, EMD Millipore). After 5 hours of stirring, samples were taken of both the NA and bulk solutions and diluted with pure water to concentrations in the range of sensitivity of the AES spectrometer.

The Na^+ ion excess per nucleotide, $\bar{\Gamma}_{\text{ss}}$, was determined from the molar concentrations

of sodium in the nucleic acid solution, $c_{\text{Na}}^{\text{sample}}$, and in the bulk, $c_{\text{Na}}^{\text{bulk}}$, and from the concentration of nucleotides, $c_{\text{P}}^{\text{sample}}$: [35]

$$\bar{\Gamma}_{\text{ss}}^{\text{molar}} = \frac{c_{\text{Na}}^{\text{sample}} - c_{\text{Na}}^{\text{bulk}}}{c_{\text{P}}^{\text{sample}}}. \quad (\text{A.6})$$

The various concentrations were measured using inductively coupled plasma atomic emission spectroscopy (AES) (iCAP 6300, Thermo Scientific) calibrated using standard solutions of sodium and phosphorus (Fluka). The details of the calibration are given below. To keep the measurement uncertainty in the Γ_{ss} results reasonably small, it is necessary to increase the sample NA concentration as the bulk salt concentration is increased. Comparison of data obtained at different NA concentrations shows no concentration dependence; nonetheless, we were careful always to stay well below the solution overlap concentration: approximately 3.5 mM at low $c_{\text{Na}}^{\text{bulk}}$, based on salt-dependent persistence lengths from the experiments of Chen *et al.* [80]. Because we noticed a correlation between samples from separate dialysis runs read out in the same AES batch, we use the number of such batches, and not the total number of samples, in computing the reported standard errors of the mean. Each reported D-AES datum represents the results of at least three dialysis experiments, each read out by AES at least twice.

A.3.2 AES calibration

All samples collected from the equilibrium dialysis experiments were significantly diluted to fall within the working range of the spectrometer; that is, less than about 50 PPM (1 PPM = 1 μg analyte per mL solution, a molar unit). The mass of water used in the dilutions was measured by weight.

We measured the concentration of sodium by monitoring emission in three channels: 588.9 nm, 589.5 nm and 818.3 nm. We measured the concentration of phosphorus, as a

proxy for the concentration of NA monomers, by monitoring emission in five channels: 177.4 nm, 185.8 nm, 185.9 nm, 213.6 nm and 214.9 nm. We then related signal intensities in these channels to concentrations of our diluted samples via calibration curves. We prepared standard solutions for these curves by diluting Fluka stock standard solutions (Sigma-Aldrich). Each of the 7 AES experiments carried out was independently calibrated; sample calibration curves from one representative experiment are shown in Figure A.8. Note that the phosphorus calibration curves are generally linear across the range of concentrations probed; however, the sodium curves show significant nonlinear behavior at low concentration. It is thus essential to use more than two points in the calibration to capture the shape of the curve.

A.4 Poisson-Boltzmann calculations

We used Poisson-Boltzmann calculations to estimate the ion excess of the helix state, Γ_{ds} , using model A-form or B-form helices for RNA or DNA, respectively. These same helices with one strand removed were used as naive models of the single-stranded nucleic acids. Calculations were carried out using the Adaptive Poisson-Boltzmann Solver [116] over a range of NaCl concentrations. Structural models were obtained from the Nucleic Acid Builder [163] and charges were assigned using the PDB2PQR package [195, 196]. The detailed parameters used are given below. Although we use atomically resolved structural models here, Shkel and Record have demonstrated that comparable results can be obtained for NA thermodynamic properties using cylindrical models [174]. Ion excess values were obtained in both molal [35] and molar [37, 109] ensembles by appropriate integration of the resulting ion densities.

The following parameters were specified in the APBS input file: (the meaning of these parameters is described in the APBS documentation)

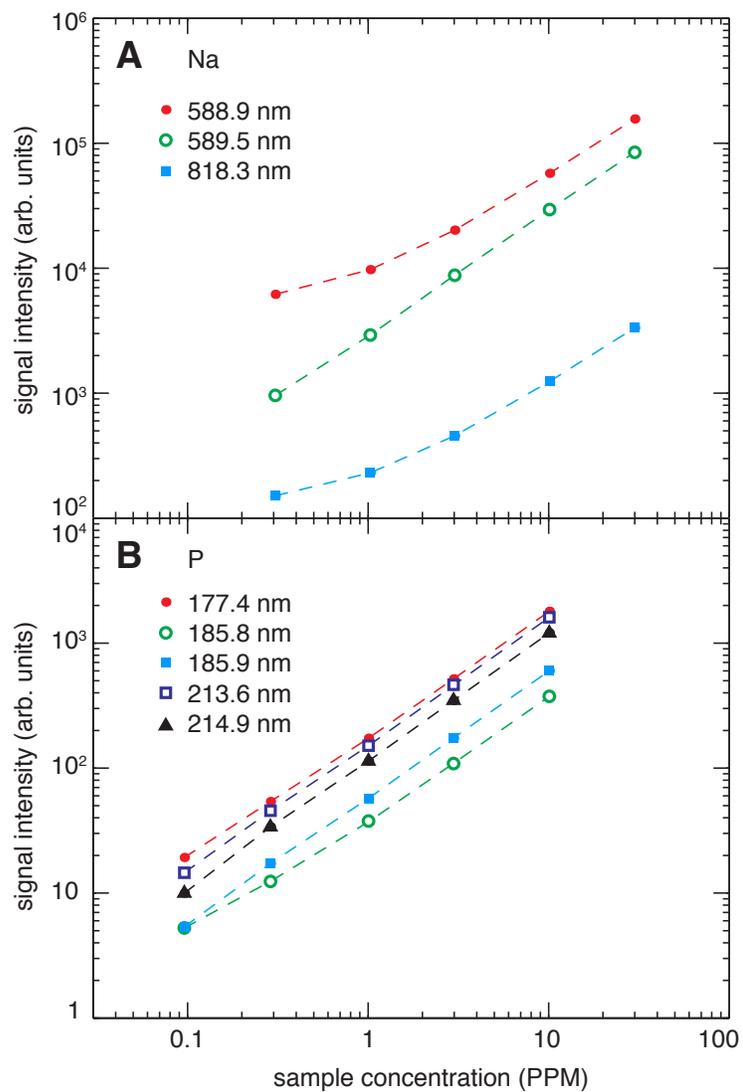


Figure A.8: AES calibration curves for (A) sodium and (B) phosphorus. Dashed lines represent the cubic interpolation function used to relate experimental intensities to concentrations.

```
mg-auto
dime 129 129 225
cglen 192 192 336
fglen 192 192 336
cgcent mol 1
fgcent mol 1
mol 1
npbe
bcfl gdh
pdie 2.000
sdie 78.5400
srfm gmol
chgm spl2
sdens 10.00
srad 1.40
swin 0.30
temp 293
calcenergy no
calcforce no
ion charge 1 conc [bulk salt concentration in M] radius 2.27
ion charge -1 conc [bulk salt concentration in M] radius 1.75
write ndens dx ndens
```

Running these instructions generates a file specifying, at each grid point in the three-dimensional volume, the total number-density of all ion species. To obtain the ion excess, it is then necessary to integrate the ion density to give the total number of ions in the volume and then to subtract either the total number expected based on the bulk salt concentration, to obtain $\Gamma_{\text{ds}}^{\text{molar}}$, or the number expected outside of the NA excluded volume, to obtain $\Gamma_{\text{ds}}^{\text{molal}}$. This corresponds to the two different limits of the integral in Equation 5.3. Finally, the cation excess and anion depletion are backed out based on the total charge of the macromolecule via the condition of charge neutrality.

A.5 Analysis of duplex thermal denaturation data

Salt-dependent DNA duplex melting temperatures are reported by Owczarzy *et al.* [154] for a library of sequences with various lengths and G-C base-pair content. Since we were interested in making a connection to the 25 bp sequence of the present study, we used Equation 5.8 to analyze only the 25 bp melting data. Increasing the G-C content of the sequences tends to increase the melting temperature, but we see no trend in $dT_m/d\ln c_{\text{bulk}}$. Therefore, we used all reported 25 bp sequences regardless of base pair composition.

For each duplex, with reported T_m values at $[\text{NaCl}] = \{69, 119, 200, 621, 1020\}$ mM, we fit a smoothing spline to the data to allow evaluation of $dT_m/d\ln c_{\text{bulk}}$. The $\Delta\Gamma_{\text{ds-ss}}$ values obtained for the various sequences were then combined into a single curve, from 69-1020 mM, by averaging. The uncertainty was taken as the sum, in quadrature, of the contribution from the sequence-to-sequence variation (standard error of the mean) and the contribution from the approximately 10% uncertainty in the value of β . As discussed in Chapter 6, $dT_m/d\ln c_{\text{bulk}}$ was assumed to be salt-independent below 69 mM and the data were extrapolated into this region under that assumption; therefore, the error below 69 mM is from the uncertainty in β alone. The $\Delta\Gamma_{\text{ds-ss}}$ values obtained from this analysis were shown in Figure 5.4B alongside the results from our single-molecule force denaturation studies (i.e., $\Delta\Gamma_{\text{ds-stretch}} + \Delta\Gamma_{\text{stretch-ss}}$).

Appendix B

Hairpin $\Delta\Gamma$ control experiments

B.1 Hairpin control experiments

Whereas we wanted to ascribe the $\Delta\Gamma_{\text{ds-stretch}}$ behavior obtained from our hairpin unfolding experiments (Figure 6.4D) to effects intrinsic to folded and unfolded NAs, and not to effects arising from the experimental setup, we performed several control experiments.

In the first control experiment, we compared the salt-dependent unfolding forces and ion excess changes of a 25 bp RNA hairpin closed by two different loops: the hexaethylene glycol spacer reported in the main text (Figure 6.4A) and an RNA loop consisting of five uridine bases. This control experiment tests if there is a significant contribution from the loop to the ion excess behavior and is especially important because, in our analysis, we compare data from the PEG-closed RNA hairpin with previously reported [96] data on a DNA hairpin with a polythymidine loop. The results of this first control experiment are shown in Figure B.1, in which we see no significant difference in behavior between the two constructs.

In the second control experiment, we compared the same quantities with respect to

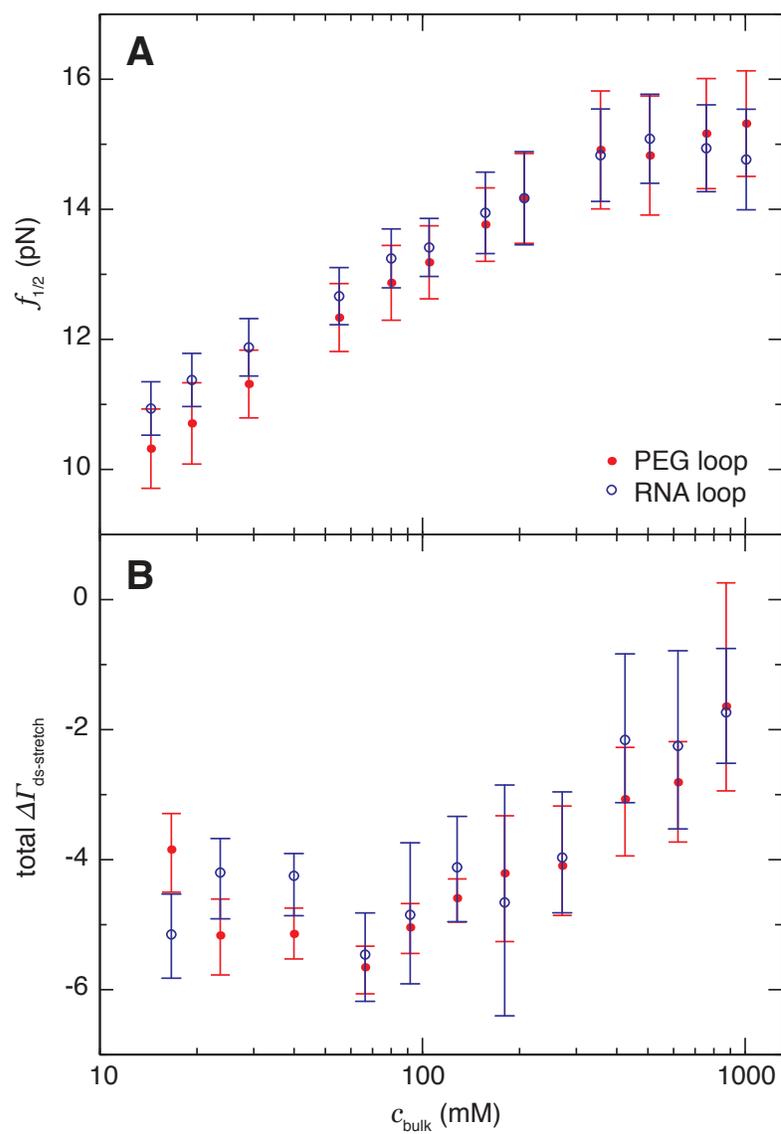


Figure B.1: Comparison of (A) unfolding force and (B) ion excess change upon unfolding determined by the Clausius-Clapeyron method for the hairpin construct shown in Figure A.1 (closed red circles) and an identical construct in which the two strands of the hairpin are connected by five uridine bases (open blue circles).

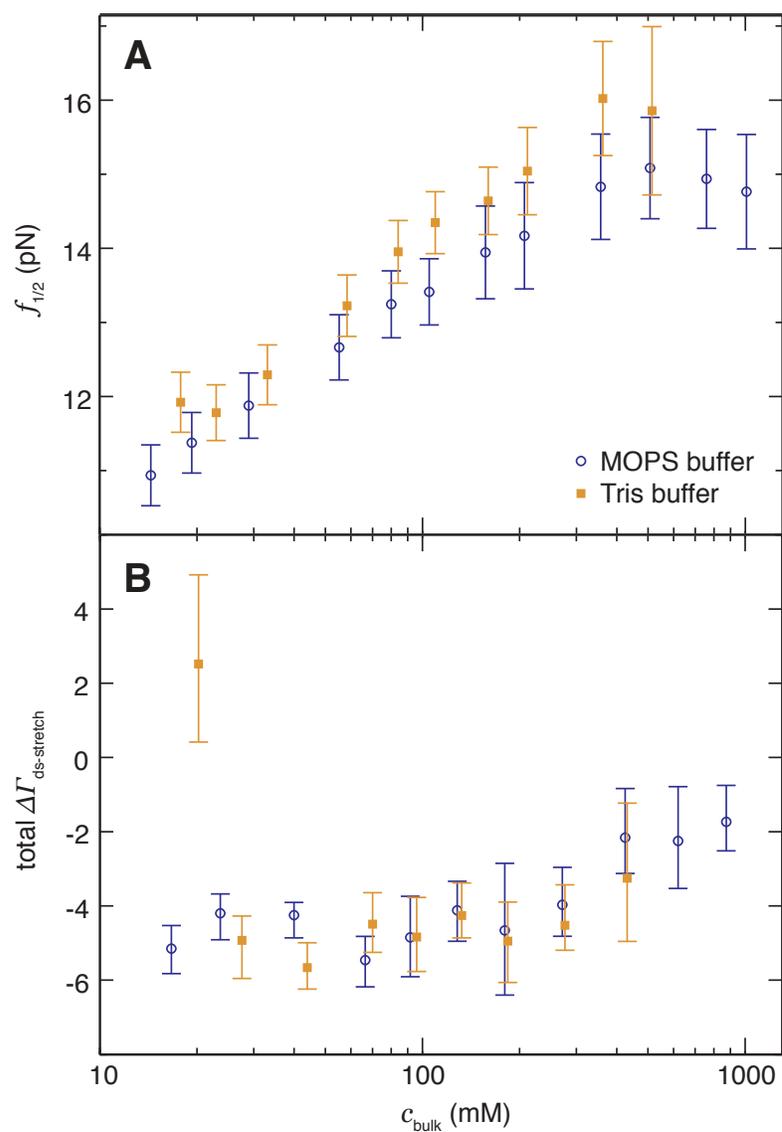


Figure B.2: Comparison of (A) unfolding force and (B) ion excess change upon unfolding determined by the Clausius-Clapeyron method for an RNA hairpin in the presence of 10 mM MOPS buffer (blue circles) and 10 mM Tris-HCl buffer (orange squares).

changes in the buffer used in the experimental solutions. In the prior DNA studies of Dittmore *et al.* [96], 10 mM Tris-HCl buffer was used. However, Leipply *et al.* [194] have warned against using cationic buffers, such as Tris, because of the potential for “idiosyncratic interactions with nucleic acids as compared with group I ions.” Therefore, in these RNA studies, we have used an anionic buffer: 10 mM MOPS. However, because we want to compare our results with the DNA data obtained in cationic buffer, we performed tests of the same RNA hairpin construct (the RNA loop construct described above) in both buffers. The results are shown in Figure B.2 and are consistent within uncertainty at all salt concentrations except the very lowest, where strongly anomalous behavior is observed. We note, however, that this lowest salt concentration (20 mM total ionic strength) is less than the lowest salt concentration in the DNA data, and so we argue that it is reasonable to directly compare the DNA/Tris data with our new RNA/MOPS data.

B.2 Agreement between $\Delta\Gamma_{\text{ds-stretch}}$ methods

Three separate thermodynamic relations can be used to extract the change in ion excess between conformations of a molecule from single-molecule pulling data: one derived from a Maxwell relation, one derived from taking the gradient of the free energy of the system, and one that is a generalized Clausius-Clapeyron relation. All three methods are described in greater detail in Appendix D; see in particular Table D.1. In Figure 6.4D we only report the results from the Clausius-Clapeyron method; however, in Figure B.3 we show that the three methods give comparable results.

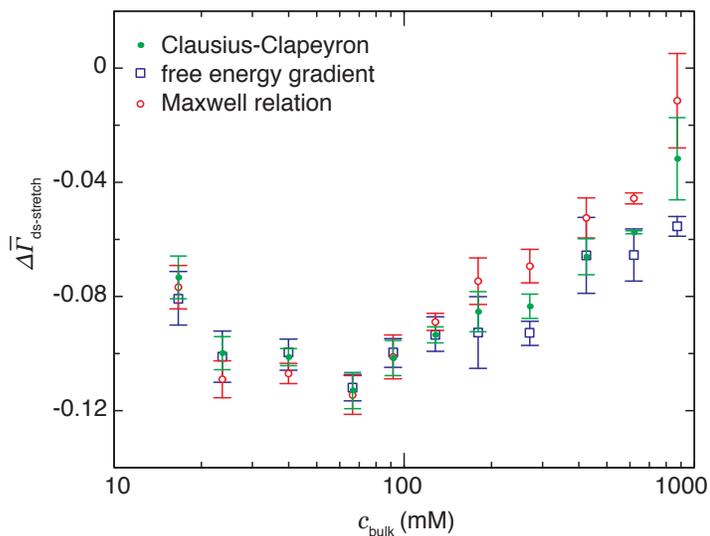


Figure B.3: Comparison of three thermodynamic methods for extracting the change in ion excess due to unfolding of the hairpin. Only the Clausius-Clapeyron data (green points) are used in Figure 6.4D.

B.3 High-salt linear limiting behavior of Γ^{molar}

In experiments conducted in molar units, both electrostatics and the excluded volume of the NA contribute to the ion excess. This is because the ions that would have occupied the volume of the NA are now displaced into the bulk. To the extent that the excluded volume of the NA is unchanged as a function of the bulk salt concentration, the number of ions displaced in this way increases linearly with c_{bulk} . This is illustrated in Figure B.4 (Figure 6.2A, replotted on linear axes), in which a linear decrease in Γ^{molar} is clearly seen at high salt.

