Understanding Nucleic Acid–Ion Interactions

Jan Lipfert,1,2 Sebastian Doniach,3,4 Rhiju Das,5 and Daniel Herschlag5

1Department of Bionanoscience, Kavli Institute of Nanoscience, Delft University of Technology, 2628 CJ Delft, Netherlands; email: Jan.Lipfert@lmu.de
2Department of Physics and Center for NanoScience, University of Munich, 80799 Munich, Germany
Departments of 3Physics, 4Applied Physics, and 5Biochemistry, Stanford University, Stanford, California 94305; email: sxdwc@slac.stanford.edu, rhiju@stanford.edu, herschla@stanford.edu

Keywords
ions, RNA/DNA, electrostatics, Poisson–Boltzmann, Manning condensation, Hill equation, free energy

Abstract
Ions surround nucleic acids in what is referred to as an ion atmosphere. As a result, the folding and dynamics of RNA and DNA and their complexes with proteins and with each other cannot be understood without a reasonably sophisticated appreciation of these ions’ electrostatic interactions. However, the underlying behavior of the ion atmosphere follows physical rules that are distinct from the rules of site binding that biochemists are most familiar and comfortable with. The main goal of this review is to familiarize nucleic acid experimentalists with the physical concepts that underlie nucleic acid–ion interactions. Throughout, we provide practical strategies for interpreting and analyzing nucleic acid experiments that avoid pitfalls from oversimplified or incorrect models. We briefly review the status of theories that predict or simulate nucleic acid–ion interactions and experiments that test these theories. Finally, we describe opportunities for going beyond phenomenological fits to a next-generation, truly predictive understanding of nucleic acid–ion interactions.
INTRODUCTION

Nucleic acids are central to the storage, transmission, processing, and regulation of genetic information, serving as carriers of the genetic code, catalysts, and mechanochemical switches. Although the importance of ions for the structure, folding, and function of DNA and RNA, and for their interactions with proteins has long been appreciated, the physical properties and energetics of ion–nucleic acid interactions are often not intuitive. Misconceptions remain common.

Nucleic acid–ion interactions provide large interaction energies. Thus, these interactions are a critical component of a complete description of the folding of functional DNAs and RNAs, of interactions of nucleic acids with ligands and macromolecule partners, and of the function of RNAs and RNA–protein complexes and machines. These processes include viral packaging and function and the assembly and function of the ribosome, spliceosome, signal recognition particle, telomerase, chromosomes, and chromatin (1–6). Quantitative
description of these interactions is a grand challenge that needs to be met to fully understand biological function and regulation involving nucleic acids and to rationally manipulate and engineer nucleic acids and their complexes.

On the surface, understanding nucleic acid–ion interactions would seem no more complex than any receptor–ligand interaction. However, the vast majority of ion–nucleic acid interactions involve the so-called ion atmosphere, a loosely associated sheath of ions surrounding the nucleic acid polyelectrolyte, and even site-bound metal ions are influenced by the ion atmosphere. Because this ion atmosphere is highly mobile, involves many ions, and varies with ionic conditions, it is not easily appreciated and explained.

Several treatments have laid out thermodynamic and mathematical frameworks of ion–nucleic acid interactions (e.g., References 7–13). However, these treatments are often steeped in thermodynamic language unfamiliar to most biochemists and molecular biologists. Recent experimental breakthroughs now provide an opportunity to present a more intuitive and empirically grounded description of the important properties of the ion atmosphere and its influence on nucleic acids and their interactions. Drawing both on theoretical models of the ion atmosphere and on the latest experimental findings, this review aims to lay out a conceptual framework to guide thinking about ion–nucleic acid interactions and aid the nucleic acid experimental practitioner. We also hope that this conceptual treatment helps clarify and focus the considerable future challenges in this area.

THE BASICS

Electrostatic Effects in Nucleic Acids: How Important Are They?

