Published in final edited form as:

J Mol Biol. 2009 July 24; 390(4): 805–819. doi:10.1016/j.jmb.2009.05.071.

Molecular simulation studies of monovalent counterion-mediated interactions in a model RNA kissing loop

Alan A. Chen[†], David E. Draper[‡], and Rohit V. Pappu^{§,*}

[†]Molecular Biophysics Program and Center for Computational Biology Washington University School of Medicine St. Louis, MO 63130

[‡]Department of Chemistry Johns Hopkins University 3400 North Charles Street Baltimore, MD 21218

[§]Department of Biomedical Engineering and Center for Computational Biology Washington University in St. Louis One Brookings Drive, Box 1097 St. Louis, MO 63130

Summary

A kissing loop is a highly stable complex formed by loop-loop base pairing between two RNA hairpins. This common structural motif is utilized in a wide variety of RNA mediated processes, including antisense recognition, substrate recognition in riboswitches, and viral replication. Recent work has shown that the Tar-Tar* complex, an archetypal kissing loop, can form without Mg²⁺, so long as high concentrations of alkali chloride salts are present. Interestingly, the stability of the complex is found to decrease with increasing cation size. In this work, we use molecular simulations to develop a molecular-level understanding of the origins of the observed counterion specificity. The ionic atmosphere of the Tar-Tar* complex was examined in the presence of 800mM NaCl, KCl, or CsCl. We used spatial free energy density profiles to analyze the differences in counterion accumulation at different spatial extents from the RNA molecule. We find that the lowest free energy levels, which are situated in the vicinity of the loop-loop interface can accommodate roughly two counterions, irrespective of counterion type. However, as we move into higher free energy levels, which are farther from the loop-loop interface, we observe increased differences in the numbers of accumulated counterions, with this number being largest for Na⁺ and smallest for Cs⁺. We analyzed the source of these differences and were able to attribute these to two distinct features: The extent of partial dehydration varies based on cation type and the smaller the cation, the greater the degree of dehydration. While smaller ions bind their first hydration shell water molecules more tightly than larger ions, they are also able to shed these water molecules for stronger electrostatic interactions with the RNA molecule. Secondly, we observed a distinct asymmetry in the numbers of accumulated cations around each hairpin in the Tar-Tar* complex. We were able to ascribe this asymmetry to the presence of a guanine-tract in the Tar hairpin, which facilitates partial dehydration of the counterions. However, the smaller ions compensate for this asymmetry by forming a belt around the loop-loop interface in the intermediate free energy levels. As a result, the degree of asymmetry in counterion accumulation around individual hairpins shows inverse correlation with the experimentally observed cation specificity for the stability of Tar-Tar* i.e., the smaller the asymmetry, the greater the experimentally observed stability. This in turn provides a plausible explanation for why the smaller cations help stabilize the Tar-Tar* complex better than the larger cations. These findings suggest that the specific sequence and structural features of the Tar-Tar* complex may be the source of the experimental observations regarding cation specificity in Tar-Tar* stability. Our

*Corresponding author: pappu@wustl.edu.

results lead to testable predictions for how changes in sequence might alter the observed counterion specificity in kissing loop stability.

Keywords

RNA-ion interactions; kissing loops; ion specificity; molecular simulation

Introduction

Ions play a crucial role in RNA stability and folding^{1; 2}. RNA molecules that can form stable secondary and tertiary structures in the absence of Mg^{2+} represent a useful class of molecules for biophysical studies because one can interrogate the role of ion-mediated interactions without the confounding effects of competition between mono and divalent cations. In an accompanying paper, Lambert *et al.* present their analysis of the variations in stability of different RNA motifs in the presence of chloride salts of different Group I cations (ranging from Li^+ to Cs^+). Of the different RNA systems studied by Lambert *et al.*, the Tar-Tar* kissing loop complex is interesting since its stability shows striking dependence on cation size. Also, this system is small enough to be tractable for detailed molecular dynamics simulations.

The formation of kissing loop complexes is crucial to many RNA mediated processes, including antisense recognition³, substrate recognition in riboswitches⁴, viral replication⁵, and retroviral genome dimerization^{6; 7}. Tar-Tar* is an archetypal kissing-loop complex and is composed of a pair of sixteen nucleotide hairpins, each containing a five base pair stem and a six nucleotide loop as shown in Figure 1. The two loops have complementary sequences that stabilize the complex via intermolecular Watson-Crick base pairs. This results in substantial deformation of the RNA backbone, with phosphorus atoms from opposing strands (C6 and U22) approaching within 5.25 Å (Figure 1a). This distance of approach is small in comparison to the distance of 10.4 Å, which is the smallest possible inter-strand separation between phosphorus atoms in canonical A-form duplexes. The high density of negatively charged phosphate groups at the loop-loop interface led to the hypothesis that this interface acts as a cation-binding pocket⁸. Partial support for this hypothesis comes from the observation that cations can localize to this pocket in molecular dynamics simulations of a similar kissing loop complex⁹.

Lambert *et al.* have measured the stability of the Tar-Tar* complex in a series of alkali-chloride salts using isothermal titration and thermal melt experiments. Stability estimates that result from these measurements are shown in Table 1. The stability of the Tar-Tar* complex is inversely proportional to the crystallographic radius of the counterion. Conversely, folding of the isolated Tar* hairpin shows only marginal dependence on counterion type.

Why do the experimental data indicate monovalent cation sensitivity in Tar-Tar* stability? In the classical picture, the size of an ion determines its charge density, its hydration free energy, its pairing propensity with counterions, and the standard heats of solution of crystalline salts¹⁰. Therefore, the origins of the observed cation specificity in Tar-Tar* stability must lie in the balance of counter and coion hydration and the accumulation / depletion of these species around the RNA. To investigate the details of how this balance leads to the experimentally observed cation specificities, we have used all-atom molecular dynamics simulations of the Tar-Tar* system with explicit representation of solvent molecules and ionic species. Results from analysis of our simulation data are presented next followed by a discussion that places our results in the context of experimental data.

Results

Lambert et al. have shown that the folding of isolated hairpins shows negligible dependence on the identity of the monovalent counterion in solution (see Table 1). Therefore, we reasoned that the dependence of Tar-Tar* stability on the identity of monovalent counterions must originate in counterion-specific interactions with the Tar-Tar* complex. Accordingly, in our simulations we focused on quantifying differences in the ionic atmospheres between Na⁺, K⁺, and Cs⁺ around the Tar-Tar* complex. Such simulations are computationally tractable when compared to simulations of folding or binding because we do not need to capture large-scale conformational changes if our focus is on details of the ionic atmosphere around an average structure. However, these tractable simulations are also restrictive by construction because they do not allow direct comparison between experimental and simulation data; rather, we can only ask if the simulated ionic atmospheres can help rationalize the experimentally observed trends regarding the dependence of Tar-Tar* stability on the identity of monovalent counterions. With this caveat in mind, we proceed to describe the results of our analysis of ionic atmospheres around Tar-Tar* in 800 mM solutions of NaCl, KCl, and CsCl, respectively. The choice of salt concentration in our study merits discussion. Finite size artifacts – where small numbers of ions are used to simulate a target concentration – can pose a serious challenge for obtaining convergent statistics regarding ionic atmospheres. To avoid this, we need to use large simulation cells to mimic finite salt concentrations. We find that we need $O(10^3)$ ions in the simulation cell to mimic a target salt concentration (see Appendix 1). For concentrations lower than 800 mM, this would require box sizes that are substantially larger than the 120 Å per side that we use in this work. Additionally, there are two biophysical features that justify our choice of salt concentration. The Tar-Tar* complex is stabilized in higher concentrations of monovalent salt and the accompanying work of Lambert et al shows a roughly constant difference in stability for the three counterions in the 200-800 mM excess salt range. Additionally, Lambert et al. have provided us with data for the counterion dependence of the stability of Tar-Tar* in the presence of 804 mM excess salt. These numbers are also shown in Table 1.