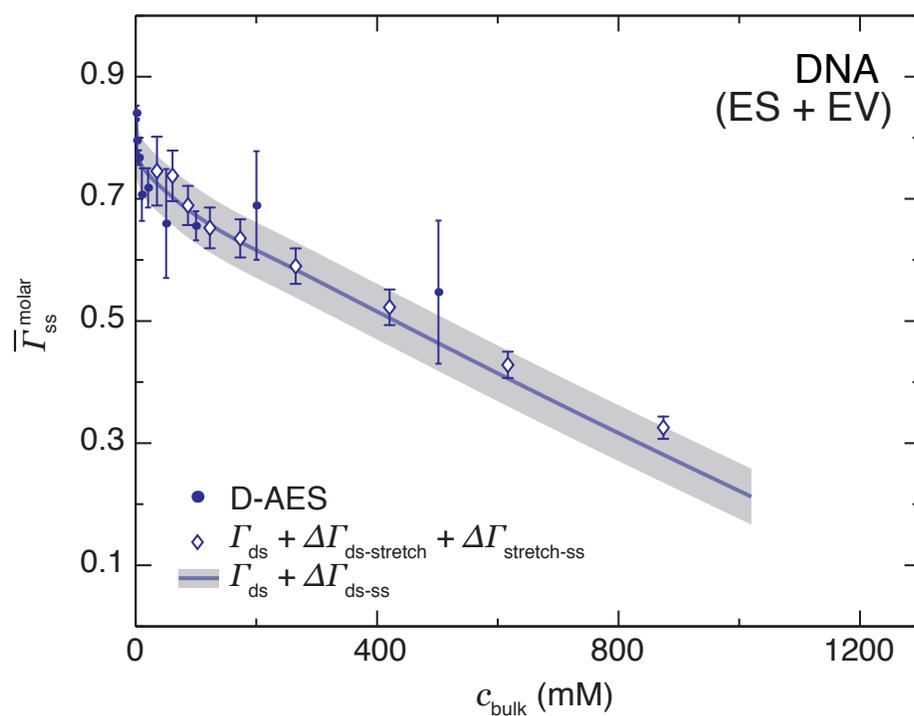


Figure B.4: Values of $\bar{\Gamma}_{ss}^{\text{molar}}$ (i.e., including both electrostatic and excluded volume effects) for DNA, determined directly by D-AES (solid points) and indirectly by thermodynamic cycle analysis using PB calculations and either single-molecule experiments (open diamonds) or melting experiments (line/shaded region). Plotted on linear axes to highlight the linear decrease in the high-salt limit.

Appendix C

Derivation of the uniform mean-field model

In deriving the uniform mean field model of Chapter 6, we consider a system consisting of two compartments, one a reservoir with bulk salt concentration c_{bulk} and the other with a uniform volume charge density ρ (Figure C.1). Because of the excess charge in the one compartment, there will be a potential difference across the interface between the two, given by the Donnan potential Φ_D . Due to this potential, work must be done on negative ions to move them into the charged compartment and, likewise, work must be done to remove positive ions from the compartment:

$$W_+ = -c_+ e \Phi_D, \quad (\text{C.1})$$

$$W_- = +c_- e \Phi_D. \quad (\text{C.2})$$

Here c_{\pm} are the concentrations of ions in the charged compartment, and thus the W values are work per volume. We can relate the ion concentrations in the charged compartment

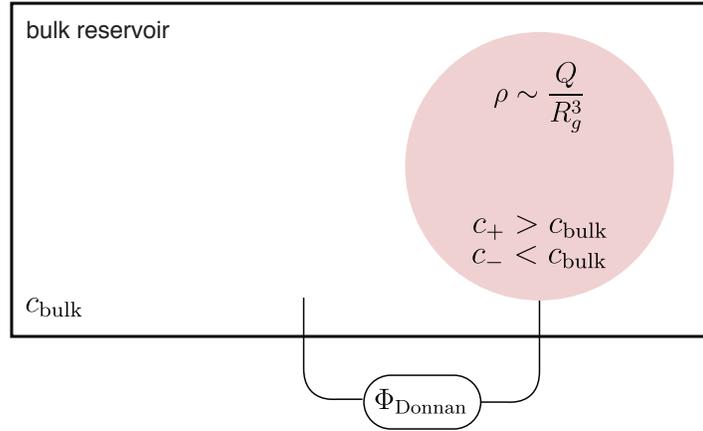


Figure C.1: Diagram of the uniform mean-field model.

to those in the bulk via the Boltzmann equation:

$$c_{\pm} = c_{\text{bulk}} \exp\left(\frac{\pm e\Phi_D}{k_B T}\right). \quad (\text{C.3})$$

Then,

$$\frac{c_-}{c_+} = \exp\left(\frac{-2e\Phi_D}{k_B T}\right). \quad (\text{C.4})$$

Now the problem is reduced to one of finding the value of the Donnan potential. We do this by invoking the Poisson equation, and thus assume mean-field behavior of the ion density. Recall the Poisson equation:

$$\nabla^2 \Phi = \frac{-\rho_{\text{tot}}}{\epsilon_0 \epsilon}. \quad (\text{MKS units}) \quad (\text{C.5})$$

Here, ρ_{tot} is the total charge density, which is, inside of the charged compartment, the sum of the fixed charge and the charge from the ions:

$$\rho_{\text{tot}} = \rho + c_+ e - c_- e. \quad (\text{C.6})$$

Plugging this into the Poisson equation:

$$\nabla^2\Phi = \frac{-1}{\epsilon_0\epsilon} \left[\rho + c_{\text{bulk}}e \exp\left(\frac{\Phi_D e}{k_B T}\right) - c_{\text{bulk}}e \exp\left(\frac{-\Phi_D e}{k_B T}\right) \right], \quad (\text{C.7})$$

$$= \frac{-1}{\epsilon_0\epsilon} \left[\rho + 2c_{\text{bulk}}e \sinh\left(\frac{\Phi_D e}{k_B T}\right) \right]. \quad (\text{C.8})$$

Sufficiently far from the boundary, the potential must be homogeneous within the charged compartment. Thus, $\nabla^2\Phi = 0$ and

$$-\rho = 2c_{\text{bulk}}e \sinh\left(\frac{\Phi_D e}{k_B T}\right). \quad (\text{C.9})$$

Solving this for the Donnan potential gives:

$$\Phi_D = \frac{k_B T}{e} \operatorname{arccsch}\left(\frac{2ce}{\rho}\right). \quad (\text{C.10})$$

We can now plug this into Equation C.4:

$$\frac{c_-}{c_+} = \exp\left[2\operatorname{arccsch}\left(\frac{2c_{\text{bulk}}e}{\rho}\right)\right]. \quad (\text{C.11})$$

We can use the hyperbolic trigonometric identity,

$$\operatorname{arccsch}z = \ln\left(\sqrt{1 + \frac{1}{z^2}} + \frac{1}{z}\right), \quad (\text{C.12})$$

to write this in the form

$$\frac{c_-}{c_+} = \left[\left(\frac{\rho}{2c_{\text{bulk}}e}\right) + \sqrt{1 + \left(\frac{\rho}{2c_{\text{bulk}}e}\right)^2} \right]^2, \quad (\text{C.13})$$

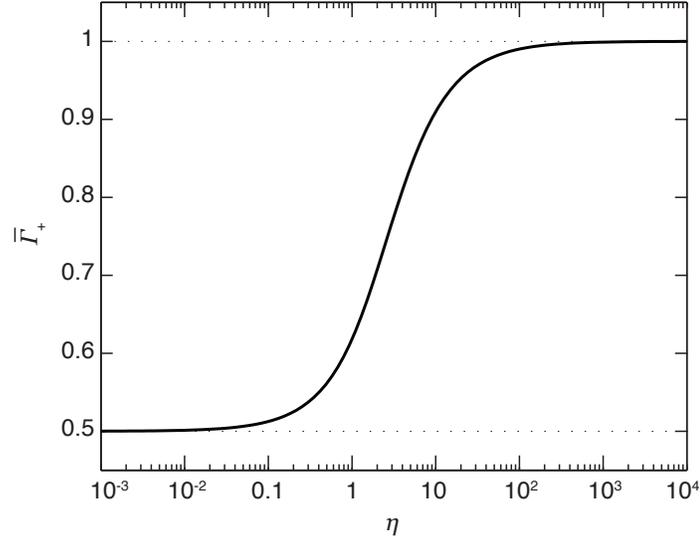


Figure C.2: Plot of Equation C.16, in which $\eta = |\rho|/c_{\text{bulk}}e$.

which is the ratio of anion to cation concentration in the charged compartment. We now want to recast this in the form of the preferential cation interaction coefficient. The c_{\pm} , which are total ion concentrations within the charged region, are related to source-charge-normalized ion excess via

$$\bar{\Gamma}_{\pm} = \frac{(c_{\pm} - c)e}{\rho}. \quad (\text{C.14})$$

Since the combined effect of excess cations and excluded anions must be to offset the source charge, ρ , we have that

$$1 = \bar{\Gamma}_{+} - \bar{\Gamma}_{-}. \quad (\text{C.15})$$

Using Equations C.13 and C.15, we can arrive at Equation 6.3 of the main text:

$$\bar{\Gamma}_{+} = \frac{1}{2} - \frac{1}{\eta} + \sqrt{\frac{1}{4} + \frac{1}{\eta^2}}, \quad (\text{C.16})$$

in which we have made the substitution $\eta = |\rho|/c_{\text{bulk}}e$.

Note that the source charge density and the bulk salt concentration only enter through η as a ratio. This agrees with our intuition, as discussed in Section 5.2.1, that Debye-Hückel behavior (i.e., weak potential/small ρ) will tend also to occur at high c_{bulk} . In Figure C.2 we see the expected molal high-salt limiting behavior: $\bar{\Gamma}_+ \rightarrow 0.5$ at small η (i.e., small ρ /large c_{bulk}). Because the model includes no information about the linear character of the NA, the Oosawa-Manning limit is not recovered at large η ; instead, $\bar{\Gamma}_+ \rightarrow 1$.

Appendix D

Theory of differential ion counting measurements

This chapter is adapted with permission from Ref. [30]. Copyright 2015, American Chemical Society.

In Section 3.1.3 I discussed measurements of the change in ion excess when an NA is stretched (a continuous transition); similarly, in Section 5.3.3 I introduced experiments to measure such a change for NA unfolding (a discrete transition). In this appendix, I rigorously derive and extend the theory of those previously discussed ion counting techniques, which, as will be shown below, are also suitable for measuring the change in binding of any ligand to a macromolecule.

Molecular biology abounds with systems in which ligand-macromolecule binding plays a crucial role, such as in active transport and allosteric regulation in proteins, in aptamer binding in riboswitches [19], and in divalent ion chelation to folded RNAs [177]. Additionally, charged biological macromolecules, such as RNA and DNA, possess a nonspecif-

ically associated atmosphere of excess counterions and depleted coions that, through electrostatic screening, strongly modulates the macromolecule's ability to adopt compact conformations [177]. Detailed, quantitative descriptions of ligand binding interactions, in their various forms, are thus central to understanding a wide variety of biological phenomena. At the experimental level, such descriptions come in the form of stoichiometries and free energies of ligand association.

A host of techniques are used to obtain these data. For macromolecule-ligand interactions, these include isothermal titration calorimetry, surface plasmon resonance, and techniques involving fluorescently or radioactively labeled ligands [197]. Macromolecule-ion interactions can be measured using dialysis or titration in the presence of an ion-binding dye [71], ASAXS [141], and nuclear magnetic resonance-based techniques [198]. Each technique has its own advantages and limitations, which can include suitability for only certain ligands or specially labeled ligands, the ability to probe only certain states of the macromolecule, and the inability to probe both stoichiometry and energetics simultaneously.

Recently, a new technique has emerged for measuring ligand binding stoichiometry, including nonspecific association of ions, through single-molecule mechanical stretching of biological polymers [75, 76, 96, 157]. The technique rests on the thermodynamic relationship between the two experimental control parameters, stretching force and bulk chemical potential, and their conjugates, polymer extension and number of associated ligands. The principal advantage of this method is that, because it relies only on measuring the force vs. extension relationship for the macromolecule as a function of ligand concentration, it is general for any ligand and no special modification of the ligand is required. In addition, the measurement is model-independent in terms of the details of the ligand binding; e.g., number of binding sites, etc. However, because of the nature of the underlying thermodynamic argument, only *changes* in ligand stoichiometry between

conformational states are accessible, not absolute values. That being said, such results are still of interest, since in many systems ligand binding is coupled to a conformational change. For example, binding of the target ligand to a riboswitch aptamer can modulate the aptamer’s folding pathway, skewing the equilibrium between the unfolded and folded states [199].

In Section D.1 we review the derivation of the three thermodynamic relations that underlie this method [75, 76, 96, 157]. In the process we discuss new generalizations of the theory to allow for systems with an arbitrary number of ligand species and for systems that are studied in a thermodynamic ensemble besides the constant-force ensemble. In Section D.2 we further extend the theory to consider “force-ramp” or “rupture force” experiments, which are performed under non-equilibrium conditions. In Section D.3 we show that the same mechanical manipulation data also yield the difference in free energy of ligand binding between different conformations of the macromolecule. In Section D.4 we analyze previously reported single-molecule nucleic acid data to obtain these free energy differences; we then interpret these newly computed results.

D.1 Equilibrium stoichiometry measurements

The experiment under consideration is the salt-dependent measurement of molecular elasticity that has been discussed throughout this dissertation. In such a study, a polymeric macromolecule is stretched in the presence of a bulk concentration of ligand characterized by its chemical potential, μ . In one type of such an experiment, a constant applied force, f , stretches the macromolecule while the time-average of the fluctuating extension, $\langle X \rangle$, is measured. This isotensional experiment can be realized using an MT apparatus or through OT or AFM [29] in a “force clamp” configuration [200]. Alternatively, the ends of the macromolecule can be fixed a certain distance, X , apart while the

Table D.1: Ligand counting formulae for a single ligand species, in the isotensional ensemble

method	formula ^a	data required
Maxwell relation ^b	$\Delta\Gamma = \frac{1}{k_B T A(c_{\text{bulk}})} \int_{f_1}^{f_2} \left(\frac{\partial \langle X \rangle}{\partial \ln c_{\text{bulk}}} \right) df$	series of extension vs. force curves at different bulk ligand concentrations (e.g., Figure 1.2B)
free energy gradient ^c	$\Delta\Gamma = \frac{-1}{A(c_{\text{bulk}})} \left(\frac{\partial K}{\partial \ln c_{\text{bulk}}} \right)$	constant-force, two-state folding/unfolding traces at various pulling forces and ligand concentrations (e.g., Figure 1.2C)
Clausius-Clapeyron ^c	$\Delta\Gamma = \frac{\Delta X}{k_B T A(c_{\text{bulk}})} \left(\frac{\partial f}{\partial \ln c_{\text{bulk}}} \right)_K$	slope of a curve of constant K in force-concentration space

^a Variables used: k_B , Boltzmann constant; T , temperature; A , ligand activity correction factor; X , macromolecule extension; f , applied force; c_{bulk} , bulk ligand concentration; K , two-state equilibrium constant.

^b Applies generally; here $\Delta\Gamma$ is the change in ligand excess when the macromolecule is stretched from f_1 to f_2 .

^c Applies to a macromolecule exhibiting two discrete conformational states; here $\Delta\Gamma$ is the change in ligand excess between those states.

time-average of the fluctuating tension, $\langle f \rangle$, is measured. This isometric experiment can be realized using an optical trap or AFM in its standard, passive configuration. Below, we first consider an experiment performed in the isotensional ensemble and derive a series of relations needed to extract ligand binding stoichiometry from those data (results summarized in Table D.1); we then go on to consider the isometric case.

D.1.1 Thermodynamic description of the isotensional experiment

As described by Landy *et al.* [75], we consider a volume, V , containing a single macromolecule, populations of the solute species (e.g., ligands, ions) and the solvent.

This volume, at equilibrium with a bulk reservoir, is large enough that at its periphery all local concentrations of solvent and solute are equal to their bulk concentrations; i.e., all macromolecule-induced perturbations have dissipated. The macromolecule is characterized by its fluctuating end-to-end extension, X , and by the fixed applied force, f ; each molecular species is characterized by the fluctuating number present in V , \mathcal{N}_i , and by its fixed chemical potential in the bulk, μ_i . Since this is a grand canonical ensemble, the partition sum is written

$$\mathcal{Z} = \sum_{\{\mathcal{N}_i\}} \int_X e^{-[F_{\text{int}}(X, \{\mathcal{N}_i\}) - fX - \sum_i \mu_i \mathcal{N}_i] / k_B T}, \quad (\text{D.1})$$

where F_{int} is the internal free energy of the volume at fixed X and $\{\mathcal{N}_i\}$. From this partition sum, we can isolate the relevant thermodynamic potential (potential of mean extension [201]), the variation of which is given by

$$d\Omega = -\langle X \rangle df - \sum_{\text{species}} \langle \mathcal{N}_i \rangle d\mu_i, \quad (\text{D.2})$$

where the brackets denote ensemble averages. In experimental practice, such ensemble averages are replaced by time averages under an assumption of ergodicity.

The molecular species indexed by i in Equation D.2 include the various solutes as well as the solvent. We eliminate explicit consideration of the solvent by invoking the Gibbs-Duhem relation at constant temperature and pressure,

$$\sum_i \mathcal{N}_i d\mu_i = 0 \implies d\mu_w = \sum_{\text{solutes}} \frac{c_i}{c_w} d\mu_i, \quad (\text{D.3})$$

where the $\{c_i\}$ are number concentrations and the subscript w refers to the solvent; e.g., water. Using this result, the free energy differential can be written in terms of a sum

over only the solutes,

$$d\Omega = -\langle X \rangle df - \sum_{\text{solute}} \langle \Gamma_i \rangle d\mu_i, \quad (\text{D.4})$$

where the $\{\Gamma_i\}$ are the molecular excesses,

$$\Gamma_i \equiv \mathcal{N}_i - \frac{c_i}{c_w} \mathcal{N}_w = \mathcal{N}_i - c_i V - \frac{c_i}{c_w} (\mathcal{N}_w - c_w V). \quad (\text{D.5})$$

The third term in this definition is typically small, except in the case of very large solute concentrations or strong solute excluded volume effects. In the limit where this third term can be neglected, Γ_i can be thought of as the excess number of molecules of species i present in the vicinity of the macromolecule beyond the number that would be expected due to the bulk concentration alone; thus, it is the quantity we seek: a measure of the stoichiometry of ligand association.

The ion-association literature often invokes a related quantity, the preferential ion interaction coefficient given, in terms of macroscopic parameters, by Equation 5.1. This quantity, in the $m_{\text{NA}} \rightarrow 0$ limit relevant to single-molecule studies, is equivalent to the molecular excess of Equation D.5 in the limit discussed above. For this reason, we represent both quantities with the symbol Γ throughout this dissertation and refer to them, interchangeably, as either the preferential ion interaction coefficient or the ion excess.