Although electrostatic interactions are important for all biological macromolecules, their energetic consequences are enormous for nucleic acids. Consider that only 5 of the 20 amino acids that make up proteins are charged at physiological pH, and as there are both positively and negatively charged amino acids, the net charge on most proteins tends to be small (Figure 1) (14, 15). In contrast, the phosphate backbone of nucleic acids carries one negative charge per residue. As a result, nucleic acids are highly charged (Figure 1) (16). Bringing these charges in close proximity during RNA folding or DNA compaction requires overcoming an enormous electrostatic energy barrier (9, 17). For example, in the absence of counterions, the electrostatic repulsion encountered in the folding of the 400-nt Tetranychus self-splicing intron RNA would correspond to ~600 kcal/mol, approximately a thousand times the thermal energy $k_B T$ (0.62 kcal/mol at 37°C) (17).

Counterions reduce the electrostatic repulsion, which is referred to as screening. The concentration and charge of the cation(s) present greatly affect the extent of charge screening and thus whether an RNA molecule will fold or whether a protein will bind a stretch of DNA. Most functional RNA molecules will not fold under low salt conditions (18) and, indeed, addition of Mg$^{2+}$ continues to be the most common way to induce folding in RNA biochemical experiments. Conversely, favorable electrostatic interactions due to formation of DNA–protein and RNA–protein complexes are reduced at higher salt concentrations.

Why Are the Effects of Screening Ions So Difficult to Experimentally Dissect?

The fluctuating nature of the ion atmosphere around nucleic acids has made quantitative experimental studies challenging. In particular, the ion atmosphere is (mostly) invisible to traditional structural biology techniques, such as X-ray crystallography, NMR, or (cryo–) electron microscopy. Furthermore, folding events or conformational transitions induced by changing the ion atmosphere also involve other physical phenomena, such as tertiary structure formation and ion binding, so that the specific effects of the atmosphere typically cannot be readily isolated, quantified, and understood.
Figure 1

Structures of (a) three typical proteins, (b) the DNA mimetic pentapeptide repeat protein MfpA, and (c) three nucleic acids. Nucleic acids are highly negatively charged compared with proteins. Negatively charged residues are shown in red, positively charged residues in blue, polar residues in green, and nonpolar residues in white. The proteins shown are cytochrome \( c \) [Protein Data Bank (PDB) identifier 1CRC, 188 residues, and total estimated charge \(+9e\)], ribose binding protein [PDB 1URP, 271 residues, total charge \( -2e \)], and human serum albumin [PDB 1E7H, 585 residues, total charge \( -15e \)]. (b) The pentapeptide repeat protein MfpA [PDB 2BM4, 186 residues, total charge \( -5e \)] inhibits gyrase by binding as a DNA mimetic (191, 192) despite having only a modest overall negative charge. (c) The nucleic acids shown are double-stranded DNA (dsDNA) [PDB 2BNA, 24 residues, total charge \( -22e \)], double-stranded RNA (dsRNA) [adapted from PDB 3CIY, 30 residues, total charge \( -28e \)], and the P4–P6 domain from the \( Tetrahymena \) group I intron ribozyme [PDB 1GID, 158 residues, total charge \( -157e \)].

Can the ion atmosphere be identified though X-ray crystallography? X-ray crystallography is a powerful tool to study the specific contacts made by ions and nucleic acids. RNA crystal structures often reveal bound ions that can be assigned in the crystallographic electron density map (19–29). The crystallographically resolved ions are typically bound to specific sites on the RNA molecule, where they are at least partially dehydrated and often in close
proximity to one or several phosphoryl groups (Figure 2) (19, 20, 23, 25, 30). However, although high-resolution RNA structures have provided invaluable details about local RNA–ion interactions, they do not describe the ion atmosphere. Using the very simple rule of overall charge neutralization (see the section below titled Are There Any Universal Principles Governing Nucleic Acid–Ion Interactions?), one can see that cations identified in X-ray structures account for only a small fraction of the positive charge that is present; there must be other ions in the ion atmosphere (Figures 2 and 3). Apparently these ions form a fluctuating and extended layer around the DNA or RNA that cannot be captured by crystallography (1, 2, 10, 31–33). It is also worth noting that crystallization conditions can affect the position and occupancy of metal ions. For example, the same position can be occupied by different ions in different crystal structures of the same RNA (19, 34–37), and ion occupancy in solution may be different from that in the crystal structure.