Spatial distribution functions

Within the timescale of our simulations we did not see any dissociation of the Tar-Tar* complex. This is consistent with data from surface plasmon resonance studies¹¹ that estimate the dissociation rate for the Tar-Tar* complex to be $2.6 \times 10^{-3} \text{ s}^{-1}$ in 150 mM NaCl. To quantify the ionic atmospheres around the Tar-Tar* complex we calculated three dimensional spatial distribution functions for counterions. In order to calculate spatial distribution functions, a consistent frame of reference must be used. In the Tar-Tar* complex, the loop-loop interface is rigidly formed while the two stems show discernable fluctuations on the simulation timescales. The nucleotides at the rigid loop-loop interface are therefore ideally suited for use as a frame of reference. Successive frames were rotated and translated to minimize the root-mean-square-deviation (RMSD) of the loop-loop backbone phosphates. A 1 Å³ cubic grid was then overlaid and the average density of cations in each cell was calculated. Only the location of the center of each cation was considered, such that each cation could at most occupy one cell. The spatial distribution functions were converted to spatial free energy density profiles using the relation in equation 1 shown below:

$$\Delta G(x, y, z) = -RT \ln \left[\frac{\rho(x, y, z)}{\rho_0} \right] \quad (1)$$

Here, $\rho(x, y, z)$ is the density within each cell, ρ_0 is the bulk ion density as measured in the periphery of the simulation cell, $R = 1.987 \times 10^{-3} \text{ kcal/mol-K}$ and $T = 298 \text{ K}$. The resulting

free energy contours provide a quantitative assessment of the spatial variations of the strengths of ion interactions with Tar-Tar*. Isocontours of the spatial free energy density profiles at -0.5 , -1.0 , -1.5 , and -2.0 kcal/mol were then calculated and visualized using the VMD software package¹². The cumulative ion numbers within each cutoff were calculated by summation of $\rho(x,y,z)$ over all the cells that result in a ΔG value above a certain cutoff value in the spatial free energy density profiles. The relationship shown in equation (1) is based on the fact that in an equilibrium simulation, the distributions of ions obey Boltzmann statistics, and this is true for ions in local and bulk partitions. The ratio within the square brackets of equation (1) quantifies the relative probability – vis-à-vis bulk – of realizing a specific ion density within a particular spatial envelope around the RNA. If the ratio is unity, then the logarithm of the ratio is zero. This sets the bulk milieu as the reference state. Deviations from unity will be either due to ion depletion or accumulation and the quantity shown in equation (1) quantifies the free energy change associated with depleting or accumulating ions vis-à-vis the bulk.

Unlike Poisson-Boltzmann (PB) calculations¹³ where one typically uses grid sizes of $0.5\text{\AA}^3$ to analyze ion occupancy statistics, the chosen grid size of $1.0\text{\AA}^3$ accounts for the finite sizes of ions. In our analysis, there is an inverse trade-off between grid-size and the statistical errors associated with ion occupancies calculated for each grid. Since the error grows with diminishing grid size, especially when the grid size becomes significantly smaller than the size of a single ion, we used $1.0\text{\AA}^3$ grids as an optimal choice for balancing two considerations namely, high spatial resolution and low statistical error. It is worth noting that the crystallographic diameter of the smallest counterion studied, Na^+ , is $\sim 2.3\text{\AA}$, and therefore a $1.0\text{\AA}^3$ grid is accurate for the ions included in the simulation.

Figure 2a shows the calculated spatial free energy density profiles for the accumulation of Na^+ , K^+ , and Cs^+ around Tar-Tar*. To facilitate our analysis, the free energy spectra for the ionic atmospheres were divided into discrete intervals as shown by the color bar accompanying Figure 2b. For any given free energy level, both the extents and locations of ion accumulation around Tar-Tar* vary with counterion type. Figure 3 shows the total number of ions that accumulate within an energy interval. These numbers are computed as cumulative distribution functions by integrating over the free energy contours. Ion accumulation around Tar-Tar* follows the trend of $cc(\text{Na}^+) > cc(\text{K}^+) > cc(\text{Cs}^+)$ for all free energy levels. Here, cc refers to the cumulative number of counterions within a free energy level and is the ordinate in Figure 3. Within the most negative free energy contour ($\Delta G \leq -2$ kcal/mol), $cc(\text{Na}^+) = 2.3 \pm 0.2$, $cc(\text{K}^+) = 2.0 \pm 0.1$, and $cc(\text{Cs}^+) = 1.7 \pm 0.1$. Although the differences are small, they are statistically significant. Differences in the cumulative numbers of accumulated counterions increase as the magnitude of the free energies decrease as shown in Figure 3.

Figure 2b focuses on specific differences between different ionic atmospheres. The deep pink contours correspond to free energy levels $\Delta G \leq -2$ kcal/mol. The red arrows point to specific regions within this interval that are present when Na^+ accumulates around Tar-Tar* and are absent for K^+ and Cs^+ . In Figure 2 a-c, the free energy interval $-2 \text{ kcal/mol} < \Delta G \leq -1.0 \text{ kcal/mol}$ is depicted as orange contours. There are clear differences in the spatial extents of ion accumulation within this interval with the extents being largest for Na^+ and smallest for Cs^+ . There are more regions for the favorable accumulation of Na^+ when compared to K^+ and Cs^+ within this intermediate free energy interval. As we progress through the free energy intervals, the increased accumulation of Na^+ is manifest in the form of a “belt” that forms an envelope around the loop-loop interface. This belt, which is highlighted by the green arrows in Figure 2b, is only weakly present for the accumulation of K^+ and Cs^+ around Tar-Tar*. An important feature of ion accumulation around Tar-Tar* is the continuous variation in the cumulative number of accumulated counterions as a function

of free energy relative to the bulk (Figure 3). Differences in accumulation levels between different ions exist across all free energy levels, not just the most negative free energy levels.

Counterion Hydration Analysis

To understand why Na^+ , K^+ , and Cs^+ exhibit different levels of accumulation around Tar-Tar* we analyzed the hydration characteristics of the three ions around the RNA and compared this to ion hydration statistics in bulk solution. In each trajectory frame, counterions were classified as belonging to the most negative free energy interval ($\Delta G \leq -1.5$ kcal/mol), the intermediate free energy interval (-1.5 kcal/mol $< \Delta G \leq -0.5$ kcal/mol), or bulk based on the spatial free density profiles shown in Figures 2 a-c. Note that these intervals effectively combine the four free energy isocontours displayed in Figure 2 into two larger intervals, a simplification made necessary by the highly overlapping volumes of adjacent free energy isocontours. Hydration statistics were obtained by counting the number of oxygen atoms of water molecules found within a cutoff radius corresponding to the first minimum of the cation-water oxygen radial distribution function. The cation-water distance cutoffs used were 3.2, 3.6, and 4.0 Å for Na^+ , K^+ , and Cs^+ , respectively.

The bottom row in Figure 4 shows the hydration statistics (populations of ions with different numbers of first hydration shell water molecules) for each of the three counterions in bulk solution. These statistics are shown as pie charts and indicate significant non-idealities that are consistent with the high concentration (800 mM) of ions in the bulk. Previous work¹⁴ showed that these non-idealities match the estimates from conductometric measurements¹⁵ and arise due to the presence of ion pairs and clusters that are in dynamic equilibrium with dissociated, fully hydrated ions. As the bulk concentrations of ions decrease, the hydration statistics match those of dilute / ideal solutions¹⁶. Population of ion pairs and clusters results from partial dehydration of ions and it is important to recognize that for a given concentration, the extent of ion pairing and clustering varies with ion type. This is because ion size controls its charge density, which in turn dictates the free energy of hydration and the cost of partial or complete desolvation of an ion. Experimental data¹⁰ and simulations¹⁴ indicate that the smaller ions that bind their water molecules most tightly are also capable of giving up their water molecules in exchange for stronger interactions between oppositely charged species.