We can express the chemical potentials in terms of the experimentally relevant number-concentrations, $\{c_i\}$, and the activity coefficients [202], $\{\gamma_i(c)\}$, specifying non-ideal mixing behavior:

$$\mu_i = k_B T \ln(\gamma_i(c)c_i) \implies d\mu_i = k_B T A_i(c_i) d \ln(c_i), \quad (\text{D.6})$$

where $A_i(c_i) \equiv 1 + c_i \frac{\partial}{\partial c_i} \ln(\gamma_i)$. At last, substituting this result into Equation D.4 gives

the free energy of the system in terms of experimentally relevant variables:

$$d\Omega = -\langle X \rangle df - k_B T \sum_{\text{solutes}} \langle \Gamma_i \rangle A_i(c_i) d \ln(c_i). \quad (\text{D.7})$$

This is the general thermodynamic description of the constant-force pulling experiment that we will now use as the starting point to derive relations for determining ligand molecular excess.

D.1.2 Ligand binding relation for a continuously deformable macromolecule (Maxwell relation)

We first consider a macromolecule that deforms continuously (i.e., with no sudden conformational transitions) as the applied stretching force is varied; e.g., the ssRNA stretching experiments of Chapter 3. As did Zhang and Marko [76], for general ligand binding, and Landy *et al.* [75], for ion association, we approach this case through the derivation of Maxwell relations. Equating mixed partial derivatives of the free energy (Equation D.7) with respect to f and $\ln c_i$ gives

$$\left. \frac{\partial \langle X \rangle}{\partial \ln c_i} \right|_f = k_B T \left[A_i(c_i) \frac{\partial \langle \Gamma_i \rangle}{\partial f} + \sum_{j \neq i} A_j(c_j) \frac{\partial \langle \Gamma_j \rangle}{\partial f} \frac{\partial \ln c_j}{\partial \ln c_i} \right]_{\{c_i\}}, \quad (\text{D.8})$$

which represents a set of equalities, one for each of the solute species indexed by i .

When there is only a single ligand species present, or when the bulk concentrations of the various ligands do not depend upon one another, the second term on the right-hand side of Equation D.8 is zero. In this case integral expressions can be obtained for the change in ligand molecular excess between conformations of the macromolecule occurring

at applied forces f_1 and f_2 and under constant bulk ligand concentrations $\{c_i\}$:

$$\Delta\Gamma_{i,f_1 \rightarrow f_2}|_{\{c_i\}} = \frac{1}{k_B T A_i(c_i)} \int_{f_1}^{f_2} \left(\frac{\partial \langle X \rangle}{\partial \ln c_i} \right)_f df, \quad (\text{D.9})$$

where $\Delta\Gamma_{i,f_1 \rightarrow f_2} = \langle \Gamma_i \rangle_{f_2} - \langle \Gamma_i \rangle_{f_1}$.

If the solute concentrations cannot be varied independently, the second term on the right-hand side of Equation D.8 becomes important and more complicated integral expressions will be obtained that no longer allow for the unique determination of each $\Delta\Gamma_i$. Instead, only linear combinations of the molecular excess changes will be accessible, unless there is an additional constraint on the values of $\{\Gamma_i\}$.

The association of monovalent ions to a polyelectrolyte [75] (i.e., the case of ion counting considered in this dissertation) is an example of such an instance of linearly dependent solute concentrations. To preserve bulk charge neutrality in this system, all positive cations must be offset by negative anions; therefore, their two concentrations, c_+ and c_- , must be equal, implying that $\partial \ln c_+ / \partial \ln c_- = 1$ in Equation D.8. In terms of the single overall concentration of ion pairs, c_{bulk} , an expression can be derived for the combined ion excess change of both species,

$$A_+ \Delta\Gamma_+ + A_- \Delta\Gamma_-|_c = \frac{1}{k_B T} \int_{f_1}^{f_2} \left(\frac{\partial \langle X \rangle}{\partial \ln c_{\text{bulk}}} \right)_f df, \quad (\text{D.10})$$

where $\Delta\Gamma_{\pm}$ and A_{\pm} refer to the ion excess and activity factor values, respectively, for the cationic and anionic species. In general, we would have to stop here knowing only the value of the combined quantity on the left-hand side of Equation D.10; however, in this case we have an additional condition of local charge neutrality: that $\langle \Gamma_+ \rangle - \langle \Gamma_- \rangle = Q_o$, where Q_o is the charge of the macromolecule. This condition implies that $\Delta\Gamma_+ = -\Delta\Gamma_-$, which allows the individual ion excess changes to be backed out of Equation D.10.

D.1.3 Ligand binding relations for a macromolecule inhabiting discrete conformational states

We now consider macromolecules that jump between discrete conformational states, characterized by distinct average extensions, as the applied stretching force is varied; e.g., the folding/unfolding of an RNA hairpin in Chapter 6. We will derive two relations for measuring the change in ligand excess between conformational states in such a system: one arising from taking gradients of the free energy with respect to the various ligand concentrations and the other a generalized Clausius-Clapeyron relation in force-concentration space. While thermodynamically equivalent, these relations use the experimental data in different ways (Table D.1) and are thus complementary.

In addition, we note that the Maxwell relation-based treatment discussed above remains valid in the case of discrete states. This is because, as illustrated in Figure 1.2B, thermal fluctuations broaden the transitions between the states in the $\langle X \rangle (f)$ plot, yielding smooth curves that are still amenable to the application of Equation D.8.

Free energy gradient method

From the free energy differential (Equation D.7), we can write the change in the total free energy with respect to changes in the concentration of any of the solutes: (generalization of the results of Dittmore *et al.* [96])

$$\left. \frac{\partial \Omega}{\partial \ln c_i} \right|_{f, \{c_{j \neq i}\}} = -k_B T \left[A_i(c_i) \langle \Gamma_i \rangle + \sum_{j \neq i} A_j(c_j) \langle \Gamma_j \rangle \frac{\partial \ln c_j}{\partial \ln c_i} \right]. \quad (\text{D.11})$$

We are considering systems exhibiting discrete transitions between any countable number of states. However, due to the principle of microscopic reversibility [202], transitions of arbitrary complexity between such states can be treated as the composition of many

two-state transitions, $I \rightarrow II$, each characterized by a molecular excess change, $\Delta\Gamma_i = \langle\Gamma_{i,II}\rangle - \langle\Gamma_{i,I}\rangle$.

If the solute concentrations do not depend explicitly on one another, we can drop the last term in Equation D.11 and obtain, by specific evaluation for two states,

$$\Delta\Gamma_i = \frac{-1}{k_B T A_i(c_i)} \left(\frac{\partial(\Omega_{II} - \Omega_I)}{\partial \ln c_i} \right)_{f, \{c_{j \neq i}\}}. \quad (\text{D.12})$$

The free energy difference, $\Delta\Omega_{I \rightarrow II} = \Omega_{II} - \Omega_I$, can be written in terms of the ratio of the probabilities of observing the system in each of the two states, $P_I/P_{II} = e^{-\Delta\Omega_{I \rightarrow II}/k_B T}$, which defines the two-state equilibrium constant, K , in the single-molecule context and is accessible from traces like those in Figure 1.2C. Then, Equation D.12 can be rewritten as

$$\Delta\Gamma_i = \frac{-1}{A_i(c_i)} \left(\frac{\partial K_{I \rightarrow II}}{\partial \ln c_i} \right)_{f, \{c_{j \neq i}\}}, \quad (\text{D.13})$$

which is, again, the change in ligand excess between two conformational states, I and II , at constant bulk ligand concentration.

If there is a linear dependence between the various solute concentrations, such as when considering ions as the ligands, the last term of Equation D.11 cannot be neglected and, as for the continuously deformable macromolecule, we obtain coupled equations that can only be separated if there is an additional constraint on the values of $\{\Gamma_i\}$:

$$\Delta\Gamma_i = \frac{-1}{k_B T A_i} \frac{\partial K_{I \rightarrow II}}{\partial \ln c_i} + \sum_{j \neq i} \frac{A_j}{A_i} \Delta\Gamma_j \frac{\partial \ln c_j}{\partial \ln c_i}. \quad (\text{D.14})$$

Note also that if it were already known that one of the two conformational states were incompetent for ligand binding, then we would have that, e.g., $\langle\Gamma_{i,I}\rangle = 0$ and $\Delta\Gamma_i = \langle\Gamma_{i,II}\rangle$. Then, in this limit the differential ligand excess change measured by this technique would be equal to the absolute excess of bound ligands.

Clausius-Clapeyron method

The canonical Clausius-Clapeyron relation is an expression for the slope of the coexistence curve in a pressure-temperature phase diagram. The relation derives from the fact that, at coexistence, the free energies of two phases must be equal. The single-molecule equivalent of the bulk phenomenon of coexistence occurs when two states occur with equal probability, $P_I = P_{II}$; this is approximated in the middle pane of Figure 1.2C and will also be referred to as “coexistence”. Under this single-molecule coexistence condition, the free energies of the two states, Ω_I and Ω_{II} , must also be equal, since $P_I/P_{II} = e^{-(\Omega_I - \Omega_{II})/k_B T}$. Thus, when the parameters f and $\{c_i\}$ are varied such that equal coexistence is preserved, or such that K is held at any constant value, we must have that $d\Omega_I = d\Omega_{II}$.

Applying this constraint to the free energy differential (Equation D.7), we derive a relation between infinitesimal changes in force, df , and concentration, $d \ln c_i$, akin to the canonical Clausius-Clapeyron relation between infinitesimal changes in pressure and temperature: (generalization of the results of Todd and Rau [157] and Dittmore *et al.* [96])

$$\Delta X df = k_B T \sum_{\text{solutes}} \Delta \Gamma_i A_i(c_i) d \ln c_i, \quad (\text{D.15})$$

where $\Delta X = \langle X_I \rangle - \langle X_{II} \rangle$ is the difference in mean extension between the states. The allowed values of f and $\{c_i\}$ satisfying the condition of constant K can be thought of as a generalized surface embedded within the higher-dimensional force-concentration space. If the solute concentrations are linearly independent, then $f \cup \{c_i\}$ constitutes a basis of this space and it is possible to consider variations of f and one particular c_i at a time, holding the other $\{c_{j \neq i}\}$ constant. In this case we obtain an uncoupled expression for the

molecular excess change of each ligand between the two conformational states,

$$\Delta\Gamma_i = \frac{\Delta X}{k_B T A_i(c_i)} \left(\frac{\partial f}{\partial \ln c_i} \right)_K. \quad (\text{D.16})$$

If the solute concentrations are linearly dependent, there exists a lower-dimensional basis of the force-concentration space. Coupled expressions for the $\{\Delta\Gamma_i\}$ can then be obtained by rewriting Equation D.15 in terms of this basis.

While in this method, again owing to the principle of microscopic reversibility, many-state systems are treated as compositions of two-state systems, additional information can be extracted from triple, or higher order, points where multiple states are occupied with equal probability; i.e., the intersection of multiple coexistence surfaces. In analogy to the latent heat in the canonical Clausius-Clapeyron system [203], the change in molecular excess must be zero when integrated in a closed path around such a point. Thus, if $N - 1$ values of $\Delta\Gamma$ are known in the vicinity of an N -point, the N^{th} value can be obtained according to

$$\Delta\Gamma_{1,2} + \Delta\Gamma_{2,3} + \dots + \Delta\Gamma_{N,1} = 0, \quad (\text{D.17})$$

where the subscript numbers refer to the various states.

D.1.4 Thermodynamic description of the isometric experiment

Our above results derived from the differential of the thermodynamic potential relevant to constant-force experiments, the potential of mean extension (Equation D.2). In another relevant class of single-molecule experiments, it is the force, rather than the extension, that is the fluctuating thermodynamic quantity; e.g., in passive optical trapping experiments. In this case, the relevant thermodynamic potential is the potential of mean

force, F [201, 204]. Akin to Equation D.2, the variation in this potential is

$$dF = \langle f \rangle dX - \sum_i \langle \mathcal{N}_i \rangle d\mu_i, \quad (\text{D.18})$$

from which ligand counting relations analogous to those for the isotensional ensemble can be derived through the same arguments.

More correctly, in optical trapping experiments it is generally the trap position, and not the macromolecule extension, that is fixed. In this “mixed” ensemble, both f and X are fluctuating quantities; although, they are coupled through the stiffness of the trap. Thus, ligand counting relations would derive from a thermodynamic potential having the trap separation as the independent variable; the thermodynamics of this ensemble are discussed extensively by Manosas and Ritort [205].

D.2 Non-equilibrium measurements

The ligand-counting relations we have derived thus far (Equations D.9, D.13 and D.16, for linearly independent ligand concentrations) depend upon measurement, in the isotensional ensemble, of the mean equilibrium extension, $\langle X \rangle (f)$, or the two-state equilibrium constant, $K(f)$. However, for various reasons (e.g., slow kinetics), single-molecule pulling experiments cannot always be performed at equilibrium; in these cases non-equilibrium methods must be used. An example is a force-ramp experiment, in which extension data are collected as the applied force is increased at a uniform rate faster than the characteristic equilibration time of the macromolecule.

In order to analyze such non-equilibrium data using our ligand-counting relations, which derive from equilibrium thermodynamics, we can employ the Jarzynski equality,

[206]

$$e^{-\Delta\mathcal{F}/k_B T} = \langle e^{-W/k_B T} \rangle, \quad (\text{D.19})$$

which relates the Boltzmann factor of a thermodynamic potential, $\Delta\mathcal{F}$, to the average, over sufficiently many repetitions, of the Boltzmann factor of a suitably defined non-equilibrium work, W . This equality has been employed previously to analyze non-equilibrium single-molecule pulling experiments; e.g., by Liphardt *et al.* [207] and Harris *et al.* [208]. It should be noted that the number of repetitions needed, such that averaging over W converges to ΔF , increases with the dissipated work, and thus the extent of non-equilibrium behavior that can be probed is limited by the number of repetitions that are experimentally feasible [207].

In a force-ramp experiment, the relevant thermodynamic potential is the potential of mean extension, $\Delta\Omega$ (Equation D.2), which is related to an integral of the equilibrium extension vs. force curve, $\langle X \rangle(f)$, [201, 204]

$$\Delta\Omega(f) = \int_{f_0}^f \langle X \rangle(f') df', \quad (\text{D.20})$$

where f_0 and f are the starting and ending forces of the ramp, respectively. In analogy to the traditional definition of mechanical work, $\int f dx$, corresponding to the potential of mean force, we can write down a generalized work corresponding to the potential of mean extension (see Equation 3 of Jarzynski [206]):

$$W(f) = \int_{f_0}^f X_{\text{ne}}(f') df', \quad (\text{D.21})$$

where X_{ne} is the measured non-equilibrium extension during the force ramp.

The Jarzynski equality (Equation D.19), through Equations D.20 and D.21, can be used to relate the equilibrium extension vs. force curve, which is needed for the ligand

counting relations, to an average over non-equilibrium traces, which are measured in a force-ramp experiment:

$$\int_{f_0}^f \langle X \rangle (f') df' = -k_B T \ln \left[\left\langle e^{-\int_{f_0}^f X_{\text{ne}}(f') df' / k_B T} \right\rangle \right]. \quad (\text{D.22})$$

If the system is brought to equilibrium at $f = f_0$ prior to the start of non-equilibrium pulling, then $\langle X \rangle (f_0) = X_{\text{ne}}(f_0)$ and the equilibrium extension at all subsequent forces can be found by differentiation of the above:

$$\langle X \rangle (f) = -\frac{d}{df} \left[k_B T \ln \left(\left\langle e^{-\int_{f_0}^f X_{\text{ne}}(f') df' / k_B T} \right\rangle \right) \right] + X_{\text{ne}}(f_0). \quad (\text{D.23})$$

Once obtained, the equilibrium extension vs. force curves, $\langle X \rangle (f)$, can be used directly with the integrated Maxwell relation for ligand counting (Equation D.9). Application of multi-state models to these curves, like the dashed lines in Figure 1.2B, can be used to extract the ΔX and $K(f)$ parameters needed for the free energy gradient (Equation D.13) and Clausius-Clapeyron (Equation D.16) methods. Such discrete state analysis can be simplified by assuming that conformational rearrangement of the unstructured biopolymer occurs on a much faster timescale than folding/unfolding behavior; non-equilibrium effects would then occur only at the transitions between states, manifesting as rupture forces that differ from the corresponding equilibrium two-state coexistence forces. This condition on the hierarchy of relevant timescales, which is reflected by a lack of hysteresis in the force-extension curves away from the transitions, is frequently met; e.g., in the RNA stretching experiments of Bizarro *et al.* [209].

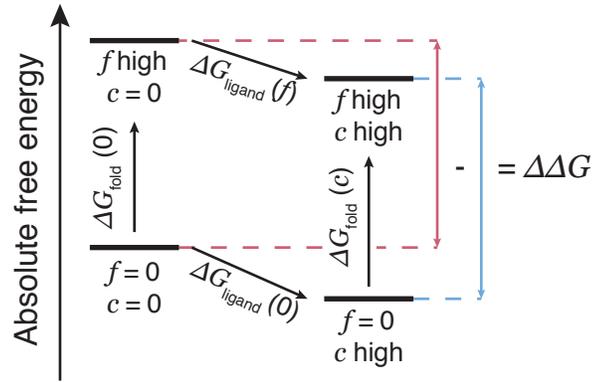


Figure D.1: Diagram of the thermodynamic cycle discussed. Based upon Figure 1 of Grilley et al. [210].

D.3 Extraction of free energies of ligand binding

The theory we have discussed so far enables the extraction of changes in ligand stoichiometry as a single molecule undergoes a force-induced conformational transition; we now seek to determine the free energy change associated with this change in stoichiometry.

Following Draper and colleagues [210, 211], who build upon the arguments of Schellman [212], we can write the free energy change associated with the binding of a particular ligand, ΔG_{ligand} , by integrating the total free energy of the system (Equation D.7), holding force and the other ligand concentrations constant:

$$\Delta G_{\text{ligand}} = -k_B T \int_{-\infty}^{\ln c} \Gamma(\ln c') d \ln c'. \quad (\text{D.24})$$

For simplicity in the ensuing discussion, we will assume that the system is in the limit where $A \approx 1$. A four-state thermodynamic cycle representative of the system under discussion is shown in Figure D.1. ΔG_{ligand} is the free energy change associated with transitioning from a system without ligand to a system with some bulk ligand concentration, holding f constant; these are the diagonal arrows in Figure D.1. The molecular

excess changes, $\Delta\Gamma$, discussed in the preceding sections were for transitions between conformational states at constant bulk concentration; these transitions correspond to the vertical arrows in Figure D.1.

Note, however, that our stoichiometry analysis (Section D.1) does not yield the absolute molecular excess, Γ , but only the change in that value between conformational states, $\Delta\Gamma$. We must therefore break the integral into two parts:

$$\Delta G_{\text{ligand}} = -k_B T \int_{-\infty}^{\ln c} [\Gamma_0(\ln c') + \Delta\Gamma(\ln c')] d \ln c', \quad (\text{D.25})$$

where $\Delta\Gamma$ depends on the conformation of the macromolecule, whereas Γ_0 does not. We then define the quantity

$$\Delta\Delta G \equiv -k_B T \int_{-\infty}^{\ln c} \Delta\Gamma(\ln c) d \ln c, \quad (\text{D.26})$$

which is the binding free energy of only those ligands that are associated with the conformation change corresponding to $\Delta\Gamma$; i.e., the difference in stability, between two conformations, due to ligand binding. In Figure D.1, $\Delta\Delta G$ corresponds to the difference between $\Delta G_{\text{ligand}}(f)$ and $\Delta G_{\text{ligand}}(0)$. In other words, it is the free energy of ligand binding to a particular conformation, referenced to the free energy of ligand binding to the zero-force conformation.

