Published in final edited form as:
J Am Chem Soc. 2012 March 14; 134(10): 4607–4614. doi:10.1021/ja208466h.

Electrostatics of nucleic acid folding under conformational constraint

Peter C. Anthony[†], **Adelene Y.L. Sim**[‡], **Vincent B. Chu**[‡], **Sebastian Doniach**[‡,§], **Steven M. Block**[‡,¶,*], and **Daniel Herschlag**[‖,*]

[†]Biophysics Program, Stanford University, Stanford, CA 94305

[‡]Department of Applied Physics, Stanford University, Stanford, CA 94305

[§]Department of Physics, Stanford University, Stanford, CA 94305

[¶]Department of Biology, Stanford University, Stanford, CA 94305

[‖]Department of Biochemistry, Stanford University, Stanford, CA 94305

Abstract

RNA folding is enabled by interactions between the nucleic acid and its ion atmosphere, the mobile sheath of aqueous ions that surrounds and stabilizes it. Understanding the ion atmosphere requires the interplay of experiment and theory. However, even an apparently simple experiment to probe the ion atmosphere—measuring the dependence of DNA duplex stability upon ion concentration and identity—suffers from substantial complexity, because the unfolded ensemble contains many conformational states that are difficult to treat accurately with theory. To minimize this limitation, we measured the unfolding equilibrium of a DNA hairpin using a single-molecule optical trapping assay, in which the unfolded state is constrained to a limited set of elongated conformations. The unfolding free energy increased linearly with the logarithm of monovalent cation concentration for several cations, such that smaller cations tended to favor the folded state. Mg^{2+} stabilized the hairpin much more effectively at low concentrations than did any of the monovalent cations. Poisson-Boltzmann theory captured trends in hairpin stability measured for the monovalent cation titrations with reasonable accuracy, but failed to do so for the Mg^{2+} titrations. This finding is consistent with previous work suggesting that Poisson-Boltzmann and other mean-field theories fail for higher valency cations where ionion correlation effects may become significant. The high-resolution data herein, because of the straightforward nature of both the folded and unfolded states, should serve as benchmarks for the development of more accurate electrostatic theories that will be needed for a more quantitative and predictive understanding of nucleic acid folding.

INTRODUCTION

RNA molecules carry genetic information and regulate gene expression in all kingdoms of life, and their cellular roles often depend on the intricate structures they can adopt and the interactions they make[1]. The acquisition of a functional form involves the pairing of single-stranded RNA into helices and, often, the juxtaposition of these helices into tertiary structure. For the folding process to be favorable, the energy of base pairing, stacking, and

[*]Corresponding authors. sblock@stanford.edu, herschla@stanford.edu.

Supporting Information Available. Additional experimental procedures; figures illustrating details of instrument calibration and PB calculations; figures presenting additional analyses of data and comparisons with HEL, PB, SMPB and *mfold* results; and tables summarizing experimental data and theoretical results. This material is available free of charge via the Internet at http://pubs.acs.org.

tertiary interactions must offset a concomitant reduction in the overall conformational entropy and an increase in the electrostatic repulsion between phosphate groups brought into proximity. Because RNA is strongly negatively charged, intra-molecular electrostatic repulsion would present an insurmountable barrier to folding, were it not mitigated by ions present in the immediate aqueous environment[2]. This environment, termed the 'ion atmosphere,' is enriched in cations and depleted of anions relative to the bulk solvent[3,4], and rearranges in the event of folding or protein binding[5].

To understand the role of the ion atmosphere in nucleic acid folding, an interplay of experimental studies and electrostatic theories is needed. The ion atmosphere is fluctuating and unstructured, making it challenging to probe both experimentally and theoretically[6-8]. The Poisson-Boltzmann (PB) theory takes a mean-field approach, treating the solvent as a continuum dielectric and the ions as non-interacting point charges[6,9]. This approach has been widely implemented, because of its relative simplicity and because several early studies comparing PB predictions with experimental observations found reasonable agreement (e.g., refs 10-12).

Studies of complex RNAs have provided useful insights into general folding behaviors, but they also involve caveats that limit their ability to evaluate theories rigorously. The electrostatic 'relaxation' of RNA molecules upon the addition of monovalent or divalent cations[13-15] was shown to be broadly consistent with expectations from polyelectrolyte theories, with much lower concentrations of divalent cations than monovalent cations required to attain equivalent relaxation effects[13,14,16]. Furthermore, some measured aspects of Mg^{2+}-induced tRNA[17], 23S rRNA[18] and RNA pseudoknot[19] folding have been fit by PB calculations with reasonable success. However, such success may arise not because PB provides an accurate description of ion atmosphere-related energetics, but because the fits involved simplified treatments of molecular ensembles (e.g., the unfolded and intermediate states were represented by small subsets of the full ensemble of possible conformers, or by *ad hoc* models) and adjustable parameters (e.g., the energetic contribution from tertiary structure formation). Moreover, in some cases, the fits did not take into account energetic contributions from specific metal-ion binding.

In contrast to RNAs containing complex structures, smaller and simpler nucleic acids facilitate more direct comparisons between measurement and theory. Short DNA duplexes are structurally rigid, permitting direct probing of the shape[20-22] and content[3,23] of their ion atmospheres. However, a fully predictive understanding of the ion atmosphere should also describe its consequences for the thermodynamics of folding, conformational transitions, and binding events. To probe energetics in the context of a simple folding process and further evaluate PB, one recent study examined two DNA duplexes tethered by a simple neutral linker[24]. This system showed that PB theory can reasonably account for at least some aspects of the thermodynamic roles of monovalent cations, but cannot account for the effects of divalent cations in screening charges and relaxing conformational ensembles. Similar conclusions were drawn from studies of the ion atmosphere in the context of DNA duplex melting[25-29]. The two thermodynamic states in duplex melting correspond to all duplex base pairs being intact or broken, and define a measurable free-energy change (ΔG). But even for this apparently simple system, the comparison of experimental and theoretical trends in free energies is hindered by the extensive structural heterogeneity of the unfolded duplex ensemble[30-32] (Figure 1).

Here, we introduce a new approach to evaluate ion atmosphere effects on folding, where the unfolded state is well defined. We followed the unfolding of a DNA hairpin using a single-molecule optical trapping assay that exerts tension on the 3′ and 5′ ends of the molecule (Figure 2A). Because the unfolded ensemble is physically constrained, it is extended and

less structurally heterogeneous than a thermally or chemically denatured ensemble (Figure 1). This constraint allowed us to more fully represent the unfolded ensemble with all-atom models in PB calculations, rather than with the single cylindrical[25,26,29,33,34] or 'grooved' models[27,28] used previously. We measured trends in ΔG across a range of concentrations of several monovalent cations and Mg^{2+}. Comparison of our experimental data and predictions from PB theory recapitulates the previously reported strengths and shortcomings of PB. More generally, these and additional force-induced unfolding data may provide simple and therefore powerful benchmarks for further tests and development of electrostatic theories.