We now turn from hydration of ions in bulk solutions to the hydration of cations around Tar-Tar*, as summarized in figure 4. Each column shows the hydration statistics for a specific cation. Pie charts on the top row show hydration statistics for the lowest free energy intervals ($\Delta G \leq -1.5$ kcal/mol) while the pie charts in the middle row show hydration statistics for the intermediate free energy intervals (-1.5 kcal/mol $< \Delta G \leq -0.5$ kcal/mol). Hydration statistics for the lowest free energy intervals ($\Delta G > -0.5$ kcal/mol) are identical to those for the bulk electrolyte, which are shown in row 3 of figure 4.

We used a figure of merit to quantify the extent to which accumulated ions differ from their behavior in bulk solutions. This figure of merit is defined as: $P_{j,k}^{(i)}$

$$F_{\text{non-idealities}(k)}^{(i)} = 100 \sqrt{\sum_{j=2}^{12} [P_{j,k}^{(i)} - P_{j,\text{bulk}}^{(i)}]^2} \quad (2)$$

Here, i refers to the cation type (Na^+ , K^+ , or Cs^+), j denotes the number of first shell water molecules around the ion in a snapshot, and k denotes a specific free energy interval; $P_{j,k}^{(i)}$ is

the fractional population of ion i in hydration state j and free energy interval k such that

$\sum_{j=2}^{12} p_{j,k}^{(i)} = 1$; $p_{j,\text{bulk}}^{(i)}$ is the distribution of hydration states in the bulk. $F_{\text{non-idealities}(k)}^{(i)} \approx 0$ if the hydration statistics for ion i within free energy interval k are similar to that of the bulk

solution. Larger values of $F_{\text{non-idealities}(k)}^{(i)}$ correspond to larger deviations of hydration statistics of accumulated ions in free energy interval k from the bulk solution.

Computed values for $F_{\text{non-idealities}(k)}^{(i)}$ are shown in Figure 5. Hydration statistics for counterions that accumulate in the most negative free energy regions ($\Delta G \leq -1.5$ kcal/mol) show the largest deviations from bulk hydration statistics. Differences in the spatial extents and cumulative numbers of accumulated ions in the range $\Delta G \leq -1.5$ kcal/mol may be rationalized in terms of the differences in hydration statistics that result from interactions of the ions with the RNA. Figure 4 shows that there is an increased likelihood for Na^+ to access tri-, tetra- and penta-hydrate states in the most negative free energy interval. The likelihoods of populating these states decrease systematically with increasing ion size. The smaller, higher charge density ions that “bind” their first hydration shell waters more tightly are also capable of shedding these waters in exchange for stronger interactions with the macroion. In the intermediate free energy levels ($-1.5 < \Delta G \leq -0.5$ kcal/mol), the ions become more bulk-like. Hence, in these intervals, the differences in ion accumulation must originate in the generic differences between the ions. These differences are also present in the bulk solution, and are amplified at high concentrations.

Counterions within the most negative free energy levels ($\Delta G \leq -1.5$ kcal/mol) have at least one inner-sphere contact with either an O6 or N7 atom from a purine base, regardless of counterion type. Counterions in the intermediate free energy intervals (-1.0 kcal/mol $< \Delta G \leq -0.5$ kcal/mol) are always associated with backbone phosphate oxygen atoms, either as solvent-separated, fully hydrated ion pairs or as partially hydrated contact ion pairs. For Na^+ around Tar-Tar* we also observe a low free energy region at the junction between the two hairpin backbones (U22 & C6) because the phosphate oxygen atoms at this site are able to partially dehydrate the ion. The “belt” of intermediate to high free energy Na^+ ions that straddles the loop-loop interface is comprised of bulk-like hydrated ions. This is because there are no proximal RNA atoms in this region that can facilitate partial dehydration. The preceding analysis emphasizes the fact that ions use the entire spectrum of hydration states as they accumulate around the macroion. The accumulation of partially hydrated ions suggests that the interplay between desolvation and interaction with the macroion is rather subtle and may be difficult to capture using mean field models.

Asymmetry in counterion accumulation

Figures 2 a-c indicate an asymmetry in the accumulation of ions around the Tar and Tar* hairpins in the complex. To quantify this difference in counterion accumulation between hairpins, we computed a “charge asymmetry ratio” for each counterion type. This quantity was estimated using the spatial free energy density profiles in Figure 2a with the addition of an imaginary plane separating the two loops. The charge asymmetry ratio is unity if the number of counterions is the same on both sides of the dividing plane. Conversely, when this parameter assumes values greater than unity, it denotes preferential accumulation of counterions on the Tar side of the complex. Figure 6 plots the asymmetry ratios for different counterions parsed across different free energy intervals. For each of the ions, the asymmetry ratio approaches unity as the magnitude of the free energy approaches zero, *i.e.*, as the ions become more bulk-like.

At the sequence level, regions of most negative free energy ion accumulation are dictated by the arrangement of purine residues, as ions in these regions make inner-sphere contacts with

electronegative N7 or O6 atoms from purine bases. The AMBER-99 force field assigns these atoms a partial charge of approximately $-0.6e$ and this explains the ability of these atoms to facilitate partial dehydration of oppositely charged ions. Guanine residues contain both N7 and O6 atoms. The Tar loop contains a tract of three guanines (Figure 1b), all of which are situated such that their N7 and O6 atoms face the inside of the loop in the Tar-Tar* structure. These atoms can substitute for first shell water molecules that are lost when counterions occupy the region created by the high local density of phosphate groups. The Tar* loop lacks the guanine tract and this asymmetry at the sequence level provides an intrinsic driving force for asymmetry in counterion accumulation.

However, within the most negative and intermediate free energy levels, the extent of asymmetry in counterion accumulation varies with ion type and follows the trend of being greatest for Cs^+ and smallest for Na^+ . For example, in the most negative free energy interval the asymmetry ratio is 5.2 for Cs^+ , 3.2 for K^+ , and 1.4 for Na^+ . This is because Na^+ accumulates around additional regions (see red arrows in Figure 2b) that are absent for K^+ and Cs^+ . One such region is located at the closest junction between the phosphate backbones of opposing loops. In Figure 2b we identify this region by a red arrow with the label 1. The presence of more regions for Na^+ occupancy in the most negative free energy interval appears to be responsible for reduced asymmetry ratios vis-à-vis K^+ and Cs^+ .

Preferential interactions coefficients

The extent of counterion accumulation and coion depletion around a macroion can be quantified in terms of the preferential interaction coefficient Γ_i , defined as the change in the number of cosolute molecules upon changing the macromolecule concentration¹⁷:

$$\Gamma_i = \left. \frac{\partial m_i}{\partial m_m} \right|_{T, \mu_i, \mu_m} \quad (3)$$

Here, m_i denotes the concentration of cosolute i and m_m denotes the concentration of the macroion. In an aqueous electrolyte, a macroion will preferentially accumulate counterions while depleting coions, leading to positive values for Γ_+ and negative values for Γ_- . In addition, the constraint of net electroneutrality requires that $\Gamma_+ + |\Gamma_-|$ must equal the net charge of the macroion. At low salt concentrations (< 0.01 M), all polyelectrolyte theories predict extensive counterion accumulation¹⁷, that is $\Gamma_+ \gg |\Gamma_-|$. In the high salt limit, the situation is reversed (*i.e.* $|\Gamma_-| \gg \Gamma_+$) on account of the excluded volume of the macroion¹⁸.