Why Are the Effects of Screening Ions So Difficult to Model?

Positive ions are attracted to nucleic acids due to their negative charge, and negative ions are repelled. As a consequence, the local environment of a nucleic acid in solution has a higher concentration of positive ions than the unperturbed bulk ion concentration far away from the nucleic acid (counterion accumulation) and a lower local concentration of negatively charged ions (coion depletion). At this level, electrostatic interactions are simple.

But despite a seemingly simple basis in charge repulsion and attraction, modeling nucleic acid–ion interactions presents a formidable theoretical challenge. There are a very large number of ions, and they can occupy many different positions. Furthermore, interaction with each of the other ions must be considered in addition to interactions with the nucleic acid. In modeling approaches, this complexity is exacerbated by the long-range nature of electrostatic interactions (38) given

---

**Figure 2**

Crystal structures reveal details of ion–nucleic acid interactions but fail to give a complete picture of the ion atmosphere. (a) Crystal structure of the P4–P6 domain of the *Tetrahymena* group I intron ribozyme [Protein Data Bank (PDB) identifier 1GID] (21). Twelve crystallographically resolved Mg$^{2+}$ ions are shown as blue spheres. (b) Crystal structure of the tandem aptamer of a glycine riboswitch from *Fusobacterium nucleatum* (PDB 3P49) (195). The two glycine molecules bound to the tandem aptamer are shown in pink. Thirteen crystallographically resolved Mg$^{2+}$ ions are shown as blue spheres. (c) Close-up of the metal ion core of the P4–P6 RNA. Metal ions (spheres) interact with the adenine (A)-rich bulge, making a number of contacts and organizing the RNA residues. (d) Detailed rendering of the glycine binding pocket of aptamer I of the structure in panel b. The two crystallographically resolved Mg$^{2+}$ ions (spheres) coordinate a series of backbone residues in the binding pocket as well as the ligand glycine. These RNAs have 158 and 169 residues, corresponding to total charges of the RNAs of $-157$ and $-168$, respectively. The structures resolve 12 and 13 Mg$^{2+}$, respectively, which means that only ~15% of the ion atmosphere required for charge neutrality is accounted for by the explicitly resolved ions. In addition, the crystallographically resolved ions can be replaced by ions of other identities; both RNAs can undergo partial folding in high concentrations of monovalent ions (62, 83). Docking of the P5abc metal ion core in the P4–P6 RNA (c) and glycine binding (d) and folding to the native state for the tandem aptamer riboswitch require divalent ions, however, with limited specificity between different divalent species (52, 62, 103).
electrostatic repulsion solute due to negatively charged negative ions around a concentration of decreased local coion depletion:

electrostatic attraction solute due to negatively charged positive ions around a concentration of increased local accumulation:

Counterion measured set (by dialysis) or can be experimentally nucleic acids, which far away from the concentration of ions concentration:

Counterion accumulation: increased local concentration of positive ions around a negatively charged solute due to electrostatic attraction

Coion depletion: decreased local concentration of negative ions around a negatively charged solute due to electrostatic repulsion

\[ U_{1,2} = \frac{1}{4\pi\varepsilon_0\varepsilon} \frac{q_1 q_2}{r} \]

For a sense of scale, the dielectric constant is 80 in water, and the repulsive energy of two phosphoryl groups, each with charge \(-1e\) at the distance typical of nearest-neighbor nucleotides (7 Å), is 0.6 kcal/mol. This value—for just a single pair of nucleic acid groups—is already close to the thermal energy \(k_B T\), so that removing it (by screening) or doubling it would have a nonnegligible effect on the energetic balance between nucleic acid conformations. When summed over all pairs of phosphoryl groups, counterions, and coions in an RNA or DNA system, the total electrostatic interaction energies are enormous. Further complicating the analysis of nucleic acid–ion behaviors and energetics is the presence of large ensembles of conformations of the nucleic acids of interest. Thus, in addition to determining the nucleic acid–ion energetics of one nucleic acid conformation, any modeling effort must repeat this calculation many times, for each of the nucleic acid conformations that represent the overall ensemble of conformers present (39–43).