EXPERIMENTAL SECTION

Optical trapping assay

The 20R55/4T DNA hairpin[35] (Figure 2B) was attached to ~1 kb double-stranded DNA (dsDNA) handles and ~1 μm-diameter polystyrene beads to form dumbbells, which were measured in ionic solutions at 23 ± 0.5 °C using a dual-beam optical trapping instrument, as described (ref 35; see also Supporting Experimental Procedures). Trap stiffness was calibrated regularly as described[36] by suspending beads in water and measuring Brownian fluctuations, Lorentzian roll-off frequencies and displacements under flow. When measuring in buffer, each dumbbell was first stretched to obtain a force-extension curve (FEC), with intensities of the traps adjusted to provide stiffnesses in the range of 0.2-0.3 pN/nm. The portion of the FEC corresponding to 0-8 pN was fit to the worm-like chain (WLC) interpolation formula[37,38], and the dumbbell was discarded if the persistence length obtained was <20 nm (indicative of a multiple tether[35]). Otherwise, the dumbbell was measured in the passive optical force clamp[39], with one bead pulled into the zero-stiffness region (ZSR) of the weaker trap, which was maintained ~3-fold less stiff than the other. As the intensity of this trap (and therefore the force) was stepped down sequentially, the bead moved within the ZSR and the hairpin was observed to fold and unfold stochastically, spending proportionately more time in the folded state (~ ≤1% to ≥99%). The force was held constant for the duration of each step (10 s). The displacements of the beads from the traps, from which forces and extensions were calculated, were sampled at 20 kHz, low-pass Bessel-filtered online at 10 kHz and median-smoothed offline over 3.1 ms, a window size determined empirically to preserve the briefest hairpin folding events under the experimental conditions.

Data analysis

The data obtained for each hairpin molecule consisted of several (≥5) 10 s records of extension and force. As previously[35], each record was partitioned into two states, corresponding to the hairpin being fully open or closed, and sorted into 0.1 nm bins. Histograms of extension derived from the two states were well fit by Gaussians (Figure 2C), and the distance between the Gaussian peaks was interpreted as the raw hairpin opening distance. The final opening distance (Δx) for the record was obtained by rescaling this raw distance by an amount determined by the displacement of the bead in the weak trap from the trap center (see ref 39 and Figure S1). This rescaling factor averaged 0.97 (± 0.02 standard deviation) for all records from all molecules. Δx for each molecule was the average of the Δx values for the individual records, excluding those in which the hairpin spent >99% of the time in one state.

$F_{1/2}$, the force at which the hairpin molecule spends equal times in the open and closed states, was determined from the midpoint of a Boltzmann fit to a plot of the fraction of time in the unfolded state versus force, with each point representing one record. Raw forces were based on trap stiffness calibrations performed in water. It is known that trap stiffness falls when solutes are added to an aqueous medium, arising mainly from an increase in the

refractive index[40]. We re-measured the trap stiffness for every monovalent working buffer using the method of Brownian fluctuations, which is insensitive to minor differences in viscosity that may exist between buffers, and normalized the stiffness in each buffer to that measured in water (Figure S2). The forces for all molecules measured in each buffer were then scaled to the normalized stiffness. The stiffness did not change significantly when 10 mM $MgCl_2$ was added to the 40 mM and 200 mM K^+ buffers; we therefore used the normalized stiffness for the corresponding monovalent buffer when rescaling the divalent force data.

ΔG, the free energy of opening the hairpin, was the product of Δx and $F_{1/2}$. Uncertainties in these parameters are discussed in the Supporting Experimental Procedures.

Hairpin energy landscape (HEL) model predictions

For each monovalent and divalent salt concentration, the model[35,41] was run using the appropriate [Na^+] and [Mg^{2+}]-dependent DNA *mfold* energies for the hairpin base pairs and loop. (The sum of the *mfold* energies equaled, within error, the sum of the corresponding energies recently determined by Huguet et al.[42] across the range of monovalent cation concentrations we tested. We therefore would not expect the HEL model to generate significantly different results using the latter set of energies.) Also input into the model were combinations of WLC parameters corresponding to integral values of force between 10 and 19 pN (inclusive), the approximate range of all experimental $F_{1/2}$ values. Values for force-dependent extension per nucleotide (D_b) of single-stranded DNA (ssDNA) were taken from previously published measurements of charomid DNA under low-salt conditions[43]. The per-nucleotide contour length (D_p) was determined by dividing D_b by the fractional extension of ssDNA at the appropriate force and ionic strength, as modeled analytically by Cocco and co-workers[44]. ssDNA persistence length (L_p) was estimated using six previously published formulations that embody different dependencies on force and ion concentration (see Figure S3). The HEL model was run a total of 60 times per salt concentration (once per L_p estimate per force, at 10 different forces). As was done previously[35], model results corresponding to different input parameters were averaged together for each condition. The reported model values of $F_{1/2}$, Δx and ΔG (Figure 3, Figure 4, Table S1) represent the means ± standard deviations of these parameters from all successful runs (2290 out of 2340 for all conditions).

Atomic models for Poisson-Boltzmann calculations

An initial atomic model of the folded DNA hairpin attached to two 25-bp handle fragments was created using Nucleic Acid Builder (NAB)[45], using the same sequences that were present experimentally, and assuming B-form structures for the helical regions. The abasic sites separating the hairpin from its handles were replaced by thymidine residues in the model. To represent the system at forces just below the hairpin unfolding force, the handle fragments were placed as far apart as possible without causing unreasonable bond lengths. From this point, the handle fragments were moved a further 18 nm apart, in accordance with a previous measurement of this hairpin sequence[35], to generate the initial unfolded model.

Control calculations performed on models incorporating a 17 nm distance change, a 5 Å/bp rise of the handle DNA, 10 bp handle fragments, or no handle DNA at all did not produce substantially different trends in unfolding energies (Table S2). Molecular dynamics simulations were conducted to generate molecular ensembles from the initial structures (see Supporting Experimental Procedures).

Poisson-Boltzmann calculations

We next calculated the electrostatic contributions to the unfolding energy of the DNA hairpin for each possible pair of folded and unfolded models, at each experimentally tested

monovalent and divalent ion concentration. We assumed a thermodynamic cycle (Figure S4) in which the hairpin can fold in solution or in a vacuum, and determined the $[M^+]$-dependent (or $[Mg^{2+}]$-dependent) ΔG of unfolding in solution (ΔG_{el}) as the appropriate sum of the ΔG of solvation (ΔG_F, ΔG_U) and folding in a vacuum ($\Delta G_{Coulomb}$). The ΔG_F and ΔG_U values were calculated using the Adaptive Poisson-Boltzmann Solver (APBS) 1.1.0[47] and the AMBER force field, as implemented in the PDB2PQR utility[48,49], while the $\Delta G_{Coulomb}$ values were evaluated using the APBS accessory program *coulomb* (see Supporting Experimental Procedures for further details). Substituting the CHARMM force field for the AMBER force field in the non-linear Poisson-Boltzmann calculations did not significantly affect trends in ΔG_{el} (Table S2).