The calculated values for Γ_+ and Γ_- around Tar-Tar* are presented in Table 2. The net charge of the Tar-Tar* complex is $-30e$; therefore the quantity $\Gamma_+ + |\Gamma_-|$ must equal 30, and the data shown in Table 2 satisfy this constraint within error. Analysis of the individual preferential interaction coefficients for the Tar-Tar* complex reveal that, in 800 mM salt, charge neutralization is achieved via a nearly even split between counterion inclusion and coion exclusion for all three milieus, NaCl, KCl, and CsCl, respectively (see Table 2). In contrast to continuum electrostatics models, atomistic simulations have the advantage of being able to predict Γ_+ and Γ_- values at high salt concentrations, where the size of the individual ions and resultant ion-ion correlations cannot be ignored.

Comparison to a short, double-stranded DNA duplex

The preceding analysis identified three salient features regarding the distributions of accumulated ions around Tar-Tar*. These are as follows: 1) The most favorable free energy levels ($\Delta G \leq -1.5$ kcal/mol) are occupied by partially dehydrated ions (see figure 4). The first hydration shell statistics of these ions are markedly different from those in the bulk (see

figures 4 and 5). For smaller cations, the extent of dehydration is greater because favorable water-ion interactions are replaced by favorable interactions with the Tar-Tar* complex. The guanine tract in the Tar hairpin facilitates the requisite partial dehydration. 2) There is asymmetry in the extent of cation accumulation around each hairpin. However, the degree of the observed asymmetry is smallest for Na⁺ and largest for Cs⁺ (see figure 6). This is because the smaller cations can form a belt across the loop-loop interface (see figure 3). 3) Finally, in the bulk-like free energy intervals ($\Delta G > -0.5$ kcal/mol), there remain differences in cation accumulation and these are explained in terms of the generic differences between the corresponding 1:1 electrolytes (NaCl, KCl, and CsCl), where at high concentrations (ca. 800 mM), the degree of non-ideality varies with cation size, and is more prominent for the smaller cations.

One concern is that the cation specificity we observe for counterion accumulation around Tar-Tar* across the three free energy levels is strictly a consequence of the generic differences between the 1:1 electrolytes at high concentrations (see point 3 above). To test for this possibility, we analyzed ionic atmospheres around an oligo-DNA duplex because the stability of small double-stranded DNA should not exhibit significant dependence on cation type at the salt concentrations used in this study. Therefore, they serve as useful negative controls for our analysis. Owczarzy *et al.*¹⁹ showed that oligo-DNAs with 10-30 base pairs (bp) show no measurable differences in melting temperatures (T_m) in equimolar solutions of NaCl, KCl, and Tris-Cl buffers for monovalent salt concentrations ranging from 55 mM to 600 mM. Nakano *et al.* have measured the melting temperatures of 6 to 14bp duplexes of varying CG contents. These studies included dsDNA, dsRNA, and DNA/RNA hybrids. They reported the absence of measurable differences in T_m between equimolar buffers of NaCl, KCl, and CsCl.²⁰ Owczarzy *et al.* and Nakano *et al.* plot T_m values as a function of salt concentration rather than activity. This strategy masks the contribution from non-idealities at moderate to high salt concentrations. Conversely, if one were to plot the T_m values as a function of activities, then it would be clear that even in seemingly simple systems such as oligo dsDNA, there is a degree of cation specificity to quantities such as thermal stability.

We simulated the ionic atmospheres around a 9bp DNA oligomer d(GCCAGTTAA/TTAACTGGC), which is one of the sequences studied by Nakano *et al.*²⁰. This sequence was chosen because of the presence of two guanine residues in each of the complementary strands, thereby allowing us to compare the role of guanines between this system and Tar-Tar*. We carried out simulations of this duplex in the presence of 800 mM NaCl and CsCl and compared the ionic atmospheres around the duplex to that around Tar-Tar* to assess if this system acts as a negative control.

The number of accumulated counterions as a function of ΔG intervals is presented in figure 7 with the Tar-Tar* data copied from figure 3 for comparison. Irrespective of the free energy interval, the numbers for the accumulated counterions within each interval around the dsDNA 9-mer are 5 times smaller than those measured for Tar-Tar*, even though the net charge of the complex differs only by a factor of 2. As noted in the previous sections, specificity in accumulation of counterions around the Tar-Tar* system is a consequence of partial ion dehydration in the most negative free energy interval ($\Delta G \leq -1.5$ kcal/mol). The 9-mer, however, is characterized by a lack of measurable accumulation in the interval $\Delta G \leq -1.5$ kcal/mol for either NaCl or CsCl. This implies that partial dehydration is not realized with this system (see point 1 above for comparison). Differences in the number of accumulated ions do appear in the interval $\Delta G > -1.5$ kcal/mol and these differences arise on account of the guanine residues in each strand. Unlike the Tar-Tar* system, where the guanine residues are involved in partial dehydration, which in turn facilitates the population of ions in the most negative free energy intervals ($\Delta G \leq -1.5$ kcal/mol), the guanine residues

in the dsDNA system engage in favorable electrostatic interactions with fully hydrated cations. Since the smaller cations have higher charge densities, this leads to the differences depicted in figure 8 and translates into differences in cumulative ion numbers in intermediate to high free energy levels ($\Delta G > -1.5$ kcal/mol). These differences also have contributions that are generic in that they are attributable to the different non-idealities between NaCl and CsCl at high concentrations¹⁴. Finally, figure 8 shows the contours for counterion accumulation around the dsDNA 9-mer. The dominant region of counterion accumulation lies in the minor groove adjacent to two guanine bases for both Na⁺ and Cs⁺ (figure 8, white arrows). However, the B-form minor groove is equidistant from either strand and therefore the structure of duplex is such that it there can be no driving force for asymmetry in cation accumulation around each of the two DNA strands. In summary, the differences between Na⁺ and Cs⁺ accumulation around the dsDNA oligomer arise entirely due the presence of guanine residues and the generic differences in non-idealities within the corresponding 1:1 electrolytes namely, NaCl and CsCl. Features such as differences in partial dehydration and differences in degree of asymmetry in cation accumulation are missing for this system because this system does not possess the sequence encoded structural features of the Tar-Tar* system. Therefore, we propose that the observed cation specific differences in counterion accumulation around the Tar-Tar* system may indeed be the cause of the experimentally observed cation specific stability.

Discussion

The accompanying paper by Lambert *et al.* describes a range of RNA systems that exhibit monovalent counterion specificity in the interactions that govern folding or complex formation. In their data, the stability of the Tar-Tar* kissing loop complex is inversely proportional to cation size. Our goal in this work was to provide a plausible explanation for the experimentally observed cation specificity in Tar-Tar* stability. To calculate ion-dependent stabilities of a complex such as Tar-Tar* is a daunting task. Therefore, we used simulations to study the distribution of counterions around Tar-Tar* to seek a correlative explanation for the experimentally observed trends regarding ion-dependent stability

Analysis of the preferential interaction coefficients for the Tar-Tar* complex reveal that, in 800 mM salt, charge neutralization is achieved via a nearly even split between counterion inclusion and coion exclusion for all three salts (Table 2). This indicates that the salt concentration in our simulations is approximately halfway between the low salt limit where $\Gamma_+ \gg |\Gamma_-|$ and the high salt limit, where $\Gamma_- \gg |\Gamma_+|$. This proposal is in agreement with Donnan equilibrium studies by Strauss *et al.*²¹, who showed that Γ_- is approximately 0.6 per phosphate group in 1 M NaBr for double-stranded DNA.

Lambert *et al.* have measured the variation of $\log(K_{\text{obs}})$ as a function of the logarithm of the activity of monovalent salt. This yields an estimate for $2\Delta\Gamma_{\pm}$, which is interpreted as the number of ions taken up or released when Tar and Tar* form a complex. Although this can resolve minute changes in ion accumulation (~ 0.1 ions), it does not reveal the absolute numbers of accumulated ions. Furthermore, $2\Delta\Gamma_{\pm}$ is a “catch-all” quantity that does not distinguish between ions at different free energy levels. For molecular simulations, the situation is reversed. The absolute number of accumulated ions around the macroion is directly measurable, but changes in ion accumulation upon complex formation can only be indirectly assessed via comparative simulations of the two end-states. Additionally, it can be seen in Figure 3 that the error depends on the free energy of ion accumulation; error bars are smallest for the most favorable free energy levels and largest for marginally favorable free energy levels. Unfortunately, the largest error bars impact calculations of preferential interaction coefficients.