Note that the lower limit of the integral in Equation D.26 goes to negative infinity. Since we are interested in the free energy of stabilization associated with the addition of ligands, it is natural to reference our measurements to zero ligand concentration, as in Figure D.1. This presents a problem, however, when the ligand concentration cannot be set arbitrarily low due to experimental limitations. An example of this arises in ion measurements, where an arbitrarily low ion concentration would lead to extreme electro-

static effects. In this case, it is necessary to choose a non-zero reference concentration, understanding that all computed $\Delta\Delta G$ will differ from their absolute values by some constant summand.

In an alternative approach, it is possible to access $\Delta\Delta G$ from integration of $X(f)$, rather than $\Gamma(\ln c)$, data, as per Bizarro *et al.* [209]. We can obtain the free energy change associated with stretching the macromolecule, $\Delta G_{\text{stretch}}$, by integrating Equation D.4 with *all* ligand concentrations held constant:

$$\Delta G_{\text{stretch}} = \int_0^f X(f')df'. \quad (\text{D.27})$$

This quantity corresponds to the vertical transitions in Figure D.1 and will, generally, depend on ligand concentration. Because of the closed nature of this thermodynamic cycle, one can show that the free energy of the conformational change associated with ligand binding, $\Delta\Delta G$, can also be written in terms of these $X(f)$ integrals: [210]

$$\Delta\Delta G = \Delta G_{\text{stretch}}(c) - \Delta G_{\text{stretch}}(0), \quad (\text{D.28})$$

again requiring information about the macromolecule under zero-ligand conditions, which may not be accessible. Inspection of the two methods (Equations D.26 and D.28) reveals that they are mathematically identical for all three ways of determining $\Delta\Gamma$.

Our $\Delta\Delta G$ is equivalent to the $\Delta\Delta G$ quantity derived by Grilley *et al.* [210]; however, whereas in the studies we describe here the conformation of the macromolecule is controlled via a pulling force, in that work the macromolecule was trapped in a particular conformation through structural mutation.

D.4 $\Delta\Delta G$ from RNA and DNA data

Using the theory of Section D.3, the NA SMFS data discussed throughout this dissertation can now also be viewed from the perspective of the free energy of stabilization afforded to the macromolecular conformation due to the ion atmosphere as a function of bulk concentration.

D.4.1 ssRNA, ssDNA and dsDNA

Figures D.2A and D.2B show the change in ion excess, as a function of bulk salt concentration, as the unstructured nucleic acids are stretched from low to high force [26, 75]. The ssRNA and ssDNA curves both decrease as a function of bulk salt, but with different slopes, a finding that has been attributed to their different stiffnesses [26, 75]. However, the dsDNA curve is an increasing function of bulk concentration. Landy *et al.* have argued that this occurs because dsDNA is dominated by intrinsic, mechanical stiffness whereas ssDNA and ssRNA are dominated by electrostatic stiffness [75].

We can now obtain the free energy change due to ions as the nucleic acids transition from a compact (low-force) to a stretched (high-force) conformation, using the methods of Section D.3. As discussed previously, it is impractical to experimentally probe the $\ln c_{\text{bulk}} \rightarrow -\infty$ limit of the integral in Equation D.26 or of the second term in Equation D.28. Instead, we integrate starting at the lowest concentration studied (50 mM Na^+ for ssDNA/ssRNA, 0.4 mM tris-H^+ for dsDNA); the $\Delta\Delta G$ values reported in Figures D.2C and D.2D are thus offset from their absolute values by some constant summand.

For all three species we obtain $\Delta\Delta G$ vs. $\ln c$ curves with negative slope. We note that, at stretching forces corresponding to both the compact and stretched states, increasing bulk salt concentration causes the nucleic acid to adopt a more compact conformation (Figure D.3A). As such, we expect that increasing bulk ion concentration has a net

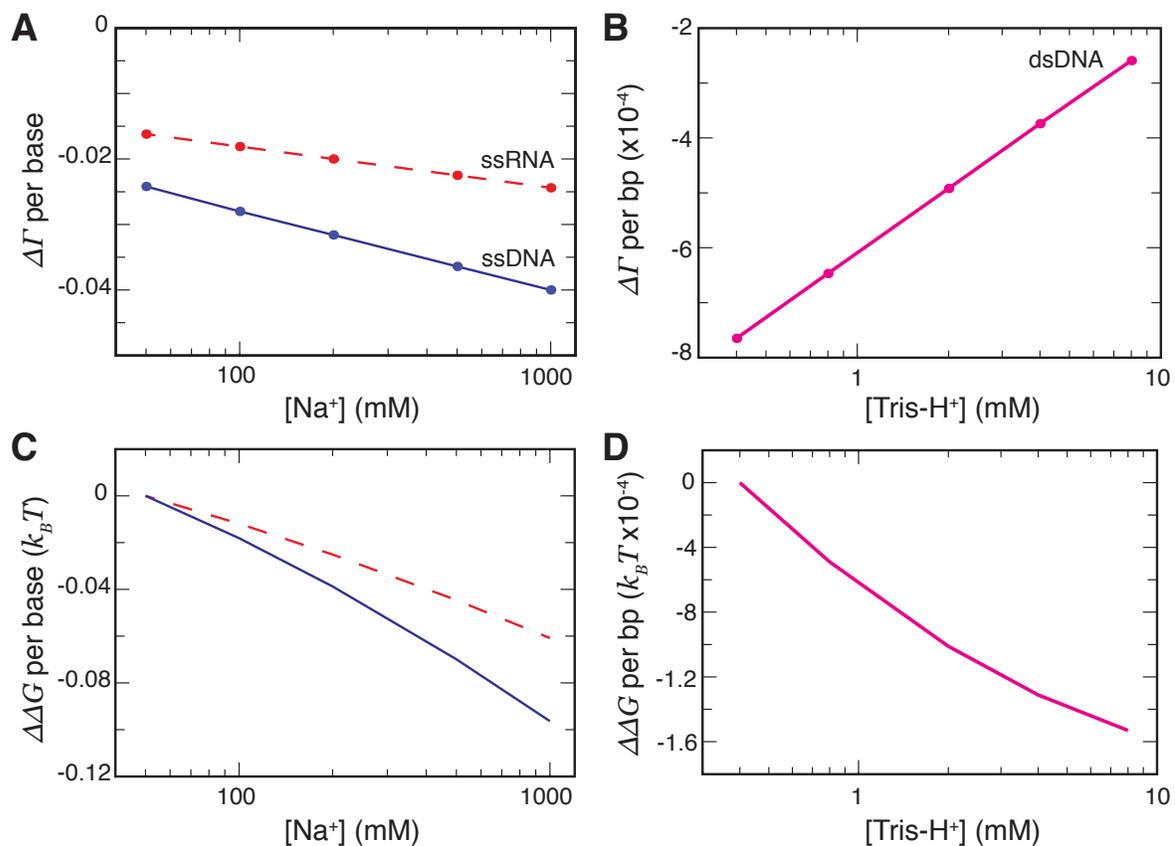


Figure D.2: Application of Equation D.26 to obtain free energy of ion stabilization of conformation for unstructured nucleic acids. (A, B) Change in number of associated monovalent cations as ssRNA [26] (dashed red) and ssDNA [75] (solid blue) are stretched from 0 to 10 pN and as dsDNA [75] (solid magenta) is stretched from 0 to 3 pN. (C, D) Corresponding values of $\Delta\Delta G$, using the lowest ion concentration as the zero-reference.

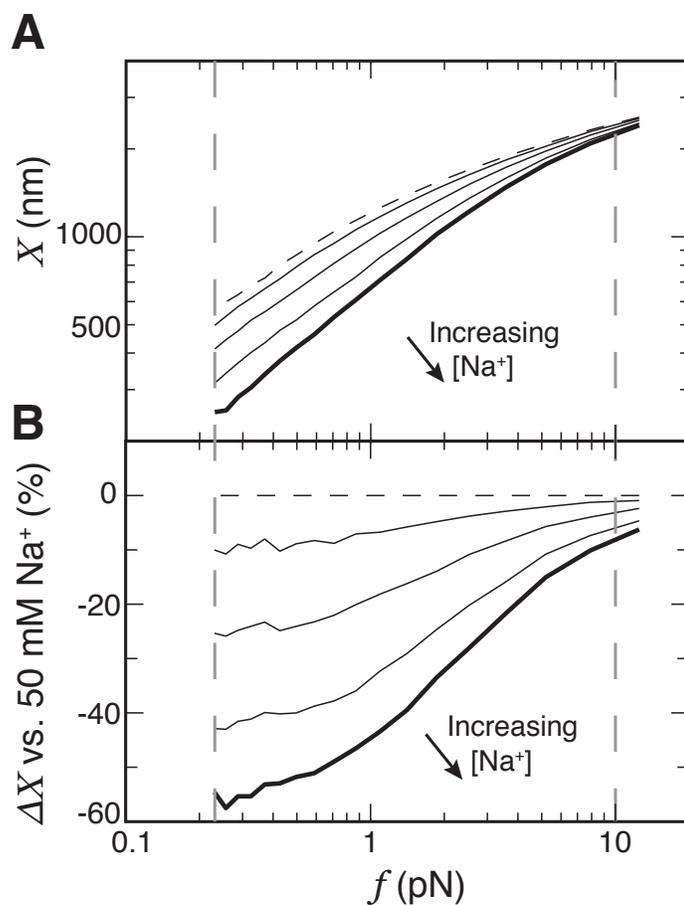


Figure D.3: Ion-induced compaction of ssRNA conformation (data of Chapter 3). (A) Extension vs. force curves at 50 (dashed line), 100, 200, 500 and 1000 (thick line) mM Na⁺. (B) Percentage deviation of each extension vs. force curve from the 50 mM Na⁺ curve. The dashed vertical lines in both plots indicate the forces corresponding to the compact and stretched states in Figures D.2A and D.2C.

stabilizing effect on both states. With the sign convention in use here, the negative slope of the $\Delta\Delta G$ vs. $\ln c_{\text{bulk}}$ plots implies that this stabilization is greater for the compact state. This is consistent with the greater degree of compaction observed at low force than at high force (Figure D.3B) and with the much shorter tensile screening length [32] compared with the Debye screening length at high force, indicating weaker electrostatics-mediated ion interactions with that state.

The inflection of the curves differ: both single-stranded species exhibit negative curvature of $\Delta\Delta G$ with respect to $\ln c_{\text{bulk}}$, whereas the dsDNA curve has positive curvature. Since, following the above arguments, $\Delta\Delta G$ arises mostly from ion interactions with the compact state, this curvature implies increasing marginal stabilization of that state, with increasing $\ln c_{\text{bulk}}$, in ssRNA and ssDNA, but decreasing marginal stabilization in dsDNA. This phenomenon can be understood in terms of the electrostatics-dependent stiffness of the polymers, as in the above explanation of the slopes of the $\Delta\Gamma$ vs. $\ln c_{\text{bulk}}$ plots [75]. Although the flexible, single-stranded polymers are known to exhibit more nuanced, length-scale-dependent conformation [26, 43, 213–216], for this qualitative argument we make the assumption of WLC behavior [8], characterized by a single persistence length. This persistence length can be written, following from the discussion of Section 2.2, as the sum $l_p = l_{p,0} + l_{p,\text{el}}$. Further, we can write $l_{p,\text{el}} \sim c_{\text{bulk}}^{-C}$, where $C = 1$ in the OSF theory [50, 51] or $C = 1/2$ in the theory of Barrat and Joanny [52]. For flexible ssDNA and ssRNA, $l_{p,\text{el}} \gg l_{p,0}$, which implies that $l_{p,\text{ss}} \sim c_{\text{bulk}}^{-C}$; for stiff dsDNA, $l_{p,0} \gg l_{p,\text{el}}$, which implies no concentration dependence of $l_{p,\text{ds}}$. To determine the strength of ion interactions with the compact state, we argue that the important quantity is the number of charged monomers per Debye sphere, g_D . The volume of the Debye sphere, v_D , scales as $v_D \sim c_{\text{bulk}}^{-3/2}$ and the density of monomers can be derived from the radius of gyration of a WLC in the long chain limit [217]. From these, we can obtain the scaling of g_D for

both ssDNA/ssRNA and dsDNA:

$$g_D \sim c_{\text{bulk}}^C \quad (\text{single-stranded}) \quad (\text{D.29})$$

$$g_D \sim c_{\text{bulk}}^{-3/2} \quad (\text{double-stranded}) \quad (\text{D.30})$$

Thus, the ion interactions with the compact conformation are predicted to be an increasing function of bulk concentration for ssDNA and ssRNA, but a decreasing function of concentration for dsDNA, consistent with the results of Figures D.2C and D.2D.

D.4.2 DNA hairpin

Figure D.4A shows the ion excess change associated with the unfolding of a 25-bp DNA hairpin from the studies of Dittmore *et al.* [96] that were discussed in Chapter 6. We apply the the same treatment of Section D.3 to these data to obtain $\Delta\Delta G$ as a function of bulk ion concentration (Figure D.4B). Others have performed similar analyses on a 20-bp RNA hairpin [209] and on RNAs exhibiting tertiary structure [210, 218], in the presence of monovalent and divalent ions.

From these data, we see that the non-monotonicity of the $\Delta\Gamma$ vs. $\ln c_{\text{bulk}}$ plot is manifest in $\Delta\Delta G$ as a changing inflection at $c^* \approx 120$ mM Na^+ . That is to say that at concentrations below c^* the marginal stabilization of the folded state over the unfolded state is increasing with bulk concentration while for concentrations above c^* the marginal stabilization is decreasing. This reverse sigmoidal character is consistent with results from other studies [209, 210, 218]; it is seen most completely in the divalent data of Li [218].

Note that the unfolded state in this hairpin system is essentially the same as the stretched state of the ssDNA system discussed above: both are ssDNA, devoid of secondary structure and stretched by ~ 10 pN applied force. In treating the unstructured nucleic acids, we argued that the contribution of this stretched state to the free energy

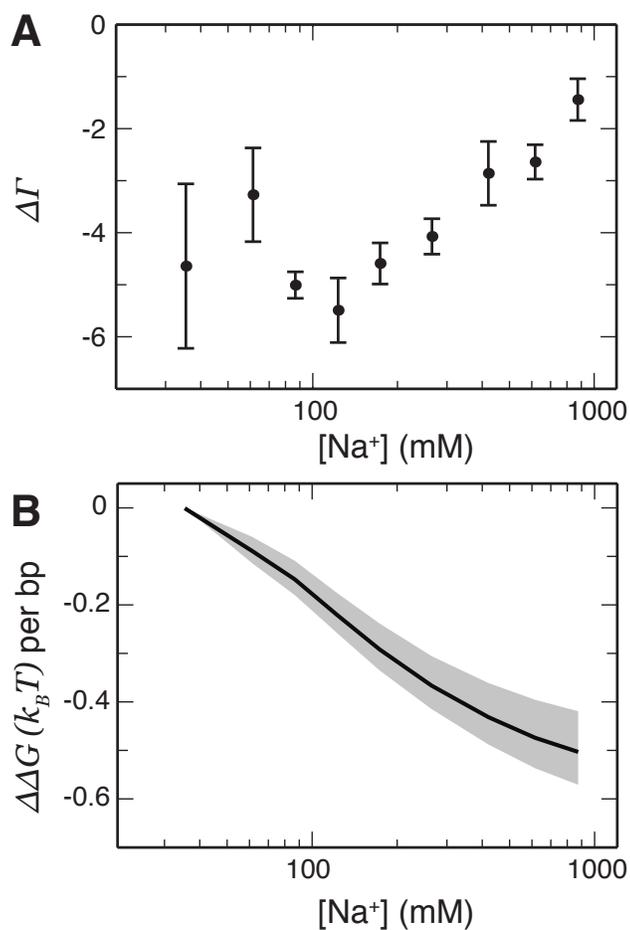


Figure D.4: Application of Equation D.26 to obtain free energy of ion stabilization of folding for hairpins. (A) Change in excess monovalent ions as a 25-bp DNA hairpin is mechanically unfolded by approx. 12–16 pN applied force. Values and uncertainties are combined from the results of the Maxwell relation, free energy gradient and Clausius-Clapeyron relation methods reported by Dittmore *et al.* [96]. (B) Line: $\Delta\Delta G$ curve of 25-bp RNA hairpin referenced to zero at 35 mM Na^+ . Shaded region denotes the propagated uncertainty from $\Delta\Gamma$ data.

of ion stabilization is minimal, since most of the electrostatic interactions are screened by the applied force. Under this assumption we can make similar arguments here: that the conformation of the hairpin's folded state below c^* acts like the ssDNA, exhibiting conformational collapse with increasing $\ln c$, whereas the conformation of the folded state above c^* acts like the dsDNA, exhibiting a conformation mostly unchanged under addition of ions.

Appendix E

Force calibration for molecules that exhibit conformational switching

This chapter is reprinted with permission from Ref. [31]. Copyright 2016, AIP Publishing.

In the hairpin folding/unfolding experiments of Chapter 6, MT force calibration was achieved indirectly by making a force calibration curve using longer PEG chains (see Section A.2.3). This was done because, when using a 60 Hz camera, the characteristic Lorentzian corner frequency of these tethers was too fast to be adequately sampled. However, more advanced cameras are now becoming available that are capable, when coupled with appropriately powerful image processing software, of real-time bead tracking at up to ~ 5000 Hz; such a system could sample the frequency response of these shorter tethers. This raises the question, however, of whether the bead position fluctuations arising from conformational switching between states will contaminate the thermal fluctuations and lead to biased force calibration results.

In this appendix, we analyze the degree to which common force calibration approaches, involving the fitting of model functions to the Allan variance or power spectral

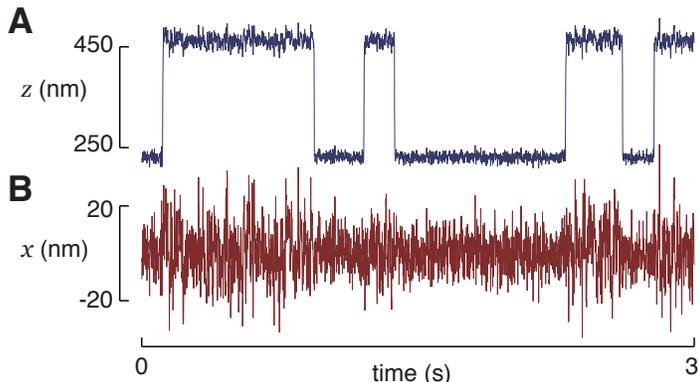


Figure E.1: Conformational switching manifests as (A) a telegraphic pattern in the bead’s vertical displacement, z , and (B) a corresponding modulation of the fluctuations in the bead’s horizontal position, x . (Simulated trajectories)

density of the bead trajectory, are biased by conformational switching. We find significant effects in two limits: that of large molecular extension changes between the two states, in which alternative fitting functions must be used, and that of very fast switching kinetics, in which the force calibration cannot be recovered due to the slow diffusion time of the magnetic bead. We use simulations and high-resolution RNA hairpin data to show that most biophysical experiments do not occur in either of these limits.