Size-modified Poisson-Boltzmann calculations were also conducted using an in-house modified version of APBS[50], and the results did not strongly differ from those of standard PB (see Table 1 and Table S3).

RESULTS

The effect of varying cation concentration and identity on hairpin stability under force

Each molecule was measured in an optical force clamp[35,39], which allows observation of the hairpin dynamically sampling the folded and unfolded states. Records of molecular end-to-end extension versus time revealed two stable states and directly yielded Δx, the change in extension between them. The value of ΔG for each molecule was calculated as the product of Δx and $F_{1/2}$, the force at which the hairpin spends equal time in the two states, which was determined by gradually varying the applied force (see Figure 2C and **Experimental Section**).

To characterize the sensitivity of ΔG to the ion atmosphere, the unfolding process was measured across a range of monovalent ion concentrations. For each ion, ΔG increased approximately linearly with $\log[M^+]$ (Figure 3A, Tables S4-S8), consistent with previous thermodynamic frameworks for duplex dissociation that separate ΔG into electrostatic (ΔG_{el}) and non-electrostatic ($\Delta G_{non\text{-}el}$) components[29,33,51-55] (Eq. 1).

$$\Delta G = \Delta G_{el} + \Delta G_{nonel} \quad (1)$$

Electrostatic calculations are best used to supply *differences* in ΔG from one ionic condition to another—i.e., $\Delta\Delta G$. Thus, we treat $\Delta G_{non\text{-}el}$ as a constant in order to evaluate the electrostatic component, ΔG_{el}, both experimentally and computationally. In practice, this approximation holds over a wide range of ionic conditions (refs 25,26,33,51; see also refs 27,56). Here, $\Delta G_{non\text{-}el}$ incorporates, for example, the energies to disrupt base pairing and stacking. ΔG_{el} arises mainly from releasing cations (and a corresponding number of anions, by charge neutrality) that are held close to the DNA to the bulk solvent[4,57] upon duplex dissociation. This release, which yields less entropic benefit as the bulk concentrations of ions increase, occurs because ssDNA has a lower density of negative charges than dsDNA. The number of cations released (Δn) has been observed to be constant and independent of DNA concentration when bulk ions are present in excess (~0.1 mM – 1 M M^+). Under these conditions, ΔG_{el} is proportional to the logarithm of M^+ [4,51,57-59]:

$$\Delta G_{el} \sim k_B T \cdot \Delta n \cdot \ln[M^+]. \quad (2)$$

A decrease in ΔG (the total energy to unfold the hairpin) indicates destabilization of the folded state relative to the unfolded state. We observed downward shifts in the measured ΔG

series as the hydrated ionic radius[46] was increased at each [M^+] (Figure 3A, Figure S5A). This trend is consistent with larger counterions associating less closely and in smaller numbers with the folded hairpin. Such tendencies were suggested by experiments designed to monitor the competition between cations associating with DNA duplexes (Figure S5B)[3,60]. Our results are also consistent with previous data revealing that duplexes[61-66] (as well as large RNAs[13,67]) are destabilized as larger cations in the ion atmosphere are substituted for smaller ones.

The slopes of lines fitted to the measured ΔG series (Figure 3A), which are proportional to Δn, were positively correlated with ionic radius (Table 1), although differences were not significant beyond experimental error, and Li^+ (the smallest cation) did not follow this trend. Previous studies found that Δn exhibits little variation among the smallest monovalent cations[62,64] but is lower for larger alkylammonium cations such as tetrame-thylammonium (TMA^+)[61,64].

Most of the differences in ΔG within each monovalent titration, and also between titrations, arose from differences in $F_{1/2}$ (Figure 3C, Tables S4-S8): the measured values of Δx exhibited little variation, with a mean and standard deviation of 17.4 ± 0.4 nm for all titrations (Figure 3D, Tables S4-S8). Previous measurements[68] have shown that the end-to-end extension of ssDNA increases with force but decreases with [M^+], as ions better screen adjacent charges along the phosphate backbone and facilitate greater DNA compaction. We saw Δx rise slightly with [M^+] over every titration. Thus, the effect of force prevails, but for most titrations, the overall change in Δx was negligible (<1 nm) and the values of Δx at the lowest and highest cation concentrations were not statistically different.

We also measured trends in ΔG as a function of the Mg^{2+} concentration (Figure 4A) in a background of 40 or 200 mM K^+. Against a 40 mM K^+ background, adding 5 mM Mg^{2+} stabilized the hairpin roughly as much as raising the K^+ concentration to 1 M (Table S6). As was the case for all monovalent titrations, most if not all of the differences in ΔG were ascribable to differences in $F_{1/2}$, and not Δx (Figure 4B and 4C).

In both K^+ backgrounds, the experimental ΔG values increased substantially with [Mg^{2+}] up to 5 mM, but only marginally thereafter (Figure 4A). In addition, the behavior of the two titrations converged at 5 mM Mg^{2+} and above. Similar patterns have been seen in the melting temperatures of DNA duplexes, which at ~10 mM Mg^{2+} reached apparent maxima and became insensitive to monovalent cation concentrations below ~100 mM[62,69,70]. The simplest explanation for our observations would be that the ion atmospheres of the folded and unfolded hairpins become saturated with Mg^{2+} when the bulk [Mg^{2+}] reaches ~5 mM, such that K^+ (whether present at 40 or 200 mM) is essentially excluded, and adding more Mg^{2+} has no effect. However, results from a cation competition experiment using a 24-bp DNA duplex cast doubt on this explanation. In the presence of 5 mM Mg^{2+} and 40 or 200 mM K^+, the numbers of K^+ ions remaining in the atmosphere were significant and differed depending on K^+ concentration[3].

More complex factors may therefore govern the energetics of DNA duplex folding in a mixture of monovalent and divalent cations. An early study using small-angle x-ray scattering (SAXS) revealed no changes in the normalized spatial distributions of monovalent and divalent cations around a DNA duplex as their relative abundances were changed, and the data were reasonably well predicted by PB[20]. However, small changes in the distributions may not have been resolvable with SAXS or PB, which does not account for finite ion size or ion-ion correlations and may inaccurately treat divalent cations near DNA[71-77]. The tightly bound ion theory, which accounts for the factors that PB does not, predicts a tighter association of Mg^{2+} with dsDNA[28,56,78]. This theory, along with

additional experimental tests of the distributions of monovalent and divalent cations around dsDNA and ssDNA, could provide further insight. The diminishing increases we observed in ΔG as $[Mg^{2+}]$ was raised are likely due to the release of fewer Mg^{2+} ions to the bulk solvent upon unfolding, which has been documented previously[69].

Hairpin energy landscape (HEL) model supports a simplified unfolded ensemble

We were motivated to study the process of hairpin unfolding by the expectation that a simpler and more defined unfolded ensemble would exist with the hairpin under force than when free in solution (Figure 1). Treating the unfolded hairpin as a worm-like chain[79], we calculate an average Δx of ~5-8 nm in the absence of force under our conditions[80]. The experimentally determined values of Δx (~17-18 nm) are substantially larger and provide strong evidence that the unfolded state is elongated. To provide an independent test of this conclusion and to better understand our data, we compared them with the energetic predictions from two sources: the empirical *mfold* program[81] and the semi-empirical hairpin energy landscape (HEL) model[35,41].