Finally, consideration of the ions themselves is highly complex. Whereas some of the ions remain solvated essentially as present in a normal salt solution, others closely approach the nucleic acid, resulting in polarization and solvation changes; and some of the closest approaching ions become partially desolvated and occasionally fully or nearly fully desolvated (1, 19, 20, 23). Situations like these are not well approximated by a simple continuum dielectric constant as in Equation 1, and state-of-the-art simulations do not reproduce the polarization effects of ions or water molecules (44–48). Additional discussion of modeling issues appears below.

**Are There Any Universal Principles Governing Nucleic Acid–Ion Interactions?**

As described above, the interactions of ions and nucleic acids can be extraordinarily complicated and in some cases extremely sensitive to experimental conditions or to model. Fortunately, there is one powerful concept that universally holds and can guide our efforts to understand these electrostatic interactions: charge neutrality. The sum of all positive and negative charges in solution must equal to zero (1, 8, 33, 49). Even small deviations from charge neutrality give rise to very large Coulomb forces that

---

**Figure 3**

Statistical analysis of ions in nucleic acid crystal structures. Distribution of the fraction of nucleic acid charges that are neutralized by ions resolved in the crystal structures of 6,238 nucleic acids obtained from the nucleic acid data base (http://ndbserver.rutgers.edu/). Nucleic acid charges were estimated by counting atoms of type P (phosphorus); ion charges were estimated by counting all elemental ions of the first and second groups as well as Mn\(^{2+}\), Fe\(^{2+}\), Cu\(^{2+}\), Ni\(^{2+}\), Cu\(^{3+}\), and Zn\(^{2+}\). The “fraction neutralized” is defined as the total positive charge due to the ions divided by the total negative charge of the nucleic acid. More than half of all crystal structures of nucleic acids do not resolve any ions, and in the structures that do report ions, the total charge of the crystallographically resolved ions, on average, accounts for only \(\sim 20\%\) of the total nucleic acid charge.

---

For a sense of scale, the dielectric constant is 80 in water, and the repulsive energy of two phosphoryl groups, each with charge \(-1e\) at the distance typical of nearest-neighbor nucleotides (7 Å), is 0.6 kcal/mol. This value—for just a single pair of nucleic acid groups—is already close to the thermal energy \(k_B T\), so that removing it (by screening) or doubling it would have a nonnegligible effect on the energetic balance between nucleic acid conformations. When summed over all pairs of phosphoryl groups, counterions, and coions in an RNA or DNA system, the total electrostatic interaction energies are enormous. Further complicating the analysis of nucleic acid–ion behaviors and energetics is the presence of large ensembles of conformations of the nucleic acids of interest. Thus, in addition to determining the nucleic acid–ion energetics of one nucleic acid conformation, any modeling effort must repeat this calculation many times, for each of the nucleic acid conformations that represent the overall ensemble of conformers present (39–43).

Finally, consideration of the ions themselves is highly complex. Whereas some of the ions remain solvated essentially as present in a normal salt solution, others closely approach the nucleic acid, resulting in polarization and solvation changes; and some of the closest approaching ions become partially desolvated and occasionally fully or nearly fully desolvated (1, 19, 20, 23). Situations like these are not well approximated by a simple continuum dielectric constant as in Equation 1, and state-of-the-art simulations do not reproduce the polarization effects of ions or water molecules (44–48). Additional discussion of modeling issues appears below.

**Are There Any Universal Principles Governing Nucleic Acid–Ion Interactions?**

As described above, the interactions of ions and nucleic acids can be extraordinarily complicated and in some cases extremely sensitive to experimental conditions or to model. Fortunately, there is one powerful concept that universally holds and can guide our efforts to understand these electrostatic interactions: charge neutrality. The sum of all positive and negative charges in solution must equal to zero (1, 8, 33, 49). Even small deviations from charge neutrality give rise to very large Coulomb forces that
strongly attract charged species such that neutrality is restored at equilibrium. (Recall that charge separation across membranes is used by cells to store and transduce energy and can be the predominant component of the proton-motive force used to synthesize ATP.) Applied to nucleic acids, charge neutrality means that overall counterion accumulation and coion depletion exactly balance the charges on the nucleic acid; in other words, the sum of all charges from the nucleic acid and its ion atmosphere is zero. This turns out to be a strong and non-trivial constraint. (Analogous constraints do not exist for the association of neutral metabolites to macromolecules.)