Conformational switching manifests as a telegraphic pattern in the time-series of molecular extension, z (Figure E.1A), and introduces two parameters into the system: the change in molecular extension between the two states, ΔX , and the kinetic rate constant for switching between the states at coexistence, k_{rate} . Coexistence occurs at a particular applied force, $f_{1/2}$. Away from coexistence there will be multiple rate constants, one corresponding to the lifetime in each conformational state. Values of $f_{1/2}$, ΔX , and k_{rate} for several simple biomolecular systems that exhibit equilibrium conformational switching are listed in Table E.1.

One of the main tasks in analyzing MT data is to calibrate the mechanical force

Table E.1: Properties of some biological model systems that exhibit cooperative conformational switching under applied force. Note that these values are generally expected to be salt-concentration-dependent.

system	$f_{1/2}$ (pN)	ΔX (nm)	k_{rate} (s^{-1})	ref.
6 bp DNA hairpin	8.1	5.2	389	[219]
30 bp DNA hairpin	14.7	26.5	3.8	[219]
DNA G-quadruplex	2.5	2.7	0.35	[220]
RNase H (I \leftrightarrow U)	5.6	15	6.9	[221]
11 kbp DNA supercoil	4.0	120	13.2	[222]

experienced by the biomolecule. This is typically done via analysis of the Brownian fluctuations of the bead orthogonal to the direction of the pulling force (\hat{x} in Figure 1.2A) by computing the variance [185], Allan variance [186], or power spectral density [186–190] of those fluctuations. In all of these cases, the force is obtained by treating the bead, tethered by the biomolecule, as a damped pendulum subject to Brownian fluctuations [185]; the “spring constant”, k , of the system for fluctuations in the \hat{x} direction is obtained from analysis of the horizontal bead fluctuations and the force, f , is found via

$$f = k \langle z \rangle, \quad (\text{E.1})$$

where $\langle z \rangle$ is the time-average vertical bead position. Associated with these fluctuations is a characteristic corner frequency that varies inversely with extension. Accurate direct force calibrations for short molecules can therefore only be obtained when sampled at a correspondingly high frame rate [186, 187, 189]; e.g., using modern high-speed MT instruments [223–225]. Traditional MT instruments, with 60–100 Hz frame rates, will give significantly biased force calibrations for such molecules. In practice, short molecules are often handled by pre-calibrating the MT using a longer tether such as double-stranded DNA; however, this introduces issues of bead-to-bead variability for which careful corrections must be made [226, 227].

Conformational switching may bias direct force calibration, since the Brownian fluctuations that are analyzed to obtain the force could be contaminated by additional fluctuations arising from the stochastic switching. These additional fluctuations are seen in Figure E.1B, where conformational switching—clearly monitored in the z trajectory (Figure E.1A)—leads to a non-constant amplitude of the x fluctuations. In Section E.1 below, we use theoretical arguments to show that conformational switching-induced bias can arise via two different mechanisms. Then, in Section E.2, we use simulations to verify these effects and to identify the portion of the experimental parameter space in which they are important.

E.1 Theory

E.1.1 Effect of conformational switching on measures of bead fluctuations

MT force calibration can be accomplished by measuring the Brownian fluctuations of the bead/tether system. The fluctuations can be quantified by a single value, the variance [185], or by computing quantities that more completely use the available data: the Allan variance (AV) or power spectral density (PSD) [186]. The conceptually simple variance-based method assigns an average energy to each degree of freedom of the system using the equipartition theorem. The expression for the calibrated force, f_{cal} , of a single-state system is: [185]

$$f_{\text{cal}} = \frac{\langle z \rangle k_B T}{\langle \delta_x^2 \rangle}. \quad (\text{E.2})$$

In the presence of conformational switching, each state has a different extension $\langle z_i \rangle$ and variance $\langle \delta_{x,i}^2 \rangle$. Assuming that the equilibration kinetics are sufficiently fast (see Section E.1.2), Equation E.2 is shown also to hold for the switching molecule under the

substitutions $\langle z \rangle \rightarrow \overline{\langle z_i \rangle}$ and $\langle \delta_x^2 \rangle \rightarrow \overline{\langle \delta_{x,i}^2 \rangle}$ (bar denotes an average over conformational states weighted by the time spent in each state).

A more robust calibration scheme utilizes the AV or PSD (in terms of a time difference or frequency, respectively), since variance-based calibration is particularly sensitive to noise. The AV and PSD are fit by functions having two free parameters: k and α , as discussed in Section A.2.3. For example, the model function for AV fitting was given as Equation A.2; a similar function exists for the PSD. Because the PSD and AV contain the same force calibration information, we will confine our discussion in this appendix to the AV; however, our findings hold for the PSD as well.

Unlike in the simple case of Equation E.2, we find that the AV and PSD fitting functions are formally modified by conformational switching. This occurs because the power of the fluctuations at a particular frequency, corresponding to a particular τ , is the time-weighted average of the power in that channel across states. Therefore, the overall AV is the time-weighted average of the AVs of the constituent states (indexed by i) plus the power from the transitions between the states:

$$\sigma^2(\tau) = \overline{\sigma_i^2(\tau)} + \text{contribution from switching.} \quad (\text{E.3})$$

So, to fit the AV of an N -state system, it is necessary to fit the average of N copies of Equation A.2—each with a different value of k and weighted by the fraction of the total time spent in that state—and to include an additional term from the switching signal. Note that this switching refers to the subtle changes in the fluctuations in x (Figure E.1B), not the telegraphic pattern in z (Figure E.1A). The simulations of Section E.2 suggest that the switching contribution can be neglected under experimentally relevant conditions. If the extensions and occupancy times for each state are known, the multi-state Equation E.3 can be fit to the data without increasing the number of fitting

parameters.

E.1.2 Effect of two-state switching kinetics

Our analysis thus far has neglected any effects that may arise from the kinetics of the system, either from the rate of switching between the conformational states or the rate at which the bead diffuses through the solution. In Section E.1.1, we assumed that the fluctuations within each state were independent. However, if the lifetime within a particular state is too short, this assumption can break down. As an example, suppose that the system begins in a compact state and then transitions to an extended state: the bead must diffuse to larger distances from the origin in order to probe the larger fluctuations expected. If the system then transitions back to the compact state before this diffusion can take place, due to a sufficiently short lifetime of the extended state, these larger fluctuations will not be adequately sampled.

We can estimate when this diffusion limitation will occur by considering the interplay between the diffusion constant, $D = k_B T / \alpha$, and the effective spring constant, $k = \langle z \rangle / f$. The spring constant sets the length scale of the fluctuations, δ_x , through Equation E.2. A simple model of diffusion can be used to obtain the expected displacement, Δx , in a particular time, Δt , given the diffusion constant: $\Delta x = \sqrt{2D\Delta t}$. This Δt is related to the kinetic rate constant for switching between the two states (assumed to be the same for both states; i.e., coexistence): $\Delta t = 1/k_{\text{rate}}$. We can expect to obtain biased force calibrations if the bead cannot diffuse far enough to sample the expected amplitude of the fluctuations in the time set by k_{rate} . Thus we can identify a critical rate constant, k_{crit} , above which the force calibration will become inaccurate:

$$k_{\text{crit}} = \frac{2f}{\langle z \rangle \alpha}. \quad (\text{E.4})$$

This result is a function of f and $\langle z \rangle$, which are properties of the biomolecule under study (see values of $f_{1/2}$ and ΔX in Table E.1), and of α , which depends on the magnetic bead size and solution viscosity. Equation E.4 assumes that the value of $\langle z \rangle$ in the extended state is much larger than that in the compact state.

E.2 Simulations

We simulated two-state bead trajectories using a Brownian dynamics approach based upon the over-damped Langevin equation governing this system:

$$\alpha \dot{x} + kx = f_L, \quad (\text{E.5})$$

where f_L is the stochastic Langevin force. The simulations were carried out iteratively at 300 kHz and then down-sampled, by averaging over bins, to $f_s = 3$ kHz to approximate the effects of video tracking with a finite frame rate. In all of our simulations, we set k based on Equation E.1—with $f = 12$ pN and $\langle z \rangle$ depending on the state occupied—and set $\alpha = 1.5 \times 10^{-5}$ pN s nm⁻¹: reasonable values for an MT experiment using micron-diameter magnetic beads. Transitions between states are modeled as stochastic events with k_{rate} and actualized by switching the spring constant between its two possible values. We model both the folding and unfolding transitions as having the same kinetic rate. This is the condition under which the effect of conformational switching is most pronounced: for unequal values of k_{rate} , the system tends toward one-state behavior. An example MATLAB script for carrying out these simulations and the fitting required for force calibration, and for computing estimated force calibration biases, may be freely downloaded at <http://www.engr.ucsb.edu/~saleh/#Code>.

E.2.1 Validation of predicted effects

To verify that the AV of a system with conformational switching behaves as the time-weighted average of the AVs of the individual states, we simulated a two-state system with extensions of 300 and 3000 nm and with $k_{\text{rate}} = 5 \text{ s}^{-1}$. The computed AV from this simulation is shown in Figure E.2A, with error bars representing the 68% confidence interval (akin to the one-sigma level) derived from the gamma distribution of the AV [186]. Also plotted are the predicted curves from modeling either a one-state system with an effective, averaged spring constant (Equation A.2) or a two-state system in which the two AVs have been averaged (Equation E.3). The models are evaluated, without fitting, using the known values of the state extensions and of f and α given above. Comparing the two models, we see that the two-state model is in good agreement with the data, whereas the one-state model disagrees across a broad domain of τ .

To confirm the switching rate-dependent effect of bead diffusion, we performed simulations over $k_{\text{rate}} = \{10, 100, 1000\} \text{ s}^{-1}$, a range encompassing the predicted critical rate for the parameters of this system: $k_{\text{crit}} = 533 \text{ s}^{-1}$ (from Equation E.4). The AV obtained from these simulations is plotted in Figure E.2B. There is little variation between the two curves below k_{crit} ; however, above k_{crit} , the maximum in the AV shifts to lower τ , commensurate with a force calibration biased towards higher forces.

E.2.2 Force calibration bias

To quantify how conformational switching biases the force calibrations obtained in an experiment, we performed simulations over a range of k_{rate} (10–1000 s^{-1}) and ΔX (0–1000 nm). Mirroring experimental practice, we treated the conformationally switching molecule as being attached to a handle of fixed length; here $\langle z \rangle_{\text{handle}} = 300 \text{ nm}$. We analyzed the resulting traces using both the effective one-state (Equation A.2) and two-state

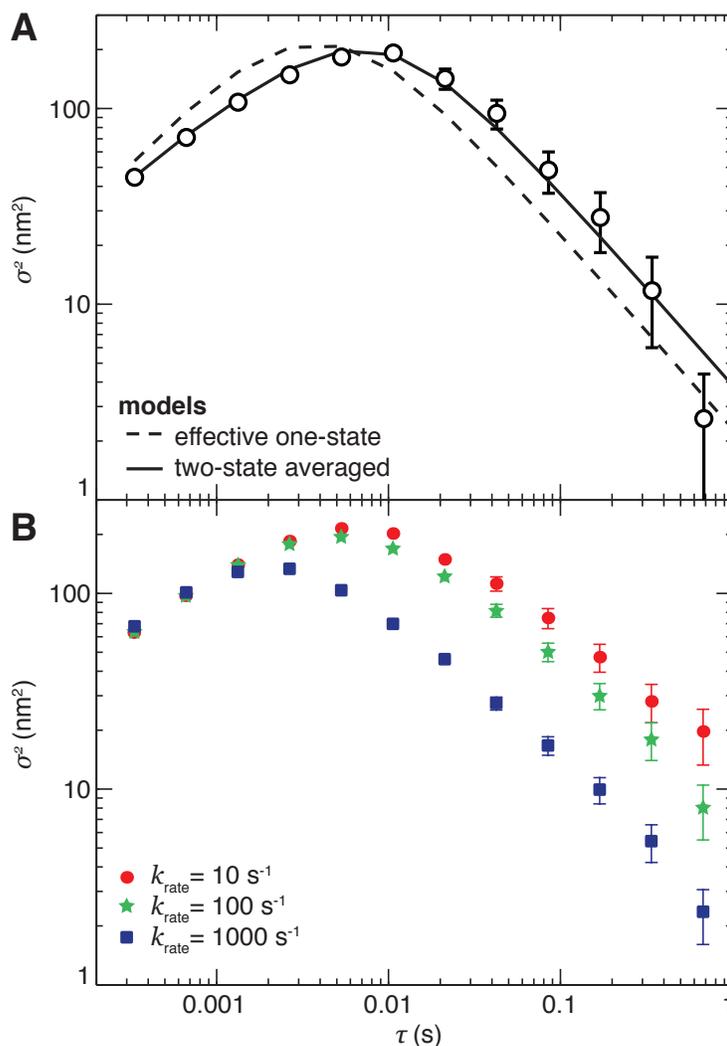


Figure E.2: (A) AV of simulated two state system with extensions 300 and 3000 nm and $k_{\text{rate}} = 5 \text{ s}^{-1}$ (circles) is in poor agreement with the single-state model (dashed line, Equation A.2), having an effective spring constant set based on the average extension of the system, but is in good agreement with the two-state model (solid line, Equation E.3) in which the AV is averaged between the states. (B) AV of the same simulated system, now with k_{rate} equal to 10 (red points), 100 (green stars) and 1000 (blue squares) s^{-1} . Distortion of the AV is seen when $k_{\text{rate}} > k_{\text{crit}}$. Error bars are the 68% confidence interval from computing the AV.

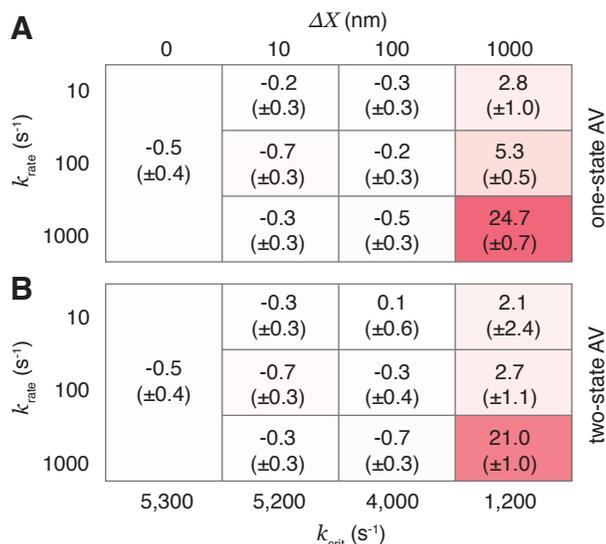


Figure E.3: Bias in force calibration based on simulated traces, analyzed using (A) the effective one-state AV model or (B) the two-state averaged AV model, each fit with two free parameters. Significant bias, due to the bead diffusion effect, is seen only when $k_{\text{rate}} \gtrsim k_{\text{crit}}$ (from Equation E.4). Reported values are percentage deviations from the expected $f = 12$ pN; uncertainties are standard errors of the mean of these values.

averaged (Equation E.3) AV fitting models. To fit the AV, we use a previously described maximum likelihood approach [186]. Both models are fit with two free parameters (k and α), since, for the two-state model, we take as known the fraction of time spent in each state; in experimental practice, these data can be obtained from telegraphic z -traces (Figure E.1A). At each pair of k_{rate} and ΔX , ten simulations, each 10 s long, were carried out. The results of this analysis are shown in Figure E.3. The bias values reported are the percentage deviations of the obtained forces from the expected force, $f = 12$ pN; the uncertainties reported are the standard errors of the mean of these biases across the ten trials.

Both methods give an acceptably unbiased calibration of the force over most of the parameter space. There is no discernible difference in the results from the single-state and two-state models up to $\Delta X = 1000$ nm at small k_{rate} . Large biases are seen in the

data with $\Delta X = 1000$ nm and $k_{\text{rate}} = 1000$ s⁻¹ for both the single- and two-state AV analyses, as expected based on the bead diffusion effect (compare k_{rate} with k_{crit} in Figure E.3). This effect is confirmed in simulations with smaller values of ΔX and f , but having the same value of k_{crit} , in which appreciable bias is still seen. The bias does decrease for smaller values of ΔX , presumably due to the breakdown of Equation E.4 when one state is no longer much more extended than the other.

While the results of these simulations are given in Figure E.3 in terms of ΔX , an accurate interpretation of the underlying effects must also account for the handle length, $\langle z \rangle_{\text{handle}}$. Because the bead diffusion effect (Section E.1.2) depends on the time it takes the bead to diffuse to the limits of its range of fluctuations in the extended state, its magnitude will depend predominately on ΔX , provided, as mentioned above, that $\Delta X \gg \langle z \rangle_{\text{handle}}$. In contrast, the non-kinetic averaging effect requiring the use of the two-state AV model (Section E.1.1) depends on there being a significant difference in the AV of the two individual states, and will thus depend predominately on $\Delta X / \langle z \rangle_{\text{handle}}$.

E.3 Sample experiment

To assess the practical experimental implications of conformational switching, we monitored the folding and unfolding of a 14 base pair RNA hairpin using a high-speed MT instrument ($f_s = 11$ kHz) [223] and then analyzed the resulting traces to obtain force calibrations. We used a molecular construct, akin to that of Ref. [38], consisting of a hairpin separated from the flow cell surface by an 828 base pair DNA spacer and from the magnetic bead by a 20 nucleotide length of single-stranded DNA. At a given bulk salt concentration, an appropriate force is applied to cause the system to enter folding/unfolding equilibrium between the two states, as in Figure E.1B.

Figure E.4A shows the AV computed from a representative trace near $f_{1/2}$ and a fit of

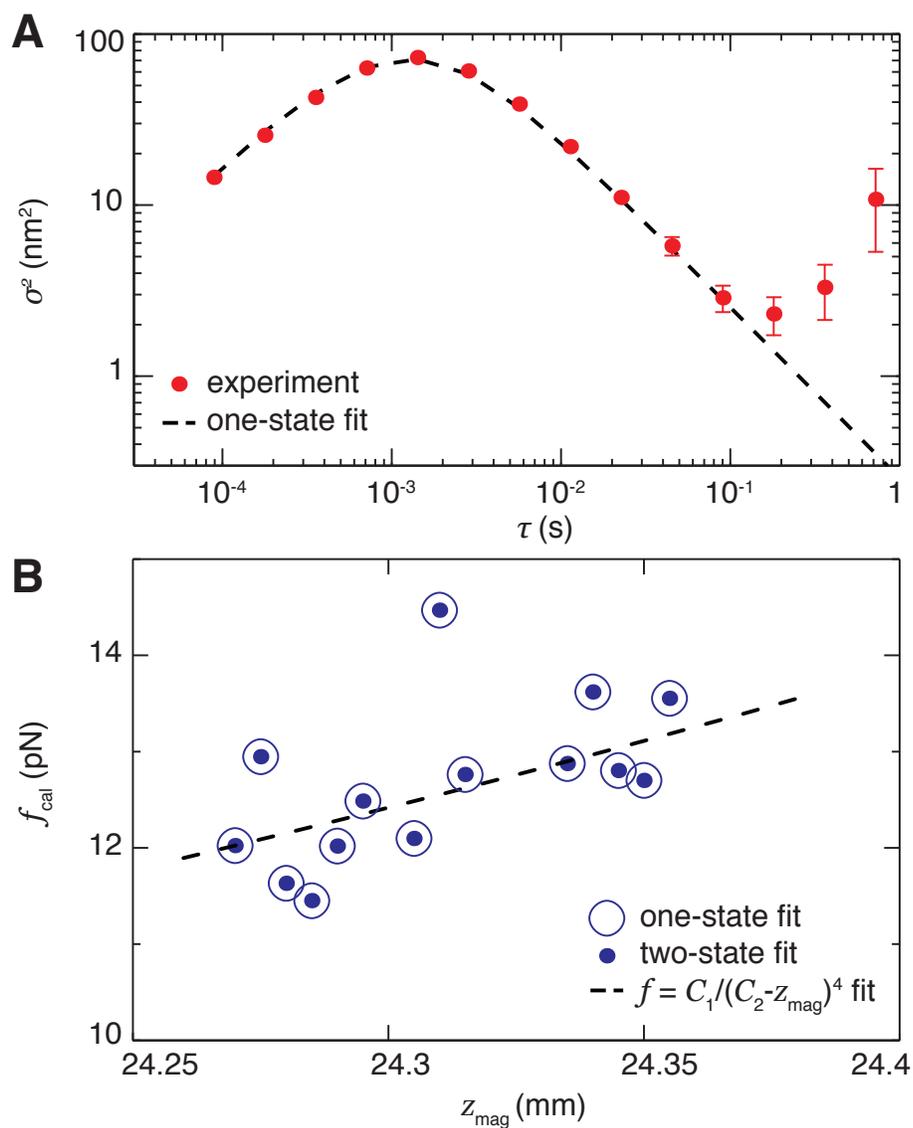


Figure E.4: Experimental force calibration of a 14 base pair RNA hairpin exhibiting two-state switching behavior is dominated by sources of error other than conformational switching. (A) AV from a representative trace (points), fit by the effective single-state model (dashed line). (B) Force calibration results for the one-state (circles) and two-state (points) methods, at various forces and salt concentrations, co-plotted with the best fit of a model function (dashed line).

the effective single-state model (Equation A.2). Because the AV and the fitting function are in good agreement, we conclude that the two-state correction is not required for such small values of ΔX . This agrees with our simulation results (Figure E.3), which predict force bias values of $< 1\%$ and standard deviations of $\sim 1\%$ for both one- and two-state fitting.