Mfold provides estimates of the energies needed to break individual base pairs in a DNA helix in solution. These energies scale with the concentrations of monovalent (NaCl) and divalent ($MgCl_2$) salts, and were summed to obtain predictions of ΔG for our hairpin. Because the *mfold* standard energies were established from large experimental datasets of duplex melting obtained in the absence of conformational constraint, we expected *mfold* to overestimate the entropy possessed by the unfolded hairpin in our experiment. A greater entropic gain upon unfolding, stabilizing the unfolded state relative to the folded state, would result in a smaller value of ΔG. Indeed, predictions of ΔG from *mfold* were consistently lower than those measured (Figure 3B, Figure S6).

The HEL model uses *mfold* as a starting point and adds the calculated energy of a hairpin as a function of its molecular end-to-end extension during mechanical unfolding. As each base pair is broken under force, the liberated nucleotides join growing strands of ssDNA on either side of the hairpin, and thereby increase the measured extension. Thus, the HEL model adds to the base-pair *mfold* energy the energy ($\Delta G_{constrain}$) necessary to keep the ends of the hairpin separated. The stretching of ssDNA is modeled by a WLC function[35]. The HEL model predicts that two distinct potential wells occur at low and high extension, corresponding to the fully folded and unfolded states of the hairpin, and that the hairpin undergoes an extension change Δx (Figure 3D) to switch states. The HEL model also predicts values for $F_{1/2}$ (Figure 3C) (by allowing the molecule to occupy these states with equal probability) and ΔG_{HEL} (Figure 3B) (the product of $F_{1/2}$ and Δx).

The HEL model generally captures trends in ΔG more accurately than *mfold* alone (Figure 3B), as it did previously for other hairpin sequences in 200 mM K^+ [35,41] We emphasize that the HEL model produces absolute estimates of ΔG that may be directly compared with measured values, while electrostatic theories such as PB only estimate ΔG_{el} and thus only changes in ΔG between ion concentrations. Here, the HEL model quantitatively reproduced values of ΔG across the full $[M^+]$ range. ΔG_{HEL} most closely matched ΔG for the Na^+ titration (Figure 3B, Figure S6, Table S1), likely because the *mfold* energies underlying the HEL calculations are based on measurements made in Na^+. Nonetheless, the slope of the ΔG_{HEL} trend was equal, within error, to slopes measured for the Li^+, Rb^+, and TMA^+ titrations (Table 1).

The HEL model does not explicitly treat electrostatics. The $[M^+]$-dependence of the calculated ΔG_{HEL} values originates from the corresponding ion dependencies of ΔG_{mfold} and the WLC parameters (see **Experimental Section**). One such parameter, the ssDNA contour length, was estimated using an analytical description[44] of ssDNA stretching in the

presence of monovalent cations, which, to our knowledge, has no analog for divalent cations. Perhaps for this reason, the HEL model underestimated the experimental increases in ΔG upon adding Mg^{2+} (Figure 4A), capturing the leveling-off of ΔG, but not the convergence between the two titrations. However, for the monovalent titrations, the HEL model predicted not only larger ΔG values than *mfold*, but also a steeper dependence of ΔG upon $[M^+]$ (Table 1, **legend**). In addition, the HEL model predicted that Δx would be largely insensitive to $[M^+]$, as measured (Figure 3D). These observations suggest that the HEL model captures the energy necessary to maintain the conformational constraint of the unfolded hairpin against ssDNA compaction as monovalent cations are added.

PB theory recapitulates trends in ΔG for monovalent but not divalent titrations

In parallel with our single-molecule measurements, we calculated electrostatic trends in ΔG versus $[M^+]$ using PB theory. he folded and constrained unfolded states of the hairpin were represented computationally by all-atom models of a B-form helix and elongated ssDNA, respectively, flanked by dsDNA handles (Figure 1). The distance changes modeled for unfolding were very close to the values of Δx measured here and in a previous study using the same hairpin sequence[35], and a sensitivity analysis showed that minor differences in Δx did not significantly affect the goodness of the fits (see **Experimental Section** and Table S2). At each value of $[M^+]$, 'virtual' ion atmospheres were established around the folded and unfolded models according to the PB potential. Changes in the associated free energy of solvation were then calculated, and ΔG_{el} was determined as part of a full thermodynamic cycle (see **Experimental Section** and Figure S4).

The calculated trend in ΔG_{el} was nearly linear in $\log[M^+]$, in agreement with the thermodynamic framework presented above [Equation (2)], and was fitted to the data aggregated across all ionic species (Figure 3A) or to the individual ionic series by varying the offset in energy (Figure 3B, Figure S6). PB does not explicitly consider ionic size, and therefore returns one ΔG_{el} trend for all monovalent cations. The calculated trend was most similar to the K^+ series (Table 1), in agreement with previous studies of folded DNA duplexes[3,24]. In general, there was good agreement between measured and calculated ΔG trends, which reflect changes in the relative stabilities of the folded and unfolded hairpin as the monovalent ion concentration is varied. As previous studies[3,21,23,24] suggest that PB treats the association of monovalent ions with the folded hairpin with reasonable accuracy, our results suggest that PB also correctly treats monovalent ions around the unfolded hairpin.

While PB worked well in reproducing the trends of ΔG_{el} for monovalent titrations, its efficacy was lower in the case of divalent titrations (Figure 4A). PB correctly predicted that the overall change in ΔG across a divalent titration would be larger against a less concentrated K^+ background, but it significantly underestimated the magnitude of this change for both titrations. The PB results for 40 mM K^+ showed ΔG saturating at high $[Mg^{2+}]$ but not increasing as much as measured for the lower Mg^{2+} concentrations, while the changes in ΔG for titrating Mg^{2+} against 200 mM K^+ were insignificant. These results are consistent with previous quantitative failures of PB in capturing the association of Mg^{2+} with DNA duplexes[3,24,27,29,78,82], and may result from ion-ion correlations, which the mean-field treatment in PB ignores. Such correlations allow Mg^{2+} ions to associate closely with folded duplexes in a coordinated manner and to displace monovalent ions to a greater extent than predicted by PB[3,74,77], leading to greater stabilization of the folded hairpin.

DISCUSSION

Nucleic acid folding is enabled by interactions between the molecule itself and the many ions comprising its ion atmosphere. The energies of these interactions govern the shape and

composition of the atmosphere, bias which conformations the molecule adopts, and substantially dictate the folding pathway of the molecule. Such energies have proven difficult to measure directly[18,83,84] but are frequently estimated using electrostatic theories. Several decades ago, Record and co-workers conducted pioneering biochemical studies of DNA duplex melting[54,85-87] and described the energetics of ion-DNA association using a theoretical framework that models the DNA backbone as a linear charge array. More recently, Draper and others have treated the folding of more complex nucleic acids using PB theory[17-19,88,89], where advances in numerical methods have allowed calculations to be performed on arbitrary molecular conformations. These studies have provided valuable insights into nucleic acid folding. Nevertheless, the comparison of experiment and theory often requires the modeling of diverse conformational *ensembles* of unfolded (and partially folded) molecules in solution, which presents significant computational and theoretical challenges. Such diversity limits the use of experimental equilibria as benchmarks for developing and evaluating theoretical treatments of the influence of the ion atmosphere on nucleic acid energetics.