A direct comparison between our calculated preferential interaction coefficients and the estimates of $2\Delta\Gamma_{\pm}$ from Lambert *et al.* could, in principle, be accomplished with additional simulations of ion accumulation around the isolated, individual hairpins. In practice, however, each calculated Γ carries an intrinsic error of ± 2 ions, making any potential agreement with experimental values (which differ only by ca. 0.3 ions) statistically insignificant.

Our simulation data indicate differences in ionic atmospheres around Tar-Tar* that depend on the identities of monovalent counterions. The trends seen in the simulations suggest a plausible explanation for the experimental data. Across all free energy levels, ion accumulation follows the trend of $cc(\text{Na}^+) > cc(\text{K}^+) > cc(\text{Cs}^+)$, where cc refers to the cumulative number of counterions within a free energy level. Ion accumulation around Tar-Tar* is characterized by continuous variation in accumulated ion densities as a function of free energy (Figure 3). The kissing loop structure creates pockets that collect high densities of phosphate groups as well as O6 and N7 atoms from purines that facilitate the accumulation of counterions around the RNA. The smaller the ion, the higher the charge density, and the lower the free energy of hydration. Such ions are amenable to partial dehydration by phosphates as well as electronegative atoms on purines because they can exchange strong ion-water (charge-dipole) interactions for stronger ionmacroion (charge-charge) interactions. We find that the smaller Na^+ has a higher probability of accessing tri-, tetra-, and penta-hydrate states vis-à-vis K^+ and Cs^+ and this accounts for the increased extents of Na^+ accumulation around Tar-Tar*. Differences in ion accumulation in the intermediate-weak free energy levels originate in the generic differences between the ions in high concentration bulk solutions that result from non-idealities such as ion pairing and clustering that were previously documented¹⁴. Although these clusters do not directly interact with the macroion, their formation is governed by the same driving forces that result in highly favorable counterion accumulation around the macroion *i.e.*, a tug-of-war between unfavorable partial dehydration and favorable unshielded coulomb interactions. Finally, the presence of a G-rich tract in Tar leads to a preferential accumulation of counterions on the Tar side of the Tar-Tar* complex. This asymmetry in counterion accumulation is most prominent in the most negative free energy levels and within these levels the asymmetry decreases with decreasing ion size.

Given our data, we propose that smaller, high charge density ions have two distinct advantages over their larger counterparts: First, they can access regions in the most negative free energy intervals (presumably regions of high electrostatic potential) around the individual hairpins as well as across the hairpin interface. Accessibility to the latter region decreases with increasing ion size, a point is made by the asymmetry ratios shown in Figure 6. Second, in the intermediate to weak free energy levels, smaller ions have the ability to form “belts” across the interface that are suggestive of being important in stabilizing the complex. These “belts” consist of fully hydrated ions, which implies that the smaller hydrated radius of Na^+ ions allows access to more electrostatically favorable regions of the distorted minor groove. If these features are responsible for the experimentally observed ion dependence in Tar-Tar* stability, then mutations that disrupt or scramble the guanine tract within the Tar hairpin will most likely destabilize the complex or at least modulate the dependence of the stability of Tar-Tar* on the identities of monovalent counterions. It would be useful to test the features predicted by the spatial distribution functions for the ionic atmospheres around Tar-Tar* and similar complexes. This seems feasible given recent advances in high resolution inelastic X-ray scattering measurements that are being applied to measure quantities related to the Fourier transforms of spatial distribution functions of counterions around highly charged macroions²². Additionally, advances in simulation methodology, including multiple long simulations, are also needed to allow quantitative

comparison between $\Delta\Gamma_{\pm}$ values estimated from simulation data and those estimated from experimental data.

Materials and Methods

Simulations and Analysis

We report results from analysis of multiple molecular dynamics simulations of the Tar-Tar* complex. All simulations contained 800 *mM* of NaCl, KCl, or CsCl, where *m* denotes molality. Error estimates for all measurements were obtained as the standard deviation across the 4 replicate trajectories acquired for each simulation condition.

Details of the forcefield

RNA was modeled using the AMBER-99 forcefield²³ employing the GROMACS implementation of Sorin and Pande²⁴. The rigid 3-site TIP3P model²⁵ was used to simulate water molecules. Ions were modeled using the parameters of Åqvist²⁶ according to the approach proposed by Chen and Pappu²⁷. It has been shown that the use of the default implementation of Åqvist's ion parameters in AMBER-99 results in unphysical lattice formation of 1:1 electrolytes at concentrations well below the solubility limit^{27; 28}. However, in recent work²⁷ we have shown that this abnormal clustering is the result of an *ad hoc* adaptation of Åqvist's ion parameters in the default AMBER-99 forcefield. We have also shown that simulations based on Åqvist's original parameters reproduce the transient ion pairing and clustering propensities of strong 1:1 electrolytes¹⁴ over a wide range of concentrations (100 *mM* to 1*M*) as adjudicated by estimates from conductometric experiments¹⁵. Use of the original Åqvist parameters with the AMBER-99 forcefield for nucleic acids results in cation-anion and cation-phosphate oxygen contact distances that are consistent with measurements from high-resolution crystal structures of model compounds²⁷.

Details of the molecular dynamics simulations

All molecular dynamics simulations were performed using version 3.3.2 of the GROMACS molecular dynamics package²⁹. All simulations used the isothermal-isobaric ensemble, *i.e.*, constant pressure of 1 bar and temperature of 298K. The weak coupling algorithms of Berendsen and coworkers were used for both the manostat and the thermostat³⁰, with coupling constants of 1 ps and 0.2 ps, respectively. The equations of motion were integrated using a 2 fs time step and the leapfrog algorithm³¹. The two bond lengths and one bond angle in each water molecule were constrained to values prescribed by the TIP3P model using the SETTLE algorithm of Miyamoto and Kollman³². Snapshots were saved for analysis once every 2 ps. Periodic boundary conditions were employed to mimic the macroscopic setting for electrolytes. Long-range electrostatic interactions between periodic images were treated using the particle mesh Ewald approach³³, with fourth-order cubic interpolation and a tolerance of 10^{-5} . Neighbor lists were updated every 10 time steps. A cutoff of 10 Å was used for van der Waals interactions, real space Coulomb interactions, and for updating neighbor lists.

Initial coordinates for the Tar-Tar* complex were obtained from the NMR structure of Chang and Tinoco⁸ (RCSB accession ID 1KIS). Note that this structure contains a van der Waals violation between the O1P atoms of residues C6 and U22 that is readily resolved via energy minimization. The Tar-Tar* complex was placed in a cubic box 120 Å on a side containing 55,118 TIP3P waters, 792 anions, and 792 + 30 cations to maintain net electroneutrality. The ion numbers were chosen to correspond to a concentration in molal units of 800 *mM*. Four independent simulations, each of length 15 ns were generated for the Tar-Tar* complex in each of the three salts (NaCl, KCl, CsCl). In each simulation the

trajectory was initialized using randomized velocities and ion positions, and allowed to equilibrate for 5 ns. We settled on the choice of 15 ns (+5 ns equilibration) trajectories based on the observation of sub-ns relaxation times observed for rearrangements of ionic atmospheres around all RNA constructs. A detailed analysis of this point is provided in Appendix 2.