Force calibrations using the one- and two-state AV models are shown in Figure E.4B for a representative molecule over a range of force and bulk salt concentrations. The calibrated forces, as a function of the position of the magnets, z_{mag} , relative to a reference position are fit by a function that approximates the force calibration behavior: $f = C_1/(C_2 - z_{\text{mag}})^4$. The standard deviation of the calibrated forces about this curve is 5.6%, larger than the 1% deviation expected from the simulations. Because the single-state and multi-state force calibrations are in such good agreement, and because the model functions fit the data so well, we conclude that this scatter arises from sources other than the determination of k .

E.4 Conclusion

We have carried out theoretical and simulation studies to determine whether conformational switching biases the calibration of force spectroscopy experiments. We have shown that perturbations to the expected behavior can occur due to the way the AV/PSD of the individual states must be combined (by averaging) and due to the competition between bead diffusion and conformational switching kinetics; the former effect can be corrected by using the appropriately averaged fitting function. We see from simulated (Figure E.3) and actual (Figure E.4) experiments that, in practice, conformational switching is not a dominant source of noise over the parameter space of contemporary biophysical studies, where ΔX values tend to be less than 100 nm (Table E.1) and $\Delta X/\langle z \rangle_{\text{handle}} \ll 1$.

In these cases, error in force calibration is dominated by other sources. Biases arising from conformational switching will become more pronounced in systems with faster switching kinetics, in solutions with higher viscosity, and in experiments employing larger beads or shorter biomolecular handles.

Appendix F

Notation used

F.1 Abbreviations used

AES	atomic emission spectroscopy
AFM	atomic force microscopy
ASAXS	anomalous small-angle x-ray scattering
AV	Allan variance
D-AES	dialysis and atomic emission spectroscopy
dGPVG	de Gennes-Pincus-Velasco-Brochard
DNA	deoxyribonucleic acid
dsDNA	double-stranded DNA
ES	electrostatics
EV	excluded volume
FJC	freely jointed chain
HEG	hexaethylene glycol
MC	Monte Carlo
MD	molecular dynamics

MN	Manning-Netz
MT	magnetic tweezers
NA	nucleic acid
OSF	Odijk-Skolnick-Fixman
OT	optical trapping
PB	Poisson-Boltzmann
PEG	polyethylene glycol
poly(U)	polyuridine
PSD	power spectral density
RNA	ribonucleic acid
SLC	snakelike chain
SMFS	single-molecule force spectroscopy
ssNA	single-stranded nucleic acid
TBI	tightly bound ion
TCEP	tris(2-carboxyethyl)phosphine
UMF	uniform mean field
WLC	wormlike chain

F.2 Mathematical symbols used

a	effective chemical bond length
A	ion activity correction factor
b	charge spacing
c	molar concentration
c_{bulk}	bulk salt concentration

c_{Θ}	solution theta concentration
c_{\pm}	cation/anion concentration
c^*	$\Delta\Delta G$ inflection concentration
C	free parameter
D	diffusion coefficient
e	elementary (electron) charge
f	force
f_{app}	applied force
f_c	inter-regime crossover force
f_{corner}	Lorentzian corner frequency
f_{cal}	calibrated force
f_{el}	electrostatic tension
f_L	Langevin force
f_o	reference force
f_s	sampling frequency
$f_{1/2}$	equilibrium unfolding force
F	potential of mean force
\mathcal{F}	generic thermodynamic potential
g_D	monomers per Debye sphere
h_{mag}	magnet height above flow cell
I	ionic strength
k	effective spring constant
k_B	Boltzmann's constant
k_{crit}	critical rate constant
k_{rate}	kinetic rate constant of two-state switching

K	equilibrium constant
\mathcal{K}_i	modified Bessel function of the second kind
l	Kuhn length
l_B	Bjerrum length
l_p	persistence length
$l_{p,el}$	electrostatic persistence length
$l_{p,0}$	bare persistence length
L_c	contour length
m	molal concentration
N	degree of polymerization
\mathcal{N}	number of molecules
\mathcal{N}_e	number of charges
q	scattering wave vector
Q	electrostatic charge
r	position in space
R	radius
R_F	Flory radius
R_g	radius of gyration
R_0	ideal gas constant
S	stretch modulus
T	temperature
T_m	melting temperature
v	monomer excluded volume
v_D	volume of Debye sphere
V	volume of system

V_{solvent}	solvent-accessible volume
W	work
x	displacement of bead orthogonal to pulling force
\hat{x}	direction orthogonal to pulling force
X	molecular extension
X_c	inter-regime crossover extension
X_{ne}	non-equilibrium extension from force ramp study
z	displacement of bead parallel to pulling force
\hat{z}	direction parallel to pulling force
z_{handle}	molecular handle length
z_{mag}	magnet position
z_{max}	magnet position of closest approach
Z	valence
\mathcal{Z}	grand-canonical partition function
α	dissipation due to viscous drag
β	empirical thermal denaturation constant
γ	ion activity coefficient
Γ	preferential ion interaction coefficient
$\bar{\Gamma}$	per-nucleotide ion excess
Γ_{\pm}	preferential cation/anion interaction coefficient
Γ_{ds}	ion excess of helical dsNA
Γ_{ES}	ion excess due to electrostatics
Γ_{EV}	ion excess due to excluded volume
Γ_{ss}	ion excess of random-coil ssNA
Γ_{stretch}	ion excess of stretched ssNA

δ	bead rotation correction distance
δ_x^2	variance of bead position
ΔG_{ligand}	free energy change of ligand binding
ΔH°	change in standard enthalpy of formation
Δt	time interval
ΔX	extension change between states
$\Delta \Gamma$	difference in ion excess between states
$\Delta \Delta G$	differential free energy of conformation-associated ligand binding
ϵ	dielectric constant
ϵ_0	permittivity of free space
ζ	Manning parameter
η	dimensionless relative charge density
κ^{-1}	Debye screening length
μ	chemical potential
ν	Flory exponent
ξ	tensile screening length
ξ_c	inter-regime crossover length scale
ξ_t	thermal blob size
ρ	charge density
σ^2	Allan variance
τ	time separation between samples
Φ	electrostatic potential
Φ_D	Donnan potential
Ω	grand-canonical (Landau) potential

Bibliography

- [1] Gennes, P. G. d. (1979) Scaling concepts in polymer physics, Cornell University Press, Ithaca, N.Y.
- [2] Saleh, O. A. (2015) Perspective: Single polymer mechanics across the force regimes. *J. Chem. Phys.*, **142**(19), 194902.
- [3] Rubinstein, M. and Colby, R. H. (2003) Polymer physics, Oxford University Press, New York.
- [4] Staudinger, H. (1920) Über polymerisation. *Ber. Dtsch. Chem. Ges.*, **53**(6), 1073–1085.
- [5] Staudinger, H. (1933) Viscosity investigations for the examination of the constitution of natural products of high molecular weight and of rubber and cellulose. *T. Faraday Soc.*, **29**(140), 18–32.
- [6] Kuhn, W. (1934) Über die gestalt fadenförmiger moleküle in lösungen. *Kolloid Z.*, **68**(1), 2–15.
- [7] Flory, P. J. (1949) The configuration of real polymer chains. *J. Chem. Phys.*, **17**(3), 303–310.
- [8] Kratky, O. and Porod, G. (1949) Röntgenuntersuchung gelöster fadenmoleküle. *Recl. Trav. Chim. Pay.-B.*, **68**(12), 1106–1122.
- [9] McIntosh, D., Ribeck, N., and Saleh, O. (2009) Detailed scaling analysis of low-force polyelectrolyte elasticity. *Phys. Rev. E*, **80**(4).
- [10] Rubinstein, M. and Papoian, G. A. (2012) Polyelectrolytes in biology and soft matter. *Soft Matter*, **8**(36), 9265–9267.
- [11] Bloomfield, V. A. (1996) DNA condensation. *Curr. Opin. Struc. Biol.*, **6**(3), 334–341.
- [12] Crick, F. (1970) Central dogma of molecular biology. *Nature*, **227**(5258), 561–563.

- [13] International Human Genome Sequencing Consortium (2004) Finishing the euchromatic sequence of the human genome. *Nature*, **431**(7011), 931–945.
- [14] Luger, K., Mäder, A. W., Richmond, R. K., Sargent, D. F., and Richmond, T. J. (1997) Crystal structure of the nucleosome core particle at 2.8 Å resolution. *Nature*, **389**(6648), 251–260.
- [15] Wingender, E., Dietze, P., Karas, H., and Knüppel, R. (1996) TRANSFAC: A database on transcription factors and their DNA binding sites. *Nucleic Acids Res.*, **24**(1), 238–241.
- [16] Roe, J.-H., Burgess, R. R., and Record, M. T. (1984) Kinetics and mechanism of the interaction of Escherichia coli RNA polymerase with the λ P R promoter. *J. Mol. Biol.*, **176**(4), 495–522.
- [17] Kozlov, A. G. and Lohman, T. M. (1998) Calorimetric studies of E. coli SSB protein-single-stranded DNA interactions. Effects of monovalent salts on binding enthalpy. *J. Mol. Biol.*, **278**(5), 999–1014.
- [18] Doherty, E. A. and Doudna, J. A. (2001) Ribozyme structures and mechanisms. *Annu. Rev. Biophys. Bio.*, **30**(1), 457–475.
- [19] Garst, A. D., Edwards, A. L., and Batey, R. T. (2011) Riboswitches: Structures and mechanisms. *Cold Spring Harbor Perspect. Biol.*, **3**(6), a003533.
- [20] Grønbech-Jensen, N., Mashl, R. J., Bruinsma, R. F., and Gelbart, W. M. (1997) Counterion-induced attraction between rigid polyelectrolytes. *Phys. Rev. Lett.*, **78**(12), 2477.
- [21] Manning, G. S. (1969) Limiting laws and counterion condensation in polyelectrolyte solutions I. Colligative properties. *J. Chem. Phys.*, **51**(3), 924–933.
- [22] Zimm, B. H. and Bret, M. L. (1983) Counter-ion condensation and system dimensionality. *J. Biomol. Struct. Dyn.*, **1**(2), 461–471.
- [23] Salamone, J., (ed.) *Polymeric Materials Encyclopedia* pp. 649–650 CRC Press (1996).
- [24] Savelyev, A. (2012) Do monovalent mobile ions affect DNA’s flexibility at high salt content?. *Phys. Chem. Chem. Phys.*, **14**(7), 2250–2254.
- [25] McIntosh, D. and Saleh, O. (2011) Salt species-dependent electrostatic effects on ssDNA elasticity. *Macromolecules*, **44**(7), 2328–2333.
- [26] Jacobson, D. R., McIntosh, D. B., and Saleh, O. A. (2013) The snakelike chain character of unstructured RNA. *Biophys. J.*, **105**(11), 2569–2576.

- [27] Abels, J., Moreno-Herrero, F., Van der Heijden, T., Dekker, C., and Dekker, N. (2005) Single-molecule measurements of the persistence length of double-stranded RNA. *Biophys. J.*, **88**(4), 2737–2744.
- [28] Marko, J. and Siggia, E. (1995) Stretching DNA. *Macromolecules*, **28**(26), 8759–8770.
- [29] Neuman, K., Lionnet, T., and Allemand, J.-F. (2007) Single-Molecule Micromanipulation Techniques. *Annu. Rev. Mater. Res.*, **37**, 33–67.
- [30] Jacobson, D. R. and Saleh, O. A. (2015) Measuring the differential stoichiometry and energetics of ligand binding to macromolecules by single-molecule force spectroscopy: An extended theory. *J. Phys. Chem. B*, **119**(5), 1930–1938.
- [31] Jacobson, D. R. and Saleh, O. A. (2016) Magnetic tweezers force calibration for molecules that exhibit conformational switching. *Rev. Sci. Instrum.*, **87**, 094302.
- [32] Pincus, P. (1976) Excluded Volume Effects and Stretched Polymer Chains. *Macromolecules*, **9**(3), 386–388.
- [33] Donnan, F. G. (1911) Theorie der membrangleichgewichte und membranpotentiale bei vorhandensein von nicht dialysierenden elektrolyten. Ein beitrag zur physikalisch-chemischen physiologie.. *Z. Elektrochem. Angew. P.*, **17**(14), 572–581.
- [34] Plum, G. E. and Bloomfield, V. A. (1988) Equilibrium dialysis study of binding of hexammine cobalt (III) to DNA. *Biopolymers*, **27**(6), 1045–1051.
- [35] Bai, Y., Greenfeld, M., Travers, K. J., Chu, V. B., Lipfert, J., Doniach, S., and Herschlag, D. (2007) Quantitative and comprehensive decomposition of the ion atmosphere around nucleic acids. *J. Am. Chem. Soc.*, **129**(48), 14981–14988.
- [36] Greenfeld, M. and Herschlag, D. (2009) Probing nucleic acid–ion interactions with buffer exchange-atomic emission spectroscopy. *Method. Enzymol.*, **469**, 375–389.
- [37] Gebala, M., Giambaşu, G. M., Lipfert, J., Bisaria, N., Bonilla, S., Li, G., York, D. M., and Herschlag, D. (2015) Cation–anion interactions within the nucleic acid ion atmosphere revealed by ion counting. *J. Am. Chem. Soc.*, **137**(46), 14705–14715.
- [38] Jacobson, D. R. and Saleh, O. A. (2016) Quantifying the ion atmosphere of unfolded, single-stranded nucleic acids using equilibrium dialysis and single-molecule methods. *Nucleic Acids Res.*, **44**(8), 3763–3771.
- [39] Baumann, C. G., Smith, S. B., Bloomfield, V. A., and Bustamante, C. (1997) Ionic effects on the elasticity of single DNA molecules. *Proc. Natl. Acad. Sci. U.S.A.*, **94**(12), 6185–6190.

- [40] Hugel, T., Rief, M., Seitz, M., Gaub, H. E., and Netz, R. R. (2005) Highly stretched single polymers: Atomic-force-microscope experiments versus ab-initio theory. *Phys. Rev. Lett.*, **94**(4), 048301.
- [41] Smith, S. B., Finzi, L., and Bustamante, C. (1992) Direct mechanical measurements of the elasticity of single DNA molecules by using magnetic beads. *Science*, **258**(5085), 1122–1126.
- [42] Seol, Y., Skinner, G., and Visscher, K. (2004) Elastic properties of a single-stranded charged homopolymeric ribonucleotide. *Phys. Rev. Lett.*, **93**(11), 118102.
- [43] Saleh, O., McIntosh, D., Pincus, P., and Ribeck, N. (2009) Nonlinear low-force elasticity of single-stranded DNA molecules. *Phys. Rev. Lett.*, **102**(6), 068301.
- [44] Dobrynin, A. V., Carrillo, J.-M. Y., and Rubinstein, M. (2010) Chains are more flexible under tension. *Macromolecules*, **43**(21), 9181–9190.
- [45] Toan, N. M. and Thirumalai, D. (2010) Theory of biopolymer stretching at high forces. *Macromolecules*, **43**(9), 4394–4400.
- [46] Debye, P. and Hückel, E. (1923) Zur theorie der elektrolyte I: Gefrierpunktserniedrigung und verwandte erscheinungent. *Phys. Z.*, **24**(9), 185–206.
- [47] de Gennes, P., Pincus, P., Velasco, R., and Brochard, F. (1976) Remarks on polyelectrolyte conformation. *J. Phys.-Paris*, **37**(12), 1461–1473.
- [48] Netz, R. (2001) Strongly stretched semiflexible extensible polyelectrolytes and DNA. *Macromolecules*, **34**(21), 7522–7529.
- [49] Manning, G. S. (2006) The persistence length of DNA is reached from the persistence length of its null isomer through an internal electrostatic stretching force. *Biophys. J.*, **91**(10), 3607–3616.
- [50] Odijk, T. (1977) Polyelectrolytes Near the Rod Limit. *J Polym. Sci. Pol. Phys.*, **15**(3), 477–483.
- [51] Skolnick, J. and Fixman, M. (1977) Electrostatic Persistence Length of a Wormlike Polyelectrolyte. *Macromolecules*, **10**(5), 944–948.
- [52] Barrat, J.-L. and Joanny, J.-F. (1993) Persistence Length of Polyelectrolyte Chains. *EPL-Europhys. Lett.*, **24**(5), 333.
- [53] Toan, N. M. and Thirumalai, D. (2012) On the origin of the unusual behavior in the stretching of single-stranded DNA. *J. Chem. Phys.*, **136**(23), 235103.
- [54] Dittmore, A., McIntosh, D., Halliday, S., and Saleh, O. (2011) Single-Molecule Elasticity Measurements of the Onset of Excluded Volume in Poly(Ethylene Glycol). *Phys. Rev. Lett.*, **107**(14), 148301.