More direct comparisons can be made using simple nucleic acid systems where the range of accessible molecular conformations is limited. The ion-counting study mentioned above[3] focused on short DNA duplexes, which are structurally rigid and do not dissociate under mild conditions. Such studies provide fundamental experimental characterizations of the ion atmosphere around nucleic acids and can provide rich datasets for comparison with theory. Nevertheless, these studies do not report directly on thermodynamic equilibria.

Here, we have returned to the thermodynamics of DNA duplex melting as a function of the ionic composition of the solution, akin to the earlier studies[25-27,29,33,34,53,54,59,62,64,69,70,78,85,87,90,91], but have used force to impose useful conformational constraints on the unfolded hairpin, rendering the modeling of this ensemble more reliable, and facilitating comparisons of experiment and theory.

Our data support conclusions from prior studies that PB theory reasonably captures the energetic effects of monovalent cations on simple nucleic acid equilibria, but not those of divalent cations (refs 24-28,78; see also refs 3,20,21,72). These conclusions thus hold despite potential difficulties in modeling the complex ensemble of unfolded single-stranded states free in solution. Improvements to PB, or alternative theories, will therefore be needed to capture the full energetic role of diffuse, divalent cations in the nucleic acid folding process, as well as to incorporate salient details such as cationic size and ion-ion correlations (see e.g., refs 50,56,73,92, and references therein). To this end, the experimental data presented here should provide benchmarks for the development and evaluation of more advanced electrostatic theories.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This work was funded by graduate fellowships from the National Science Foundation (to P.C.A.), the Agency of Science, Technology, and Research of Singapore (to A.Y.L.S), and the Stanford Bio-X Program (to V.B.C), and by NIH grants GM 49243 (to D.H.) and GM 57035 (to S.M.B.). We further acknowledge use of computational resources at the Bio-X cluster, Stanford University (NSF award CNS-619926) and the National Energy Research Scientific Computing Center (supported by the Office of Science, U.S. Department of Energy under Contract No. DE-AC02-05CH11231).

References

1. Cruz JA, Westhof E. Cell. 2009; 136:604–9. [PubMed: 19239882]
2. Chu VB, Bai Y, Lipfert J, Herschlag D, Doniach S. Curr Opin Chem Biol. 2008; 12:619–25. [PubMed: 19081286]
3. Bai Y, Greenfeld M, Travers KJ, Chu VB, Lipfert J, Doniach S, Herschlag D. J Am Chem Soc. 2007; 129:14981–8. [PubMed: 17990882]
4. Record MT Jr, Anderson CF, Lohman TM. Q Rev Biophys. 1978; 11:103–78. [PubMed: 353875]
5. Garcia-Garcia C, Draper DE. J Mol Biol. 2003; 331:75–88. [PubMed: 12875837]
6. Draper DE. Biophys J. 2008; 95:5489–95. [PubMed: 18835912]
7. Manning GS. Journal of Chemical Physics. 1969; 51:924.
8. Sharp KA, Honig B. Curr Opin Struct Biol. 1995; 5:323–8. [PubMed: 7583630]
9. Honig B, Nicholls A. Science. 1995; 268:1144–9. [PubMed: 7761829]
10. Misra VK, Draper DE. J Mol Biol. 1999; 294:1135–47. [PubMed: 10600372]
11. Misra VK, Sharp KA, Friedman RA, Honig B. J Mol Biol. 1994; 238:245–63. [PubMed: 7512653]
12. Rouzina I, Bloomfield VA. Biophys Chem. 1997; 64:139–55. [PubMed: 17029833]
13. Heilman-Miller SL, Thirumalai D, Woodson SA. J Mol Biol. 2001; 306:1157–66. [PubMed: 11237624]
14. Perez-Salas UA, Rangan P, Krueger S, Briber RM, Thirumalai D, Woodson SA. Biochemistry. 2004; 43:1746–53. [PubMed: 14769052]
15. Russell R, Millett IS, Tate MW, Kwok LW, Nakatani B, Gruner SM, Mochrie SG, Pande V, Doniach S, Herschlag D, Pollack L. Proc Natl Acad Sci U S A. 2002; 99:4266–71. [PubMed: 11929997]
16. Das R, Kwok LW, Millett IS, Bai Y, Mills TT, Jacob J, Maskel GS, Seifert S, Mochrie SG, Thiyagarajan P, Doniach S, Pollack L, Herschlag D. J Mol Biol. 2003; 332:311–9. [PubMed: 12948483]
17. Misra VK, Draper DE. J Mol Biol. 2002; 317:507–21. [PubMed: 11955006]
18. Grilley D, Misra V, Caliskan G, Draper DE. Biochemistry. 2007; 46:10266–78. [PubMed: 17705557]
19. Soto AM, Misra V, Draper DE. Biochemistry. 2007; 46:2973–83. [PubMed: 17315982]
20. Andresen K, Das R, Park HY, Smith H, Kwok LW, Lamb JS, Kirkland EJ, Herschlag D, Finkelstein KD, Pollack L. Phys Rev Lett. 2004; 93:248103. [PubMed: 15697865]
21. Das R, Mills TT, Kwok LW, Maskel GS, Millett IS, Doniach S, Finkelstein KD, Herschlag D, Pollack L. Phys Rev Lett. 2003; 90:188103. [PubMed: 12786045]
22. Pabit SA, Qiu X, Lamb JS, Li L, Meisburger SP, Pollack L. Nucleic Acids Res. 2009; 37:3887–96. [PubMed: 19395592]
23. Pabit SA, Meisburger SP, Li L, Blose JM, Jones CD, Pollack L. J Am Chem Soc. 2010; 132:16334–6. [PubMed: 21047071]
24. Bai Y, Chu VB, Lipfert J, Pande VS, Herschlag D, Doniach S. J Am Chem Soc. 2008; 130:12334–41. [PubMed: 18722445]
25. Bond JP, Anderson CF, Record MT Jr. Biophys J. 1994; 67:825–36. [PubMed: 7948695]
26. Shkel IA, Record MT Jr. Biochemistry. 2004; 43:7090–101. [PubMed: 15170346]
27. Tan ZJ, Chen SJ. Biophys J. 2006; 90:1175–90. [PubMed: 16299077]
28. Tan ZJ, Chen SJ. Biophys J. 2007; 92:3615–32. [PubMed: 17325014]
29. Korolev N, Lyubartsev AP, Nordenskiold L. Biophys J. 1998; 75:3041–56. [PubMed: 9826624]
30. Goddard NL, Bonnet G, Krichevsky O, Libchaber A. Phys Rev Lett. 2000; 85:2400–3. [PubMed: 10978020]
31. Jung J, Van Orden A. J Am Chem Soc. 2006; 128:1240–9. [PubMed: 16433541]
32. Ma H, Proctor DJ, Kierzek E, Kierzek R, Bevilacqua PC, Gruebele M. J Am Chem Soc. 2006; 128:1523–30. [PubMed: 16448122]
33. Korolev N, Lyubartsev AP, Nordenskiold L. J Biomol Struct Dyn. 2002; 20:275–90. [PubMed: 12354079]