The choice of simulation timescales merits explanation in light of data from Ponomarev *et al.*³⁴ They found the characteristic relaxation times for counterions around double stranded DNA to be ca. 5 ns. This required simulations that were longer than 60 ns to construct converged ion-DNA pair correlation functions. Our simulations differ in two important respects. First, we simulate solution conditions with high concentrations of excess salt as opposed to the salt free conditions with neutralizing counterions alone that were used by Ponomarev *et al.* Secondly, we model finite salt concentrations using atypically large numbers of ions (~800 pairs), which in turn facilitates rapid exchange between two sets of ionic species – bulk and local ions (see Appendix 1 and 2). We find that using small numbers of ion pairs with small boxes creates additional finite size artifacts that include a mismatch between bulk anion and cation concentrations (as detailed in Appendix 1) and also significantly slows the rate of ionic rearrangement (as described in Appendix 2).

Calculation of preferential interaction coefficients

Preferential interaction coefficients were calculated using the formalism of Smith³⁵ in which component 1 is the solvent (water), component 2 is the macromolecule at infinite dilution, and component 3 is a co-solute (either the cation or the anion). The preferential interaction of the component 3 around component 2 (the macromolecule) is therefore:

$$\Gamma_{23} = N_{23} - \frac{\rho_3}{\rho_1} N_{21} \quad (4)$$

Here, ρ_i denotes the density of component i and N_{2i} denotes net excess or deficit of particles of type i in a local partition surrounding the macromolecule compared to a bulk partition of equal volume. The second term on the right hand of equation 4 subtracts out any contribution to N_{23} attributable to accumulation of solvent around the macromolecule. In practical terms, this calculation reduces to defining a local volume around the RNA large enough to contain all the thermodynamically perturbed ions. The exact shape of this local volume is unimportant, as any volume of bulk ions that is accidentally included does not contribute to the net excess / deficit number N_{2i} once the appropriate bulk prior is subtracted. Using the phosphate-counterion radial distribution functions, one can estimate this volume to extend roughly 40 Å from the center of the RNA in all directions. The local volume can be defined as a sphere of radius r_c placed at the center of mass of the RNA hairpin. The excess / deficit numbers N_{2i} are defined as:

$$N_{2i} = \langle n_{2i} \rangle_{r \leq r_c}^{obs} - \langle n_{2i} \rangle_{r \leq r_c}^{ideal} \quad (5)$$

$\langle n \rangle_{r \leq r_c}^{obs}$ is the average number of cations, anions, or water within the local partition. $\langle n \rangle_{r \leq r_c}^{ideal}$ is obtained by multiplying the bulk density ρ_i by the volume of the local compartment (in this case a sphere of radius r_c minus the van der Waals volume of the Tar-Tar* complex). If a sufficiently large local volume is chosen, the results should satisfy the electroneutrality relationship: $|Z| = |\Gamma_+| + |\Gamma_-|$; Here, Z is the net charge of the RNA, which is -30e for the Tar-Tar* complex. A sensitivity analysis of the resulting Γ_+ , Γ_- , and $|\Gamma_+| + |\Gamma_-|$ reveals that

any value of $r_c \geq 40 \text{ \AA}$ results in identical values for the calculated preferential interaction coefficients.

Calculation of charge asymmetry ratios

In order to quantify the relative preference of counterions to accumulate around Tar over Tar*, we computed a charge asymmetry ratio defined as:

$$A_{f \geq f_c} = \left\langle \frac{\sum \rho_{f \geq f_c}^{Tar}(x, y, z)}{\sum \rho_{f \geq f_c}^{Tar^*}(x, z, y)} \right\rangle \quad (6)$$

Here, f_c denotes a cutoff in free energy from the spatial free energy density profiles and $\rho(x, y, z)$ is the cation spatial distribution function, which was divided into Tar and Tar* halves via an imaginary plane bisecting the loop-loop interface, as shown by the green dividing line in Figure 2a. Note that an asymmetry ratio of 1 denotes equal counterion accumulation, whereas values greater than 1 denote preferred accumulation on the Tar half of the complex.

Acknowledgments

This work was supported by grant MCB-0718924 from the National Science Foundation. The authors are grateful to Nathan Baker, Kathleen Hall, Timothy Lohman, and Andreas Vitalis for many helpful and stimulating discussions.

Appendix 1: Finite size artifacts and choice of system size

Finite size artifacts plague all molecular simulations that use periodic boundary conditions because we model macroscopic systems using small numbers of particles. This operational constraint requires that we be careful when choosing the system size for molecular simulations. This is especially important in simulations with ions where the system needs to be large enough to provide a *bona fide* partitioning between local and bulk ions. The presence of a macroion causes counterion accumulation and coion depletion in its local atmosphere. However, this accumulation / depletion should occur within a well-defined length scale around the macroion beyond which the counter and coions should be distributed like they would in a bulk electrolytic solution. This type of partitioning between local and bulk ions can only occur if the simulation box is sufficiently large to help avoid spurious accumulation of depleted coions near the edge of the box. Also, the use of smaller boxes results in artificial truncation of the ionic atmosphere around the macromolecule. This truncation needs to be avoided to extract accurate thermodynamic information regarding the ionic atmosphere around the macroion. Our findings described below suggest that a box size of $120 \text{ \AA}$ per side is adequate based on the assessment of bulk-like behavior beyond the atmosphere around the macroion. Clearly, larger boxes would further decrease errors due to finite size, but would cause significant increase in computational expense. Previous work¹⁴ focused on the calibration of strong 1:1 electrolytes as a function of concentration, all the way from 100 mM to 1 M . These calibrations guided our analysis of both finite size artifacts, especially the bulk behavior, and the analysis of partial hydration states of ions.

For our simulations in an excess salt concentration of 800 mM , we found it necessary to employ an atypically large central simulation cell ($120 \text{ \AA}$) containing 55,118 waters and 792 ion pairs. The size of the simulation cell is substantially larger than the RNA complex itself, which is roughly cylindrical measuring $15 \times 50 \text{ \AA}$. Employing smaller simulation cells leads to significant accumulation of coions near the edge of the simulation cell and excess anions in the bulk electrolyte as demonstrated in Figure 9a for an $80 \text{ \AA}$ box. The calculated

preferential interaction coefficients (Table 2) indicate that the complex will preferentially include ~15 cations while excluding ~15 anions, and in the 80 Å box this results in anion concentrations ~15% higher than the bulk cation concentration. This problem is alleviated by the use of larger simulation cells with correspondingly larger numbers of ions. We tried cells of size 90, 100, 110, and 120 Å to a side, and found that only for the 120 Å box do we obtain bulk anion and cation concentrations that agreed with each other to within 1% (Figure 9b). This demonstrates that the key determinant in choosing an appropriate box size is the ratio of the number of thermodynamically perturbed ions to the number of ions in the bath.

Appendix 2: Finite size artifacts and choice of simulation lengths

Converged estimates of ionic atmosphere rearrangements around a macroion can only be determined if the simulation is much longer than the intrinsic timescale of ionic rearrangement. In our prior work¹⁴, we determined that the reorganization time of bulk electrolytes in the absence of a macromolecule was ca. 100 ps. To assess how this timescale was altered by the presence of the RNA, we examined the time dependent relaxation of the ionic atmosphere around the Tar-Tar* complex. A coaxial cylinder enclosing the RNA was created and the total net charge enclosed was tabulated as a function of cylinder radius. This is similar to the procedure used to calculate preferential interaction coefficients. When the cylinder is sufficiently large, the net charge enclosed should be zero (this includes the macromolecule's net charge of -30e). These measurements were repeated as a function of time after initial system preparation (with randomized ion positions) to assess the timescales over which equilibrium distributions were reached. As shown in Figure 10, the relaxation in the 80 Å box behaves differently than the 120 Å box due to finite size effects. Specifically, charge fluctuations are slow and strongly coupled at all length scales in the 80 Å box, indicating that there are too few ions to create distinct local and bulk populations. In contrast, in simulations with the 120 Å box, we find fast relaxation of the local ion partition (< 1ns) while the bulk partition oscillates independently and slowly around zero net charge, indicating that these are truly distinct populations of ions. The presence of fluctuations in the bulk of ± 2 ions dictates the intrinsic error in all measurements of ion accumulation made for this system, which is manifest in the size of the error bars in the preferential interaction coefficients (Table 2). This intrinsic error is unlikely to be reduced by longer simulations or even greater numbers of simulations unless substantially larger boxes are employed. However, in light of the sub-nanosecond relaxation time of the ionic atmosphere, our choice of 15 ns simulations should be sufficient to capture multiple ionic atmosphere reorganizations, and the use of 5 ns equilibration is sufficient to cause a reduction in bias from the initial configuration of ions.