- [55] Smith, S. B., Cui, Y., and Bustamante, C. (1996) Overstretching B-DNA: The elastic response of individual double-stranded and single-stranded DNA molecules. *Science*, **271**(5250), 795–799.
- [56] Wenner, J. R., Williams, M. C., Rouzina, I., and Bloomfield, V. A. (2002) Salt dependence of the elasticity and overstretching transition of single DNA molecules. *Biophys. J.*, **82**(6), 3160–3169.
- [57] Williams, M. C., Wenner, J. R., Rouzina, I., and Bloomfield, V. A. (2001) Effect of pH on the overstretching transition of double-stranded DNA: Evidence of force-induced DNA melting. *Biophys. J.*, **80**(2), 874–881.
- [58] van Mameren, J., Gross, P., Farge, G., Hooijman, P., Modesti, M., Falkenberg, M., Wuite, G. J., and Peterman, E. J. (2009) Unraveling the structure of DNA during overstretching by using multicolor, single-molecule fluorescence imaging. *Proc. Natl. Acad. Sci. U.S.A.*, **106**(43), 18231–18236.
- [59] Storm, C. and Nelson, P. (2003) Theory of high-force DNA stretching and overstretching. *Phys. Rev. E*, **67**(5), 051906.
- [60] Marko, J. F. (1998) DNA under high tension: overstretching, undertwisting, and relaxation dynamics. *Phys. Rev. E*, **57**(2), 2134.
- [61] King, G. A., Gross, P., Bockelmann, U., Modesti, M., Wuite, G. J., and Peterman, E. J. (2013) Revealing the competition between peeled ssDNA, melting bubbles, and S-DNA during DNA overstretching using fluorescence microscopy. *Proc. Natl. Acad. Sci. U.S.A.*, **110**(10), 3859–3864.
- [62] Paik, D. H. and Perkins, T. T. (2011) Overstretching DNA at 65 pN does not require peeling from free ends or nicks. *J. Am. Chem. Soc.*, **133**(10), 3219–3221.
- [63] Rief, M., Clausen-Schaumann, H., and Gaub, H. E. (1999) Sequence-dependent mechanics of single DNA molecules. *Nat. Struct. Mol. Biol.*, **6**(4), 346–349.
- [64] Dessinges, M.-N., Maier, B., Zhang, Y., Peliti, M., Bensimon, D., and Croquette, V. (2002) Stretching single stranded DNA, a model polyelectrolyte. *Phys. Rev. Lett.*, **89**(24), 248102.
- [65] Stevens, M. J., McIntosh, D. B., and Saleh, O. A. (2012) Simulations of stretching a strong, flexible polyelectrolyte. *Macromolecules*, **45**(14), 5757–5765.
- [66] Stevens, M. and Saleh, O. (2016) Simulations of stretching a flexible polyelectrolyte with varying charge separation. *Eur. Phys. J.-Spec. Top.*, pp. doi:10.1140/epjst/e2016-60113-0.

- [67] Ke, C., Humeniuk, M., Hanna, S., Marszalek, P. E., et al. (2007) Direct measurements of base stacking interactions in DNA by single-molecule atomic-force spectroscopy. *Phys. Rev. Lett.*, **99**(1), 018302.
- [68] McIntosh, D. B., Duggan, G., Gouil, Q., and Saleh, O. A. (2014) Sequence-dependent elasticity and electrostatics of single-stranded DNA: Signatures of base-stacking. *Biophys. J.*, **106**(3), 659–666.
- [69] Seol, Y., Skinner, G. M., Visscher, K., Buhot, A., and Halperin, A. (2007) Stretching of homopolymeric RNA reveals single-stranded helices and base-stacking. *Phys. Rev. Lett.*, **98**(15), 158103.
- [70] Draper, D. (1992) The RNA-folding problem. *Accounts Chem. Res.*, **25**(4), 201–207.
- [71] Draper, D. E. (2004) A guide to ions and RNA structure. *RNA*, **10**(3), 335–343.
- [72] Draper, D. E. (2008) RNA folding: Thermodynamic and molecular descriptions of the roles of ions. *Biophys. J.*, **95**(12), 5489–5495.
- [73] Saenger, W. (1984) Principles of Nucleic Acid Structure, Springer-Verlag, New York.
- [74] Otaka, H., Ishikawa, H., Morita, T., and Aiba, H. (2011) PolyU tail of rho-independent terminator of bacterial small RNAs is essential for Hfq action. *Proc. Natl. Acad. Sci. U.S.A.*, **108**(32), 13059–13064.
- [75] Landy, J., McIntosh, D., and Saleh, O. (2012) Quantifying screening ion excesses in single-molecule force-extension experiments. *Phys. Rev. Lett.*, **109**(4), 048301.
- [76] Zhang, H. and Marko, J. F. (2008) Maxwell relations for single-DNA experiments: Monitoring protein binding and double-helix torque with force-extension measurements. *Phys. Rev. E*, **77**(3), 031916.
- [77] Morrison, G., Hyeon, C., Toan, N., Ha, B., and Thirumalai, D. (2007) Stretching homopolymers. *Macromolecules*, **40**(20), 7343–7353.
- [78] Stevens, M., McIntosh, D., and Saleh, O. (2013) Simulations of Stretching a Strong, Flexible Polyelectrolyte: Using Long Chains To Access the Pincus Scaling Regime. *Macromolecules*, **46**(15), 6369–6373.
- [79] Bolton, P. and Kearns, D. (1979) Intramolecular water bridge between the 2'-OH and phosphate groups of RNA. Cyclic nucleotides as a model system. *J. Am. Chem. Soc.*, **101**(2), 479–484.

- [80] Chen, H., Meisburger, S. P., Pabit, S. A., Sutton, J. L., Webb, W. W., and Pollack, L. (2012) Ionic strength-dependent persistence lengths of single-stranded RNA and DNA. *Proc. Natl. Acad. Sci. U.S.A.*, **109**(3), 799–804.
- [81] Kuznetsov, S., Shen, Y., Benight, A., and Ansari, A. (2001) A semiflexible polymer model applied to loop formation in DNA hairpins. *Biophys. J.*, **81**(5), 2864–2875.
- [82] Pedersen, J. and Schurtenberger, P. (1996) Scattering functions of semiflexible polymers with and without excluded volume effects. *Macromolecules*, **29**(23), 7602–7612.
- [83] Rosa, A., Hoang, T., Marenduzzo, D., and Maritan, A. (2003) Elasticity of semiflexible polymers with and without self-interactions. *Macromolecules*, **36**(26), 10095–10102.
- [84] Livadaru, L., Netz, R., and Kreuzer, H. (2003) Stretching response of discrete semiflexible polymers. *Macromolecules*, **36**(10), 3732–3744.
- [85] Barrat, J.-L. and Joanny, J.-F. (1997) Theory of Polyelectrolyte Solutions. *Adv. Chem. Phys.*, **94**, 1–68.
- [86] Zhang, Y., Zhou, H., and Ou-Yang, Z.-C. (2001) Stretching single-stranded DNA: Interplay of electrostatic, base-pairing, and base-pair stacking interactions. *Biophys. J.*, **81**(2), 1133–1143.
- [87] Neukirch, S. and Marko, J. F. (2011) Analytical description of extension, torque, and supercoiling radius of a stretched twisted DNA. *Phys. Rev. Lett.*, **106**(13), 138104.
- [88] Ubbink, J. and Odijk, T. (1999) Electrostatic-undulatory theory of plectonemically supercoiled DNA. *Biophys. J.*, **76**(5), 2502–2519.
- [89] Stroobants, A., Lekkerkerker, H., and Odijk, T. (1986) Effect of electrostatic interaction on the liquid crystal phase transition in solutions of rodlike polyelectrolytes. *Macromolecules*, **19**(8), 2232–2238.
- [90] Stigter, D. (1975) The charged colloidal cylinder with a Gouy double layer. *J. Colloid Interf. Sci.*, **53**(2), 296–306.
- [91] De Vlaminck, I. and Dekker, C. (2012) Recent advances in magnetic tweezers. *Ann. Rev. Biophys.*, **41**, 453–472.
- [92] Ribbeck, N. and Saleh, O. A. (2008) Multiplexed single-molecule measurements with magnetic tweezers. *Rev. Sci. Instrum.*, **79**(9), 094301.
- [93] Zhulina, E. and Rubinstein, M. (2012) Ionic strength dependence of polyelectrolyte brush thickness. *Soft Matter*, **8**(36), 9376–9383.

- [94] Oosawa, F. (1957) A simple theory of thermodynamic properties of polyelectrolyte solutions. *J. Polym. Sci.*, **23**(103), 421–430.
- [95] Imai, N. and Onishi, T. (1959) Analytical solution of Poisson-Boltzmann equation for two-dimensional many-center problem. *J. Chem. Phys.*, **30**(4), 1115–1116.
- [96] Dittmore, A., Landy, J., Molzon, A. A., and Saleh, O. A. (2014) Single-molecule methods for ligand counting: Linking ion uptake to DNA hairpin folding. *J. Am. Chem. Soc.*, **136**(16), 5974–5980.
- [97] Lipfert, J., Skinner, G. M., Keegstra, J. M., Hensgens, T., Jager, T., Dulin, D., Köber, M., Yu, Z., Donkers, S. P., Chou, F.-C., et al. (2014) Double-stranded RNA under force and torque: Similarities to and striking differences from double-stranded DNA. *Proc. Natl. Acad. Sci. U.S.A.*, **111**(43), 15408–15413.
- [98] Foloppe, N. and MacKerell, A. D. (1998) Conformational properties of the deoxyribose and ribose moieties of nucleic acids: a quantum mechanical study. *J. Phys. Chem. B*, **102**(34), 6669–6678.
- [99] Li, X., Schroeder, C. M., and Dorfman, K. D. (2015) Modeling the stretching of wormlike chains in the presence of excluded volume. *Soft Matter*, **11**(29), 5947–5954.
- [100] Lipfert, J., Doniach, S., Das, R., and Herschlag, D. (2014) Understanding nucleic acid–ion interactions. *Annu. Rev. Biochem.*, **83**, 813.
- [101] Kirmizialtin, S., Silalahi, A. R., Elber, R., and Fenley, M. O. (2012) The ionic atmosphere around A-RNA: Poisson-Boltzmann and molecular dynamics simulations. *Biophys. J.*, **102**(4), 829–838.
- [102] Bowman, J. C., Lenz, T. K., Hud, N. V., and Williams, L. D. (2012) Cations in charge: Magnesium ions in RNA folding and catalysis. *Curr. Opin. Struct. Biol.*, **22**(3), 262–272.
- [103] Das, R., Mills, T., Kwok, L., Maskel, G., Millett, I., Doniach, S., Finkelstein, K., Herschlag, D., and Pollack, L. (2003) Counterion distribution around DNA probed by solution X-ray scattering. *Phys. Rev. Lett.*, **90**(18), 188103.
- [104] Pollack, L. (2011) SAXS studies of ion-nucleic acid interactions. *Annu. Rev. Biophys.*, **40**, 225–242.
- [105] Kirmizialtin, S., Pabit, S. A., Meisburger, S. P., Pollack, L., and Elber, R. (2012) RNA and its ionic cloud: Solution scattering experiments and atomically detailed simulations. *Biophys. J.*, **102**(4), 819–828.

- [106] Meisburger, S. P., Pabit, S. A., and Pollack, L. (2015) Determining the locations of ions and water around DNA from x-ray scattering measurements. *Biophys. J.*, **108**(12), 2886–2895.
- [107] Wales, M. and Williams, J. (1952) Effect of solvation on sedimentation experiments. *J. Polym. Sci.*, **8**(5), 449–456.
- [108] Record Jr, M. T., Zhang, W., and Anderson, C. F. (1998) Analysis of effects of salts and uncharged solutes on protein and nucleic acid equilibria and processes: A practical guide to recognizing and interpreting polyelectrolyte effects, Hofmeister effects, and osmotic effects of salts.. *Adv. Protein Chem.*, **51**, 281.
- [109] Ni, H., Anderson, C. F., and Record, M. T. (1999) Quantifying the thermodynamic consequences of cation (M^{2+} , M^{+}) accumulation and anion (X^{-}) exclusion in mixed salt solutions of polyanionic DNA using Monte Carlo and Poisson-Boltzmann calculations of ion-polyion preferential interaction coefficients. *J. Phys. Chem. B*, **103**(17), 3489–3504.
- [110] Gross, L. and Strauss, U. (1966) Interactions of polyelectrolytes with simple electrolytes. I. Theory of electrostatic potential and Donnan equilibrium for a cylindrical rod model: The effect of site-binding. In Conway, B. and Barradas, R., (eds.), *Chemical Physics of Ionic Solutions*, p. 361 John Wiley and Sons.
- [111] Moreira, A. G. and Netz, R. R. (2000) Strong-coupling theory for counter-ion distributions. *EPL-Europhys. Lett.*, **52**(6), 705.
- [112] Alexandrowicz, Z. and Katchalsky, A. (1963) Colligative properties of polyelectrolyte solutions in excess of salt. *J. Polym. Sci. Part A*, **1**(10), 3231–3260.
- [113] Skerjanc, J. and Strauss, U. (1968) Interactions of polyelectrolytes with simple electrolytes. III. The binding of magnesium ion by deoxyribonucleic acid. *J. Am. Chem. Soc.*, **90**(12), 3081–3085.
- [114] Kirkwood, J. G. (1934) On the theory of strong electrolyte solutions. *J. Chem. Phys.*, **2**(11), 767–781.
- [115] Holst, M. J., The Poisson-Boltzmann equation: Analysis and multilevel numerical solution. Technical report, California Institute of Technology (1994).
- [116] Baker, N. A., Sept, D., Joseph, S., Holst, M. J., and McCammon, J. A. (2001) Electrostatics of nanosystems: Application to microtubules and the ribosome. *Proc. Natl. Acad. Sci. U.S.A.*, **98**(18), 10037–10041.
- [117] Granot, J. (1983) Effect of finite ionic size on the solution of the Poisson-Boltzmann equation: Application to the binding of divalent metal ions to DNA. *Biopolymers*, **22**(7), 1831–1841.

- [118] Lamm, G. and Pack, G. R. (1997) Local dielectric constants and Poisson-Boltzmann calculations of DNA counterion distributions. *Int. J. Quantum Chem.*, **65**(6), 1087–1093.
- [119] Misra, V. K. and Draper, D. E. (2000) Mg 2+ binding to tRNA revisited: The nonlinear Poisson-Boltzmann model. *J. Mol. Biol.*, **299**(3), 813–825.
- [120] Misra, V. K. and Draper, D. E. (2002) The linkage between magnesium binding and RNA folding. *J. Mol. Biol.*, **317**(4), 507–521.
- [121] Shkel, I. A., Tsodikov, O. V., and Record, M. T. (2002) Asymptotic solution of the cylindrical nonlinear Poisson–Boltzmann equation at low salt concentration: Analytic expressions for surface potential and preferential interaction coefficient. *Proc. Natl. Acad. Sci. U.S.A.*, **99**(5), 2597–2602.
- [122] Chu, V. B., Bai, Y., Lipfert, J., Herschlag, D., and Doniach, S. (2007) Evaluation of ion binding to DNA duplexes using a size-modified Poisson-Boltzmann theory. *Biophys. J.*, **93**(9), 3202–3209.
- [123] Gavryushov, S. (2008) Electrostatics of B-DNA in NaCl and CaCl₂ solutions: Ion size, interionic correlation, and solvent dielectric saturation effects. *J. Phys. Chem. B*, **112**(30), 8955–8965.
- [124] Giambaşu, G. M., Luchko, T., Herschlag, D., York, D. M., and Case, D. A. (2014) Ion counting from explicit-solvent simulations and 3D-RISM. *Biophys. J.*, **106**(4), 883–894.
- [125] Tan, Z.-J. and Chen, S.-J. (2005) Electrostatic correlations and fluctuations for ion binding to a finite length polyelectrolyte. *J. Chem. Phys.*, **122**(4), 044903.
- [126] Tan, Z.-J. and Chen, S.-J. (2007) RNA helix stability in mixed Na⁺/Mg 2+ solution. *Biophys. J.*, **92**(10), 3615–3632.
- [127] Tan, Z.-J. and Chen, S.-J. (2010) Predicting ion binding properties for RNA tertiary structures. *Biophys. J.*, **99**(5), 1565–1576.
- [128] Zhu, Y. and Chen, S.-J. (2014) Many-body effect in ion binding to RNA. *J. Chem. Phys.*, **141**(5), 055101.
- [129] Mocchi, F. and Laaksonen, A. (2012) Insight into nucleic acid counterion interactions from inside molecular dynamics simulations is “worth its salt”. *Soft Matter*, **8**(36), 9268–9284.
- [130] Mills, P., Anderson, C. F., and Record Jr, M. T. (1985) Monte Carlo studies of counterion-DNA interactions. Comparison of the radial distribution of counterions with predictions of other polyelectrolyte theories. *J. Phys. Chem.*, **89**(19), 3984–3994.

- [131] Murthy, C. S., Bacquet, R. J., and Rossky, P. J. (1985) Ionic distributions near polyelectrolytes. A comparison of theoretical approaches. *J. Phys. Chem.*, **89**(4), 701–710.
- [132] Patra, C. N. and Yethiraj, A. (1999) Density functional theory for the distribution of small ions around polyions. *J. Phys. Chem. B*, **103**(29), 6080–6087.
- [133] Howard, J. J., Lynch, G. C., and Pettitt, B. M. (2011) Ion and solvent density distributions around canonical B-DNA from integral equations. *J. Phys. Chem. B*, **115**(3), 547–556.
- [134] Hayes, R. L., Noel, J. K., Whitford, P. C., Mohanty, U., Sanbonmatsu, K. Y., and Onuchic, J. N. (2014) Reduced model captures Mg²⁺-RNA interaction free energy of riboswitches. *Biophys. J.*, **106**(7), 1508–1519.
- [135] Savelyev, A. and MacKerell Jr, A. D. (2015) Competition among Li⁺, Na⁺, K⁺, and Rb⁺ monovalent ions for DNA in molecular dynamics simulations using the additive CHARMM36 and Drude polarizable force fields. *J. Phys. Chem. B*, **119**(12), 4428–4440.
- [136] Shack, J., Jenkins, R. J., and Thompsett, J. M. (1952) The binding of sodium chloride and calf thymus desoxypentose nucleic acid. *J. Biol. Chem.*, **198**(1), 85–92.
- [137] Rubin, R. L. (1977) Spermidine-deoxyribonucleic acid interaction in vitro and in *Escherichia coli*. *J. Bacteriol.*, **129**(2), 916–925.
- [138] Strauss, U. P., Helfgott, C., and Pink, H. (1967) Interactions of polyelectrolytes with simple electrolytes. II. Donnan equilibria obtained with DNA in solutions of 1-1 electrolytes. *J. Phys. Chem.*, **71**(8), 2550–2556.
- [139] Braunlin, W., Strick, T., and Record, M. (1982) Equilibrium dialysis studies of polyamine binding to DNA. *Biopolymers*, **21**(7), 1301–1314.
- [140] Shapiro, J. T., Stannard, B. S., and Felsenfeld, G. (1969) Binding of small cations to deoxyribonucleic acid. Nucleotide specificity. *Biochemistry*, **8**(8), 3233–3241.
- [141] Pabit, S. A., Meisburger, S. P., Li, L., Blose, J. M., Jones, C. D., and Pollack, L. (2010) Counting ions around DNA with anomalous small-angle X-ray scattering. *J. Am. Chem. Soc.*, **132**(46), 16334–16336.
- [142] Lindström, T., de Ruvo, A., and Söremark, C. (1977) Determination of Donnan equilibria by gel filtration. *J. Polym. Sci. Pol. Chem.*, **15**(8), 2029–2032.
- [143] Nguyen, H. T., Pabit, S. A., Meisburger, S. P., Pollack, L., and Case, D. A. (2014) Accurate small and wide angle x-ray scattering profiles from atomic models of proteins and nucleic acids. *J. Chem. Phys.*, **141**(22), 22D508.