34. MacGillivray AD, McMullen AI. J Theor Biol. 1968; 19:159–68. [PubMed: 5726566]
35. Woodside MT, Behnke-Parks WM, Larizadeh K, Travers K, Herschlag D, Block SM. Proc Natl Acad Sci U S A. 2006; 103:6190–5. [PubMed: 16606839]
36. Neuman KC, Block SM. Rev Sci Instrum. 2004; 75:2787–809. [PubMed: 16878180]
37. Marko JF, Siggia ED. Macromolecules. 1995; 28:8759–8770.
38. Bustamante C, Marko JF, Siggia ED, Smith S. Science. 1994; 265:1599–1600. [PubMed: 8079175]
39. Greenleaf WJ, Woodside MT, Abbondanzieri EA, Block SM. Phys Rev Lett. 2005; 95:208102. [PubMed: 16384102]
40. Svoboda K, Block SM. Cell. 1994; 77:773–84. [PubMed: 8205624]
41. Woodside MT, Anthony PC, Behnke-Parks WM, Larizadeh K, Herschlag D, Block SM. Science. 2006; 314:1001–4. [PubMed: 17095702]
42. Huguet JM, Bizarro CV, Forns N, Smith SB, Bustamante C, Ritort F. Proc Natl Acad Sci U S A. 2010; 107:15431–6. [PubMed: 20716688]
43. Dessinges MN, Maier B, Zhang Y, Peliti M, Bensimon D, Croquette V. Phys Rev Lett. 2002; 89:248102. [PubMed: 12484983]
44. Cocco S, Marko JF, Monasson R, Sarkar A, Yan J. Eur Phys J E Soft Matter. 2003; 10:249–63. [PubMed: 15015107]
45. Macke TJ, Case DA. Molecular Modeling of Nucleic Acids. 1998; 682:379–393.
46. Ohtaki H, Radnai T. Chemical Reviews. 1993; 93:1157–1204.
47. Baker NA, Sept D, Joseph S, Holst MJ, McCammon JA. Proceedings of the National Academy of Sciences of the United States of America. 2001; 98:10037–10041. [PubMed: 11517324]
48. Dolinsky TJ, Czodrowski P, Li H, Nielsen JE, Jensen JH, Klebe G, Baker NA. Nucleic Acids Res. 2007; 35:W522–5. [PubMed: 17488841]
49. Dolinsky TJ, Nielsen JE, McCammon JA, Baker NA. Nucleic Acids Res. 2004; 32:W665–7. [PubMed: 15215472]
50. Chu VB, Bai Y, Lipfert J, Herschlag D, Doniach S. Biophys J. 2007; 93:3202–9. [PubMed: 17604318]
51. Frankkamenetskii MD, Anshelevich VV, Lukashin AV. Uspekhi Fizicheskikh Nauk. 1987; 151:595–618.
52. Kotin L. J Mol Biol. 1963; 7:309–11. [PubMed: 14065313]
53. Schildkraut C. Biopolymers. 1965; 3:195–208. [PubMed: 5889540]
54. Record MT Jr. Biopolymers. 1967; 5:975–92. [PubMed: 6078884]
55. Baldwin RL. Accounts of Chemical Research. 1971; 4:265.
56. Tan ZJ, Chen SJ. Methods Enzymol. 2009; 469:465–87. [PubMed: 20946803]
57. Leipply D, Lambert D, Draper DE. Methods Enzymol. 2009; 469:433–63. [PubMed: 20946802]
58. Record MT Jr, Zhang W, Anderson CF. Adv Protein Chem. 1998; 51:281–353. [PubMed: 9615173]
59. Rouzina I, Bloomfield VA. Biophys J. 1999; 77:3242–51. [PubMed: 10585946]
60. Bleam ML, Anderson CF, Record MT. Proc Natl Acad Sci U S A. 1980; 77:3085–9. [PubMed: 16592827]
61. Collins JM, Rogers KS. Chem Biol Interact. 1977; 19:197–203. [PubMed: 589700]
62. Nakano S, Fujimoto M, Hara H, Sugimoto N. Nucleic Acids Res. 1999; 27:2957–65. [PubMed: 10390539]
63. Orosz JM, Wetmur JG. Biopolymers. 1977; 16:1183–99. [PubMed: 880349]
64. Stellwagen E, Muse JM, Stellwagen NC. Biochemistry. 2011; 50:3084–94. [PubMed: 21410141]
65. Trend BL, Knoll DA, Ueno M, Evans DF, Bloomfield VA. Biophys J. 1990; 57:829–34. [PubMed: 2344467]
66. Vieregg J, Cheng W, Bustamante C, Tinoco I Jr. J Am Chem Soc. 2007; 129:14966–73. [PubMed: 17997555]
67. Lambert D, Leipply D, Shiman R, Draper DE. J Mol Biol. 2009; 390:791–804. [PubMed: 19427322]