Appendix 3: Ion residence times

We quantified ion residence times within two sets of free energy intervals using the definitions of intervals that were employed in the hydration analysis; the most negative interval ($\Delta G \leq -1.5$ kcal/mol) and the intermediate free energy interval ($-1.5 < \Delta G \leq -0.5$ kcal/mol). Residence times were tabulated as the length of time elapsed between the initial entry and exit of a counterion from any region contained within the designated free energy interval. Since snapshots were saved in 2 ps intervals, this is the shortest resolvable time interval for analysis of ion residence times. A total of 3341, 3766, and 4895 counterion entry / exit events were observed for the most negative free energy interval of Na⁺, K⁺, and Cs⁺ respectively. In the intermediate free energy interval, 6487, 15298, and 19456 entry / exit events were observed for Na⁺, K⁺, and Cs⁺ respectively.

The mean ion residence times of cations in each free-energy interval are presented in table 3. These times are systematically longer for the most negative free energy interval when

compared to the intermediate free interval. Within a free energy interval, the ion residence times are longest for Na⁺ and smallest for Cs⁺ and these trends track with the features obtained from analysis of the spatial free energy density profiles.

References

1. Bukhman YV, Draper DE. Affinities and Selectivities of Divalent Cation Binding Sites Within an RNA Tertiary Structure. *J. Mol. Biol* 1997;273:1020–1031. [PubMed: 9367788]
2. Takamoto K, He Q, Morris S, Chance MR, Brenowitz M. Monovalent cations mediate formation of native tertiary structure of the *Tetrahymena thermophila* ribozyme. *Nat. Struct. Biol* 2002;9:928–933. [PubMed: 12434149]
3. Wagner EGH, Simons RW. Antisense RNA Control in Bacteria, Phages, and Plasmids. *Ann. Rev. Microbiol* 1994;48:713–42. [PubMed: 7826024]
4. Montange RK, Batey RT. Riboswitches: Emerging Themes in RNA Structure and Function. *Ann. Rev. Biophys* 2008;37:117–133. [PubMed: 18573075]
5. Miller WA, White KA. Long-Distance RNA-RNA Interactions in Plant Virus Gene Expression and Replication. *Annu. Rev. Phytopathol* 2006;44:447–467. [PubMed: 16704356]
6. Skripkin E, Paillart J-C, Marquet R, Ehresmann B, Ehresmann C. Identification of the primary site of the human immunodeficiency virus type 1 RNA dimerization *in vitro*. *Proc. Natl. Acad. Sci. U. S. A* 1994;91:4945–4949. [PubMed: 8197162]
7. Ly H, Parslow TG. Bipartite Signal for Genomic RNA Dimerization in Moloney Murine Leukemia Virus. *J. Virol* 2002;76:3135–3144. [PubMed: 11884538]
8. Chang K-Y, Tinoco I Jr. The Structure of an RNA “Kissing” Hairpin Complex of the HIV TAR Hairpin Loop and its Complement. *J. Mol. Biol* 1997;269:52–66. [PubMed: 9193000]
9. Réblová K, Špačková N. a. Šponer JE, Koča J, Šponer J. Molecular dynamics simulations of RNA kissing-loop motifs reveal structural dynamics and formation of cation-binding pockets. *Nucleic Acids Res* 2003;31:6942–6952. [PubMed: 14627827]
10. Collins KD. Charge Density-Dependent Strength of Hydration and Biological Structure. *Biophys. J* 1997;72:65–76. [PubMed: 8994593]
11. Nair TM, Myszkowski DG, Davis DR. Surface plasmon resonance kinetics studies of the HIV TAR RNA kissing hairpin complex and its stabilization by 2-thiouridine modification. *Nucleic Acids Res* 2000;28:1935–1940. [PubMed: 10756194]
12. Humphrey W, Dalke A, Schulten K. VMD - Visual Molecular Dynamics. *J. Molec. Graphics* 1996;14:33–38.
13. Misra VK, Draper DE. A thermodynamic framework for Mg²⁺ binding to RNA. *Proc. Natl. Acad. Sci. U. S. A* 2001;98:12456–12461. [PubMed: 11675490]
14. Chen AA, Pappu RV. Quantitative Characterization of Ion Pairing and Cluster Formation in Strong 1:1 Electrolytes. *J. Phys. Chem. B* 2007;111:6469–6478. [PubMed: 17518490]
15. Fuoss RM. Conductimetric determination of thermodynamic pairing constants for symmetrical electrolytes. *Proc. Natl. Acad. Sci. U. S. A* 1980;77:34–8. [PubMed: 16592752]
16. Bouazizi S, Nasr S, Jaïdane N, Bellissent-Funel M-C. Local Order in Aqueous NaCl Solutions and Pure Water: X-ray Scattering and Molecular Dynamics Simulations Study. *J. Phys. Chem. B* 2006;110:23515–23523. [PubMed: 17107207]
17. Anderson CF, Record MT Jr. Salt-nucleic acid interactions. *Annu. Rev. Phys. Chem* 1995;46:657–700. [PubMed: 7495482]
18. Anderson CF, Record MT Jr. The Relationship Between the Poisson-Boltzmann Model and the Condensation Hypothesis: An Analysis based on the Low Salt Form of the Donnan Coefficient. *Biophys. Chem* 1980;11:353–360. [PubMed: 16997254]
19. Owczarzy R, Moreira BG, You Y, Behlke MA, Walder JA. Prediction Stability of DNA Duplexes in Solutions Containing Magnesium and Monovalent Cations. *Biochemistry* 2008;47:5336–5353. [PubMed: 18422348]
20. Nakano, S.-i.; Fujimoto, M.; Hari, H.; Sugimoto, N. Nucleic acid duplex stability: influence of base composition on cation effects. *Nucleic Acids Res* 1999;27:2957–2965. [PubMed: 10390539]