- [144] Nguyen, H. T., Pabit, S. A., Pollack, L., and Case, D. A. (2016) Extracting water and ion distributions from solution x-ray scattering experiments. *J. Chem. Phys.*, **144**(21), 214105.
- [145] Gebala, M., Bonilla, S., Bisaria, N., and Herschlag, D. (2016) Does cation size affect occupancy and electrostatic screening of the nucleic acid ion atmosphere?. *J. Am. Chem. Soc.*, **138**, 10925–10934.
- [146] Meisburger, S. P., Sutton, J. L., Chen, H., Pabit, S. A., Kirmizialtin, S., Elber, R., and Pollack, L. (2013) Polyelectrolyte properties of single stranded DNA measured using SAXS and single-molecule FRET: Beyond the wormlike chain model. *Biopolymers*, **99**(12), 1032–1045.
- [147] Holland, J. G. and Geiger, F. M. (2012) Importance of length and sequence order on magnesium binding to surface-bound oligonucleotides studied by second harmonic generation and atomic force microscopy. *J. Phys. Chem. B*, **116**(22), 6302–6310.
- [148] Walter, S. R., Young, K. L., Holland, J. G., Gieseck, R. L., Mirkin, C. A., and Geiger, F. M. (2013) Counting the number of magnesium ions bound to the surface-immobilized thymine oligonucleotides that comprise spherical nucleic acids. *J. Am. Chem. Soc.*, **135**(46), 17339–17348.
- [149] Record, M. T. (1967) Electrostatic effects on polynucleotide transitions. II. Behavior of titrated systems. *Biopolymers*, **5**(10), 993–1008.
- [150] Record, M. T. (1975) Effects of Na⁺ and Mg⁺⁺ ions on the helix–coil transition of DNA. *Biopolymers*, **14**(10), 2137–2158.
- [151] Serra, M. J., Baird, J. D., Dale, T., Fey, B. L., Retatagos, K., and Westhof, E. (2002) Effects of magnesium ions on the stabilization of RNA oligomers of defined structures. *RNA*, **8**(03), 307–323.
- [152] Bond, J. P., Anderson, C. F., and Record Jr, M. T. (1994) Conformational transitions of duplex and triplex nucleic acid helices: Thermodynamic analysis of effects of salt concentration on stability using preferential interaction coefficients.. *Biophys. J.*, **67**(2), 825.
- [153] Williams, D. J. and Hall, K. B. (1996) Thermodynamic comparison of the salt dependence of natural RNA hairpins and RNA hairpins with non-nucleotide spacers. *Biochemistry*, **35**(46), 14665–14670.
- [154] Owczarzy, R., You, Y., Moreira, B. G., Manthey, J. A., Huang, L., Behlke, M. A., and Walder, J. A. (2004) Effects of sodium ions on DNA duplex oligomers: Improved predictions of melting temperatures. *Biochemistry*, **43**(12), 3537–3554.

- [155] Stellwagen, E., Muse, J. M., and Stellwagen, N. C. (2011) Monovalent cation size and DNA conformational stability. *Biochemistry*, **50**(15), 3084–3094.
- [156] Reiling, C., Khutsishvili, I., Huang, K., and Marky, L. A. (2015) Loop contributions to the folding thermodynamics of DNA straight hairpin loops and pseudoknots. *J. Phys. Chem. B*, **119**(5), 1939–1946.
- [157] Todd, B. A. and Rau, D. C. (2008) Interplay of ion binding and attraction in DNA condensed by multivalent cations. *Nucleic Acids Res.*, **36**(2), 501–510.
- [158] Andresen, K., Qiu, X., Pabit, S. A., Lamb, J. S., Park, H. Y., Kwok, L. W., and Pollack, L. (2008) Mono- and trivalent ions around DNA: A small-angle scattering study of competition and interactions. *Biophys. J.*, **95**(1), 287–295.
- [159] Braunlin, W., Anderson, C., and Record, M. T. (1986) ^{23}Na -NMR investigations of counterion exchange reactions of helical DNA. *Biopolymers*, **25**(1), 205–214.
- [160] Li, A. Z., Huang, H., Re, X., Qi, L. J., and Marx, K. A. (1998) A gel electrophoresis study of the competitive effects of monovalent counterion on the extent of divalent counterions binding to DNA. *Biophys. J.*, **74**(2), 964–973.
- [161] Sander, C. and Ts'o, P. O. (1971) Interaction of nucleic acids: VIII. Binding of magnesium ions by nucleic acids. *J. Mol. Biol.*, **55**(1), 1–21.
- [162] Hansen, P. L., Podgornik, R., and Parsegian, V. A. (2001) Osmotic properties of DNA: Critical evaluation of counterion condensation theory. *Phys. Rev. E*, **64**(2), 021907.
- [163] Macke, T. and Case, D. A. (1998) Modeling unusual nucleic acid structures. In Leontes, N. B. and SantaLucia, J., (eds.), *Molecular Modeling of Nucleic Acids*, pp. 379–393 American Chemical Society.
- [164] Krakauer, H. (1971) The binding of Mg^{++} ions to polyadenylate, polyuridylate, and their complexes. *Biopolymers*, **10**(12), 2459–2490.
- [165] Willemsen, A. and Van Os, G. (1971) Interaction of magnesium ions with poly A and poly U. *Biopolymers*, **10**(6), 945–960.
- [166] Yoo, J. and Aksimentiev, A. (2012) Competitive binding of cations to duplex DNA revealed through molecular dynamics simulations. *J. Phys. Chem. B*, **116**(43), 12946–12954.
- [167] Olmsted, M. C., Anderson, C. F., and Record, M. T. (1991) Importance of oligo-electrolyte end effects for the thermodynamics of conformational transitions of nucleic acid oligomers: A grand canonical Monte Carlo analysis. *Biopolymers*, **31**(13), 1593–1604.

- [168] Record, M. T., Anderson, C. F., and Lohman, T. M. (1978) Thermodynamic analysis of ion effects on the binding and conformational equilibria of proteins and nucleic acids: The roles of ion association or release, screening, and ion effects on water activity. *Q. Rev. Biophys.*, **11**(02), 103–178.
- [169] Hori, N., Denesyuk, N. A., and Thirumalai, D. (2016) Salt effects on the thermodynamics of a frameshifting RNA pseudoknot under tension. *J. Mol. Biol.*, **428**, 2847–2859.
- [170] Goddard, N. L., Bonnet, G., Krichevsky, O., and Libchaber, A. (2000) Sequence dependent rigidity of single stranded DNA. *Phys. Rev. Lett.*, **85**(11), 2400.
- [171] Shkel, I. A. and Record, M. T. (2004) Effect of the number of nucleic acid oligomer charges on the salt dependence of stability (ΔG_{37}°) and melting temperature (T_m): NLPB analysis of experimental data. *Biochemistry*, **43**(22), 7090–7101.
- [172] MacKerell, A. D. (1997) Influence of magnesium ions on duplex DNA structural, dynamic, and solvation properties. *J. Phys. Chem. B*, **101**(4), 646–650.
- [173] Tan, Z.-J. and Chen, S.-J. (2006) Nucleic acid helix stability: Effects of salt concentration, cation valence and size, and chain length. *Biophys. J.*, **90**(4), 1175–1190.
- [174] Shkel, I. A. and Record, M. T. (2012) Coulombic free energy and salt ion association per phosphate of all-atom models of DNA oligomer: Dependence on oligomer size. *Soft Matter*, **8**(36), 9345–9355.
- [175] Sim, A. Y., Lipfert, J., Herschlag, D., and Doniach, S. (2012) Salt dependence of the radius of gyration and flexibility of single-stranded DNA in solution probed by small-angle x-ray scattering. *Phys. Rev. E*, **86**(2), 021901.
- [176] Landy, J., Lee, Y., and Jho, Y. (2013) Limiting law excess sum rule for polyelectrolytes. *Phys. Rev. E*, **88**(5), 052315.
- [177] Draper, D. E., Grilley, D., and Soto, A. M. (2005) Ions and RNA Folding. *Annu. Rev. Bioph. Biom.*, **34**(1), 221–243.
- [178] van den Hout, M., Skinner, G., Klijnhout, S., Krudde, V., and Dekker, N. (2011) The Passage of Homopolymeric RNA through Small Solid-State Nanopores. *Small*, **7**(15), 2217–2224.
- [179] Rio, D. C. (2011) RNA : A Laboratory Manual, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
- [180] Brockman, C., Kim, S. J., and Schroeder, C. M. (2011) Direct observation of single flexible polymers using single stranded DNA. *Soft Matter*, **7**(18), 8005–8012.

- [181] Wickersham, C. E., Kerr, D. H., and Lipman, E. A. (2010) Synthesis of extended nanoscale optical encoders. *Bioconjugate Chem.*, **21**(12), 2234–2238.
- [182] Lansdorp, B. Pushing the Envelope of Magnetic Tweezer Resolution PhD thesis University of California, Santa Barbara (2015).
- [183] Gosse, C. and Croquette, V. (2002) Magnetic tweezers: Micromanipulation and force measurement at the molecular level. *Biophys. J.*, **82**(6), 3314–3329.
- [184] Park, C.-Y., Jacobson, D. R., Nguyen, D. T., Willardson, S., and Saleh, O. A. (2016) A thin permeable-membrane device for single-molecule manipulation. *Rev. Sci. Instrum.*, **87**(1), 014301.
- [185] Strick, T., Allemand, J.-F., Bensimon, D., Bensimon, A., and Croquette, V. (1996) The elasticity of a single supercoiled DNA molecule. *Science*, **271**(5257), 1835–1837.
- [186] Lansdorp, B. M. and Saleh, O. A. (2012) Power spectrum and Allan variance methods for calibrating single-molecule video-tracking instruments. *Rev. Sci. Instrum.*, **83**(2), 025115.
- [187] Te Velthuis, A. J., Kerssemakers, J. W., Lipfert, J., and Dekker, N. H. (2010) Quantitative guidelines for force calibration through spectral analysis of magnetic tweezers data. *Biophys. J.*, **99**(4), 1292–1302.
- [188] Nørrelykke, S. F. and Flyvbjerg, H. (2010) Power spectrum analysis with least-squares fitting: Amplitude bias and its elimination, with application to optical tweezers and atomic force microscope cantilevers. *Rev. Sci. Instrum.*, **81**(7), 075103.
- [189] Wong, W. P. and Halvorsen, K. (2006) The effect of integration time on fluctuation measurements: calibrating an optical trap in the presence of motion blur. *Opt. Express*, **14**(25), 12517–12531.
- [190] Berg-Sørensen, K. and Flyvbjerg, H. (2004) Power spectrum analysis for optical tweezers. *Rev. Sci. Instrum.*, **75**(3), 594–612.
- [191] Leach, J., Mushfique, H., Keen, S., Di Leonardo, R., Ruocco, G., Cooper, J., and Padgett, M. (2009) Comparison of Faxén’s correction for a microsphere translating or rotating near a surface. *Phys. Rev. E*, **79**(2), 026301.
- [192] Upadhyayula, S., Quinata, T., Bishop, S., Gupta, S., Johnson, N., Bahmani, B., Bozhilov, K., Stubbs, J., Jreij, P., Nallagatla, P., and Vullev, V. (2012) Coatings of polyethylene glycol for suppressing adhesion between solid microspheres and flat surfaces. *Langmuir*, **28**(11), 5059–5069.

- [193] Liphardt, J., Onoa, B., Smith, S. B., Tinoco, I., and Bustamante, C. (2001) Reversible unfolding of single RNA molecules by mechanical force. *Science*, **292**(5517), 733–737.
- [194] Leipply, D., Lambert, D., and Draper, D. E. (2009) Ion–RNA interactions: Thermodynamic analysis of the effects of mono- and divalent ions on RNA conformational equilibria. *Method. Enzymol.*, **469**, 433–463.
- [195] Dolinsky, T. J., Czodrowski, P., Li, H., Nielsen, J. E., Jensen, J. H., Klebe, G., and Baker, N. A. (2007) PDB2PQR: Expanding and upgrading automated preparation of biomolecular structures for molecular simulations. *Nucleic Acids Res.*, **35**(suppl 2), W522–W525.
- [196] Dolinsky, T. J., Nielsen, J. E., McCammon, J. A., and Baker, N. A. (2004) PDB2PQR: An automated pipeline for the setup of Poisson–Boltzmann electrostatics calculations. *Nucleic Acids Res.*, **32**(suppl 2), W665–W667.
- [197] Rossi, A. M. and Taylor, C. W. (2011) Analysis of protein–ligand interactions by fluorescence polarization. *Nat. Protoc.*, **6**(3), 365–387.
- [198] Johannsen, S., Korth, M. M., Schnabl, J., and Sigel, R. K. (2009) Exploring metal ion coordination to nucleic acids by NMR. *CHIMIA*, **63**(3), 146–152.
- [199] Greenleaf, W. J., Frieda, K. L., Foster, D. A., Woodside, M. T., and Block, S. M. (2008) Direct observation of hierarchical folding in single riboswitch aptamers. *Science*, **319**(5863), 630–633.
- [200] Lang, M. J., Asbury, C. L., Shaevitz, J. W., and Block, S. M. (2002) An automated two-dimensional optical force clamp for single molecule studies. *Biophys. J.*, **83**(1), 491–501.
- [201] Keller, D., Swigon, D., and Bustamante, C. (2003) Relating single-molecule measurements to thermodynamics. *Biophys. J.*, **84**(2), 733–738.
- [202] Adamson, A. W. (1979) *A Textbook of Physical Chemistry*, Academic Press, 2nd edition.
- [203] Reif, F. (1965) *Fundamentals of Statistical and Thermal Physics*, McGraw-Hill, .
- [204] Rouzina, I. and Bloomfield, V. A. (2001) Force-induced melting of the DNA double helix 1. Thermodynamic analysis. *Biophys. J.*, **80**(2), 882–893.
- [205] Manosas, M. and Ritort, F. (2005) Thermodynamic and kinetic aspects of RNA pulling experiments. *Biophys. J.*, **88**(5), 3224–3242.
- [206] Jarzynski, C. (1997) Nonequilibrium equality for free energy differences. *Phys. Rev. Lett.*, **78**(14), 2690.

- [207] Liphardt, J., Dumont, S., Smith, S. B., Tinoco, I., and Bustamante, C. (2002) Equilibrium information from nonequilibrium measurements in an experimental test of Jarzynski’s equality. *Science*, **296**(5574), 1832–1835.
- [208] Harris, N. C., Song, Y., and Kiang, C.-H. (2007) Experimental free energy surface reconstruction from single-molecule force spectroscopy using Jarzynski’s equality. *Phys. Rev. Lett.*, **99**(6), 068101.
- [209] Bizarro, C., Alemany, A., and Ritort, F. (2012) Non-specific binding of Na⁺ and Mg²⁺ to RNA determined by force spectroscopy methods. *Nucleic Acids Res.*, **40**(14), 6922–6935.
- [210] Grilley, D., Soto, A. M., and Draper, D. E. (2006) Mg²⁺–RNA interaction free energies and their relationship to the folding of RNA tertiary structures. *P. Natl. Acad. Sci. USA*, **103**(38), 14003–14008.
- [211] Bukhman, Y. V. and Draper, D. E. (1997) Affinities and selectivities of divalent cation binding sites within an RNA tertiary structure. *J. Mol. Biol.*, **273**(5), 1020–1031.
- [212] Schellman, J. A. (1975) Macromolecular binding. *Biopolymers*, **14**(5), 999–1018.
- [213] Carrillo, J.-M. Y. and Dobrynin, A. V. (2011) Polyelectrolytes in salt solutions: Molecular dynamics simulations. *Macromolecules*, **44**(14), 5798–5816.
- [214] Ullner, M. and Woodward, C. E. (2002) Orientational correlation function and persistence lengths of flexible polyelectrolytes. *Macromolecules*, **35**(4), 1437–1445.
- [215] Nguyen, T. and Shklovskii, B. (2002) Persistence length of a polyelectrolyte in salty water: Monte Carlo study. *Phys. Rev. E*, **66**(2), 021801.
- [216] Everaers, R., Milchev, A., and Yamakov, V. (2002) The electrostatic persistence length of polymers beyond the OSF limit. *Eur. Phys. J. E*, **8**(1), 3–14.
- [217] Bloomfield, V. A., Crothers, D. M., and Tinoco Jr, I. (1974) *Physical Chemistry of Nucleic Acids*, Harper and Row, New York.
- [218] Li, P. T. (2013) Analysis of diffuse K⁺ and Mg²⁺ ion binding to a two-base-pair kissing complex by single-molecule mechanical unfolding. *Biochemistry*, **52**(29), 4991–5001.
- [219] Woodside, M. T., Behnke-Parks, W. M., Larizadeh, K., Travers, K., Herschlag, D., and Block, S. M. (2006) Nanomechanical measurements of the sequence-dependent folding landscapes of single nucleic acid hairpins. *Proc. Natl. Acad. Sci. U.S.A.*, **103**(16), 6190–6195.

- [220] Long, X., Parks, J. W., Bagshaw, C. R., and Stone, M. D. (2013) Mechanical unfolding of human telomere G-quadruplex DNA probed by integrated fluorescence and magnetic tweezers spectroscopy. *Nucleic Acids Res.*, **41**(4), 2746–2755.
- [221] Cecconi, C., Shank, E. A., Bustamante, C., and Marqusee, S. (2005) Direct observation of the three-state folding of a single protein molecule. *Science*, **309**(5743), 2057–2060.
- [222] Brutzer, H., Luzzietti, N., Klaue, D., and Seidel, R. (2010) Energetics at the DNA supercoiling transition. *Biophys. J.*, **98**(7), 1267–1276.
- [223] Lansdorp, B. M., Tabrizi, S. J., Dittmore, A., and Saleh, O. A. (2013) A high-speed magnetic tweezer beyond 10,000 frames per second. *Rev. Sci. Instrum.*, **84**(4), 044301.
- [224] Dulin, D., Cui, T. J., Cnossen, J., Docter, M. W., Lipfert, J., and Dekker, N. H. (2015) High spatiotemporal-resolution magnetic tweezers: calibration and applications for dna dynamics. *Biophys. J.*, **109**(10), 2113–2125.
- [225] Huhle, A., Klaue, D., Brutzer, H., Daldrop, P., Joo, S., Otto, O., Keyser, U. F., and Seidel, R. (2015) Camera-based three-dimensional real-time particle tracking at kHz rates and Ångström accuracy. *Nat. Commun.*, **6**.
- [226] Yu, Z., Dulin, D., Cnossen, J., Köber, M., van Oene, M. M., Ordu, O., Berghuis, B. A., Hensgens, T., Lipfert, J., and Dekker, N. H. (2014) A force calibration standard for magnetic tweezers. *Rev. Sci. Instrum.*, **85**(12), 123114.
- [227] Chen, H., Fu, H., Zhu, X., Cong, P., Nakamura, F., and Yan, J. (2011) Improved high-force magnetic tweezers for stretching and refolding of proteins and short DNA. *Biophys. J.*, **100**(2), 517–523.

Coda

A sonnet, presented as the abstract to my group meeting talk on May 12, 2016. Based upon the opening chorus of Romeo and Juliet by William Shakespeare.

Two models, both alike in dignity,
In the log regime, where we lay our scene.
From charge and bending, break new mutiny.
From curve fitting, sense we attempt to glean.
From forth fatal loins of mathematics,
Do modified wormlike chains come to life;
Whose diff'ring hand'ling of 'lectrostatics,
Are with subtlety entirely rife.
To understand and to, needs be, improve
The models herein by our data gauged—
Data that, against trials, may not prove—
Is now my one hours' traffic to be staged;
The which, if you with patient ears attend,
What here shall miss, my toil shall strive to mend.