68. Smith SB, Cui Y, Bustamante C. Science. 1996; 271:795–9. [PubMed: 8628994]
69. Owczarzy R, Moreira BG, You Y, Behlke MA, Walder JA. Biochemistry. 2008; 47:5336–53. [PubMed: 18422348]
70. Williams AP, Longfellow CE, Freier SM, Kierzek R, Turner DH. Biochemistry. 1989; 28:4283–91. [PubMed: 2765487]
71. Anderson CF, Record MT Jr. Annu Rev Biophys Biophys Chem. 1990; 19:423–65. [PubMed: 2194482]
72. Granot J, Kearns DR. Biopolymers. 1982; 21:219–32. [PubMed: 7055633]
73. Grochowski P, Trylska J. Biopolymers. 2008; 89:93–113. [PubMed: 17969016]
74. Grosberg AY, Nguyen TT, Shklovskii BI. Reviews of Modern Physics. 2002; 74:329–345.
75. Ni HH, Anderson CF, Record MT. Journal of Physical Chemistry B. 1999; 103:3489–3504.
76. Pack GR, Wong L, Lamm G. Biopolymers. 1999; 49:575–90. [PubMed: 10226502]
77. Shklovskii BI. Physical Review E. 1999; 60:5802–5811.
78. Tan ZJ, Chen SJ. Biophys J. 2008; 95:738–52. [PubMed: 18424500]
79. Murphy MC, Rasnik I, Cheng W, Lohman TM, Ha TJ. Biophysical Journal. 2004; 86:2530–2537. [PubMed: 15041689]
80. This calculation assumes WLC parameters of D_p = 0.6 nm/nt and L_p = 1-1.5 nm, and a width of 2.0 nm for the folded hairpin (Woodside et al. PNAS 2006). Any formation of secondary structure within the unfolded hairpin would only serve to lower Δx. Even at very low monovalent ion concentrations (~10 mM), which cause substantial stiffening of ssDNA (L_p ~ 3 nm; Murphy et al. Biophys. J. 2004), an average Δx of only ~13 nm is expected.
81. Zuker M. Nucleic Acids Res. 2003; 31:3406–15. [PubMed: 12824337]
82. Wensel TG, Meares CF, Vlachy V, Matthew JB. Proc Natl Acad Sci U S A. 1986; 83:3267–71. [PubMed: 3458180]
83. Grilley D, Soto AM, Draper DE. Proc Natl Acad Sci U S A. 2006; 103:14003–8. [PubMed: 16966612]
84. Rialdi G, Levy J, Biltonen R. Biochemistry. 1972; 11:2472–9. [PubMed: 4625190]
85. Record MT. Biopolymers. 1975; 14:2137–2158.
86. Record MT, Lohman TM. Biopolymers. 1978; 17:159–166.
87. Record MT Jr, Woodbury CP, Lohman TM. Biopolymers. 1976; 15:893–915. [PubMed: 4169]
88. Bokinsky G, Rueda D, Misra VK, Rhodes MM, Gordus A, Babcock HP, Walter NG, Zhuang X. Proc Natl Acad Sci U S A. 2003; 100:9302–7. [PubMed: 12869691]
89. Misra VK, Shiman R, Draper DE. Biopolymers. 2003; 69:118–36. [PubMed: 12717727]
90. Dove WF, Davidson N. Journal of Molecular Biology. 1962; 5:467.
91. Privalov PL, Ptitsyn OB, Birshtei Tm. Biopolymers. 1969; 8:559.
92. Chen AA, Marucho M, Baker NA, Pappu RV. Methods Enzymol. 2009; 469:411–32. [PubMed: 20946801]

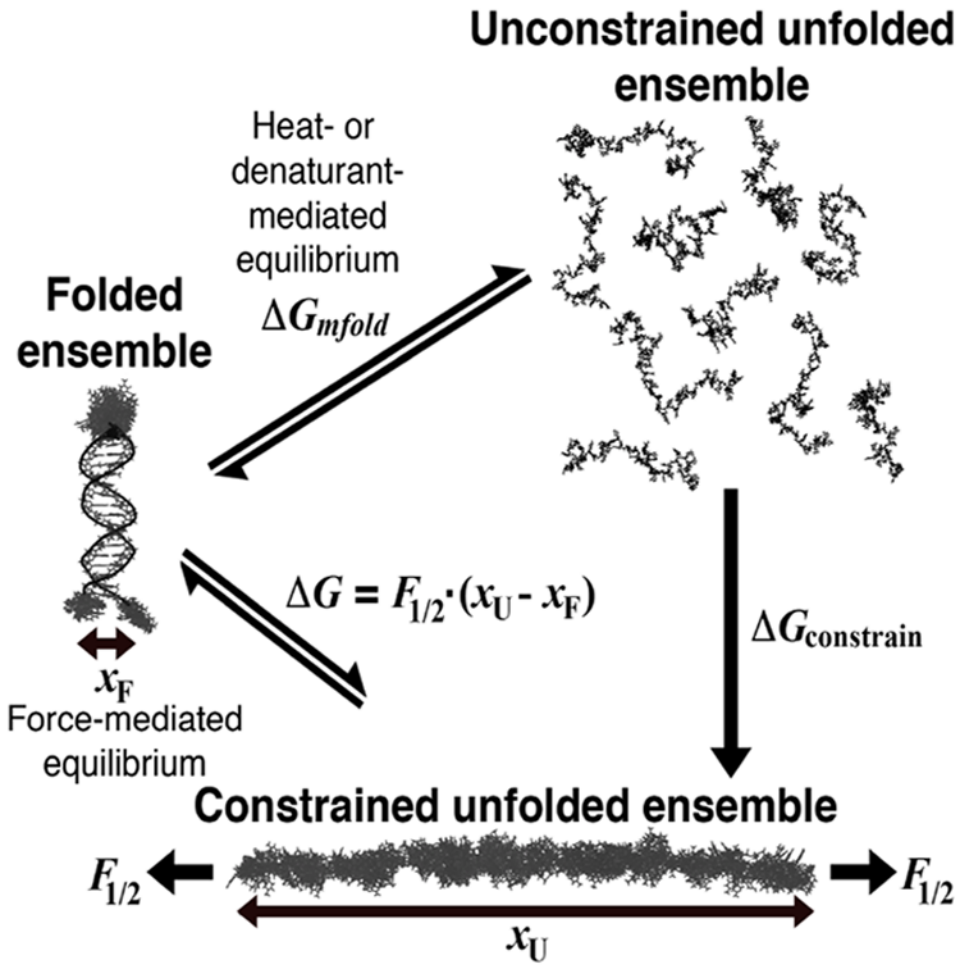

Figure 1. Illustration of hairpin folding in the presence or absence of conformational constraint

PB calculations were performed on ensembles of folded (left) and constrained unfolded (bottom) DNA hairpin structures representing the endpoints of folding in our single-molecule optical trapping assay. Positions of the nucleotides connecting the hairpin to the handles, in the folded hairpin loop, and in the entire unfolded hairpin were randomized by molecular dynamics. See text for explanations of variables.