21. Strauss UP, Helfgott C, Pink H. Interactions of polyelectrolytes with simple electrolytes. II. Donnan equilibria obtained with DNA in solutions of 1-1 electrolytes. *J. Phys. Chem* 1967;71:2550–2556. [PubMed: 6063711]
22. Angelini TE, Golestanian R, Coridan RH, Butler JC, Beraud A, Krisch M, Sinn H, Schweizer KS, Wong GCL. Counterions between charged polymers exhibit liquid-like organization and dynamics. *Proc. Natl. Acad. Sci. U. S. A* 2006;103:7962–7967. [PubMed: 16690742]
23. Cornell WD, Cieplak P, Bayly CI, Gould IR, Merz KM Jr. Ferguson DM, Spellmeyer DC, Fox T, Caldwell JW, Kollman PA. A Second Generation Force Field for the Simulation of Proteins, Nucleic Acids, and Organic Molecules. *J. Am. Chem. Soc* 1995;117:5179–5197.
24. Sorin EJ, Pande VS. Exploring the Helix-Coil Transition via All-Atom Equilibrium Ensemble Simulations. *Biophys. J* 2005;88:2472–2493. [PubMed: 15665128]
25. Jorgensen WL, Chandrasekhar J, Madura JD, Impey RW, Klein ML. Comparison of simple potential functions for simulating liquid water. *J. Phys. Chem* 1983;79:926–935.
26. Åqvist J. Ion Water Interaction Potentials Derived from Free-Energy Perturbation Simulations. *J. Phys. Chem* 1990;94:8021–8024.
27. Chen AA, Pappu RV. Parameters of Monovalent Ions in the AMBER-99 Forcefield: Assessment of Inaccuracies and Proposed Improvements. *J. Phys. Chem. B* 2007;111:11884–11887. [PubMed: 17887792]
28. Auffinger P, Cheatham TE III, Vaiana AC. Spontaneous Formation of KCl Aggregates in Biomolecular Simulations: A Force Field Issue? *J. Chem. Theory Comput.* 2007
29. van der Spoel D, Lindahl E, Hess B, Groenhof G, Mark AE, Berendsen HJC. GROMACS: Fast, Flexible, and Free. *J. Comput. Chem* 2005;26:1701–1718. [PubMed: 16211538]
30. Berendsen HJC, Postma JPM, van Gunsteren WF, DiNola A, Haak JR. Molecular dynamics with coupling to an external bath. *J. Chem. Phys* 1984;81:3684–90.
31. Hockney, RW.; Eastwood, JW. *Computer Simulations Using Particles*. McGraw-Hill; New York: 1981.
32. Miyamoto S, Kollman PA. SETTLE: An analytical version of the SHAKE and RATTLE for rigid water models. *J. Comput. Chem* 1992;13:952–962.
33. Darden T, York D, Pedersen L. Particle mesh Ewald: An Nlog(N) method for Ewald sums in large systems. *J. Chem. Phys* 1993;98:10089–92.
34. Ponomarev SY, Thayer KM, Beveridge DL. Ion motions in molecular dynamics simulations on DNA. *Proc. Natl. Acad. Sci. U. S. A* 2004;101:14771–14775. [PubMed: 15465909]
35. Smith PE. Cosolvent Interactions with Biomolecules: Relating Computer Simulations Data to Experimental Thermodynamic Data. *J. Phys. Chem. B* 2004;108:18716–18724.

Figure 1.

(a) Model for the Tar-Tar* kissing loop complex depicted using coordinates obtained from NMR data (pdb 1KIS). The Tar hairpin is shown in red, while the Tar* hairpin is shown in blue. Yellow spheres indicate selected backbone phosphorus atoms that approach within 5.25 Å at the loop-loop interface. **(b)** Schematic of secondary structure for the Tar-Tar* kissing loop complex. The schematic shows that the Tar loop is G-rich, while the complementary Tar* loop is deficient in guanines.

Figure 2. Spatial free energy density profiles for counterion accumulation around the Tar-Tar* complex. Free-energy isocontours (surfaces of zero thickness at specific free-energy values) are colored according to the key shown in panel **b**. The volume encompassed by each isocontour, however, contain regions of more negative free energy density than prescribed by the isocontour itself. **(a)** All four isocontours visualized around a static RNA structure; Tar* is on top in gray while Tar is the bottom hairpin, in white. **(b)** A horizontal slice looking down at the Tar loop from the Tar-Tar* interface, with perspective as indicated by the blue line in panel **a**. Red arrows indicate areas of most negative free energy counterion accumulation that are present for Na⁺ and absent in K⁺ and Cs⁺. The Na⁺ atmosphere also

exhibits a “belt” of intermediate to weak density counterions encircling the loop-loop interface (green arrows), which is not as prominent in either K^+ or Cs^+ . (c) Depiction of the three most negative free energy isocontours situated in the interior of the loop-loop interface; the RNA solvent accessible surface is shown in transparent gray.

Figure 3. Cumulative distribution functions to assess the amount of counterion accumulation around the Tar-Tar* complex. The ordinate quantifies the number of counterions accumulated around Tar-Tar* within contours that have ΔG less than or equal to the value on the abscissa.

Figure 4. Cation hydration analysis. Distribution of hydration numbers of cations in the most negative ($\Delta G \leq -1.5$ kcal/mol) and intermediate ($-1.0 < \Delta G \leq -0.5$ kcal/mol), regions of the free energy density profiles.

Figure 5.

Deviation of counterion hydration statistics from bulk distribution, as defined in equation 2. A value of 0 indicates hydration statistics that are identical to those found in the bulk, 800mm electrolyte, while larger values indicate greater dissimilarity vis-à-vis the bulk.

Figure 6.

Bar graphs of charge asymmetry ratios calculated to quantify preferential accumulation of counterions around the Tar as opposed to Tar* hairpin in the Tar-Tar* complex. The same free energy cutoffs are used as in Figure 2, therefore the purple bar represents all ions in the range $\Delta G \leq -2.0$ kcal/mol, the red bar $-2.0 < \Delta G \leq -1.5$ kcal/mol, the orange bar $-1.5 < \Delta G \leq -1.0$ kcal/mol, and the yellow bar $-1.0 < \Delta G \leq -0.5$ kcal/mol. An asymmetry ratio of 1 denotes equivalent accumulation of counterions around both hairpins, while values greater than 1 denote preferential accumulation around the Tar hairpin.

Figure 7.

Comparison of cumulative counterion accumulation as a function of ΔG cutoff around a dsDNA 9-mer duplex in 800 mM NaCl and CsCl, respectively compared to the cumulative counterion numbers for Tar-Tar*. (Data for the Tar-Tar* complex were taken from figure 3).

Figure 8.

Spatial free energy density profiles for accumulation of Na^+ (left) and Cs^+ (right) around a 9-mer DNA duplex. Isocontours are colored at the same values as in figure 2; note the absence of the most negative purple and red contours.

Figure 9.

Illustration of how finite size artifacts lead to concentration mismatches in the bulk: **(a)** RNA phosphate-counterion radial distribution functions for the Tar-Tar* complex in a 80 Å box of 800 mM NaCl. The anion concentration is ~15% different than the cation concentration at large separations. **(b)** RNA phosphate-counterion radial distribution functions for the Tar-Tar* complex in a 120 Å box of 800mM NaCl. The anion and cation concentrations agree with each other to within 1% at large radii.

Figure 10.

Illustration of finite size effects on relaxation times for ionic atmospheres. **(a)** Net charge within a cylinder of various radii around the Tar-Tar* complex in 800 mM NaCl in a 80 Å box as a function of time after initial system preparation. Note that fluctuations are slow and strongly coupled at all length scales. **(b)** Net charge within a cylinder of various radii around the Tar-Tar* complex in 800 mM NaCl in a 120 Å box as a function of time after initial system preparation. Note that fluctuations are rapid in the local partition and slow but small in the bulk partition, indicating a true separation of ionic populations.

Table 1

Monovalent counterion dependence of the stability of the Tar-Tar* complex. Experimentally derived estimates for the stabilities (ΔG° in kcal/mol at 25°C) of the Tar-Tar* complex and the Tar hairpin in three different alkali-chloride salts as measured by UV-vis thermal melts. All experiments were carried out in 400 mM monovalent salt, at 25° C, with no divalent counterions. Data are taken from accompanying manuscript of Lambert *et al.*

	Tar-Tar*	Tar*
NaCl	-8.3 ± 0.2	-6.2
KCl	-7.4 ± 0.2	-6.1
CsCl	-6.4 ± 0.1	-6.0

Table 2

Calculated preferential interaction coefficients around the Tar-Tar*complex. Γ_+ and Γ_- denote degree of counterion inclusion and coion exclusion respectively, and the sum $\Gamma_+ + |\Gamma_-|$ recapitulates the total macromolecular charge being compensated.

	$\Gamma_+ : \text{Tar-Tar}^*$	$\Gamma_- : \text{Tar-Tar}^*$	$\Gamma_+ + \Gamma_- $
NaCl	15.0 ± 2.3	-14.8 ± 2.7	29.8 ± 3.5
KCl	13.9 ± 2.1	-16.3 ± 2.4	30.2 ± 3.2
CsCl	13.9 ± 2.2	-15.9 ± 2.3	29.8 ± 3.2