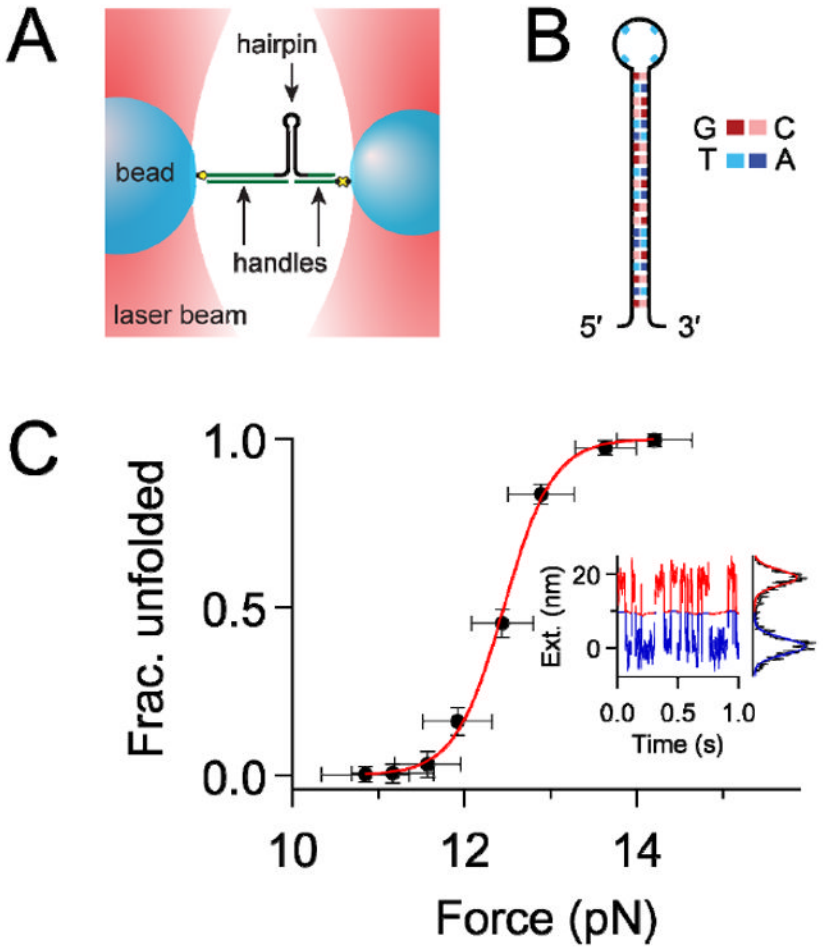

Figure 2. Set-up of optical trapping experiments and sample data

(A) Dumbbell single-molecule optical-trapping assay (not to scale). The DNA hairpin (black) was attached to two polystyrene beads (blue) through dsDNA handles (green) and specific chemical linkages (yellow). Each bead was held in a separate optical trap (red), and force was exerted on the hairpin by moving the traps apart. **(B)** The hairpin sequence used throughout this work. **(C)** Representative single-molecule data for a hairpin measured in 40 mM Na^+. $F_{1/2}$ was determined as the inflection point of a Boltzmann fit to the fraction of time in the unfolded state versus force. *Inset*, 1 s excerpt of the 10 s extension record corresponding to the point nearest to $F_{1/2}$. The record was partitioned into the unfolded (red) and folded (blue) states, and a Gaussian function was fit to the histogram of points on each side of the partition. Δx (~18 nm) was determined as the distance between the peaks of the Gaussians.

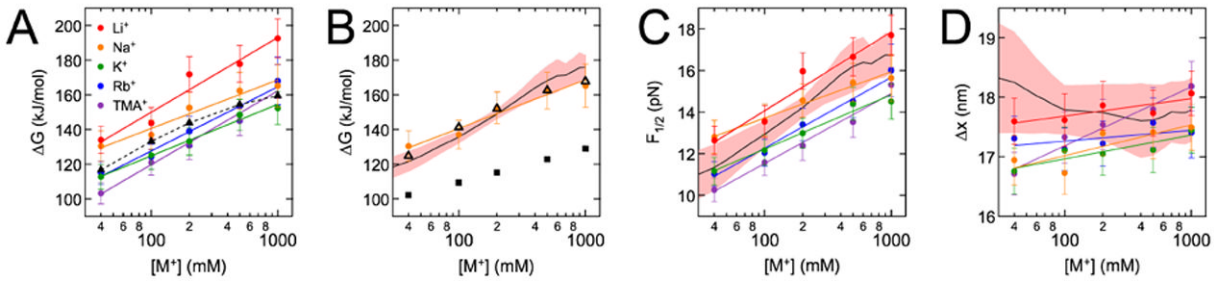

Figure 3. Monovalent titrations

ΔG **(A, B)**, $F_{1/2}$ **(C)**, and Δx **(D)** were measured (circles) for hairpin molecules in the presence of one of five monovalent ions (see legend). Ionic radii (in Å) are 3.6, 3.8, 4.2, 4.6 and 5.9 for Li^+, Na^+, K^+, Rb^+ and TMA^+, respectively. (*N.B*: here we consider the radius inclusive of the first hydration shell[46], since most ions in the ion atmosphere remain hydrated[6].) Experimental error bars represent total uncertainty, i.e., the standard error on the mean of each measurement added to estimates of systematic uncertainty. Linear fits to the data are shown in color, with $\Delta G_{\text{non-el}}$ treated as an adjustable parameter. Values of ΔG_{el} calculated with the PB equation (triangles) were fitted by χ^2 minimization to all five data sets pooled together (A) or to the Na^+ data set only (B). Uncertainties in the PB energies are smaller than the heights of the symbols. Predictions of experimental parameters were also calculated using the HEL model (black trace; uncertainty envelope in pink) and *mfold* (black squares, 5% uncertainties not shown).

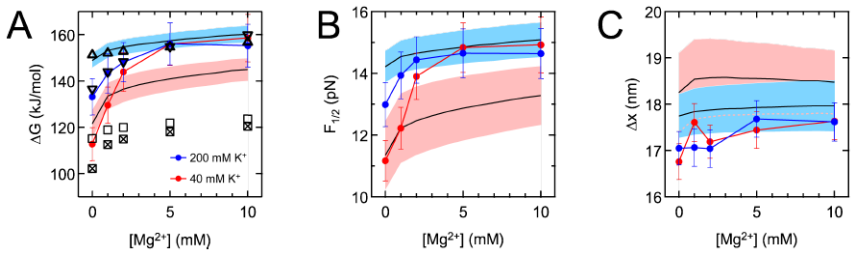

Figure 4. Divalent titrations

ΔG and ΔG_{el} **(A)**, $F_{1/2}$ **(B)**, and Δx **(C)** were measured (circles) and calculated for titrations of Mg^{2+} against K^+ backgrounds. Experimental error bars are as reported in Figure 3. Data points are connected to guide the eye. PB energies (downward and upward-pointing triangles for 40 and 200 mM K^+, respectively) were fitted by χ^2 minimization to the experimental energies at 5 and 10 mM Mg^{2+} in each titration. Predictions from the HEL model and *mfold* are depicted by black traces (uncertainty envelopes for 40 and 200 mM K^+ in pink and blue) and squares (crossed and open for 40 and 200 mM K^+), respectively. Values of Δx (C) predicted by the HEL model overlapped substantially for the 40 and 200 mM K^+ series; a pink dotted line denotes the bottom of the 40 mM uncertainty envelope.

Table 1

Summary of fitting parameters for Δ*G* trends obtained from experimental monovalent titrations.[a]

Ion	$\mathbf{m_{data}}$	χ^2_{HEL}	χ^2_{PB}	χ^2_{mfold}
Li^+	43 ± 9	1.4	3.3	9.1
Na^+	28 ± 9	2.2	0.75	1.4
K^+	30 ± 7	2.2	0.46	2.2
Rb^+	37 ± 8	0.60	1.5	5.2
TMA^+	43 ± 7	1.5	4.0	11.9

[a] A line was fitted to the plot of experimental Δ*G* values versus log[M^+] for each titration (Figure 3A) and the slope (**$\mathbf{m_{data}}$**, in kJ/mol/log(mM)) was extracted. Similar linear fits to unfolding energies calculated from HEL, PB and *mfold* yielded slopes of 41 ± 3, 32 ± 1, and 20 ± 5, respectively. The calculated unfolding energies were also fitted to each experimental Δ*G* series using χ^2 minimization, yielding the final χ^2 values (with four degrees of freedom) shown.