BE-AES: buffer equilibration–atomic emission spectroscopy

Recent experiments enable an accounting of the full content of the ion atmosphere. The approach involves buffer equilibration–atomic emission spectroscopy (BE-AES) (Figure 4c) and is commonly referred to as “ion counting” (51–53). BE-AES relies on carefully equilibrating the buffer around a nucleic acid sample and on quantifying the elemental composition of the solution using atomic emission spectroscopy (51, 53), that is, counting how much of each element is present in solution. By comparing the ion concentration in the nucleic acid containing sample with the flow-through, buffer-only sample, it is possible to count the total number of excess ions that are present in the ion atmosphere (Figure 4c). This approach is able to achieve high precision because all ions are quantified from a single sample (via their separate atomic emission lines), and each sample can be normalized by the amount of nucleic acid present by quantifying the phosphorus atoms.

Most fundamentally, BE-AES has experimentally confirmed overall charge neutrality (Figure 4d) (51). In addition, as described below, it has been used to characterize properties of the ion atmosphere and to test predictions from electrostatic theories. Other counting approaches, utilizing ion-specific fluorescent dyes, have been used to count specific ions present in the ion atmosphere and how they vary with conditions and conformational transitions but cannot assess the atmosphere in its entirety (33, 52–56).

Does Cation Accumulation Fully Account for the Nucleic Acid–Ion Atmosphere?

We emphasize above that the ion atmosphere achieves charge neutrality. Naively one might have expected then that the sum of charges from the cations in the atmosphere is equal to the charge on the nucleic acid that it surrounds. However, this is not the case for a simple reason: In addition to accumulating cations in the
Figure 4

Quantifying the total composition of the ion atmosphere. (a) Excess ion density as a function of distance from a nucleic acid, approximated here as a cylinder of radius 10 Å, determined from Poisson–Boltzmann theory. The ion atmosphere does not penetrate into the cylinder; the ion density is highest close to the nucleic acid and equals the bulk concentration far away from the nucleic acid. (b) The excess numbers of ions determined from a molecular dynamics simulation and the total charge of these ions around a 24-bp DNA duplex. The total charge of the ion atmosphere is equal and opposite to the charge of the DNA of $-46e$ (dashed line). The example shown is for 5 mM MgCl$_2$ and 40 mM NaCl. Data were taken from Reference 183. (c) Scheme of the buffer equilibration–atomic emission spectroscopy (BE-AES) approach that enables complete quantification of the total ion atmosphere (51, 53). (d) The number of associated excess Na$^+$ ions (blue circles), depleted Cl$^-$ ions (orange triangles), and their total charge (green squares) around a 24-bp DNA duplex determined by BE-AES. The total charge of the ion atmosphere is equal, to within experimental error, to the inverse of the DNA charge (dashed line). Data are from Reference 51.
atmosphere, there is “negative binding” of anions (also called coions), or preferential exclusion of anions from the atmosphere. This exclusion makes physical sense in that negatively charged coions are repelled from the negatively charged nucleic acid in solution. Thus, the total charge of the accumulated cations does not match the full charge of the nucleic acid.

What is harder to predict is the scale of this coion exclusion effect. For weakly charged objects (such as most proteins) in monovalent salt, simple theories predict that the number of coions excluded from the object is equal to the number of counterions brought in to screen the charge. However, nucleic acids are highly charged, and most theories about so-called polyelectrolytes predict that the number of counterions brought into the resulting atmosphere exceeds the number of excluded coions because of the strong attraction to the polyelectrolyte. In qualitative agreement with this intuition, BE-AES experiments indicate measurable exclusion of anions but much less than the accumulation of cations of the same valency (Figure 4d, shown schematically in Figure 5) (51).

What Is the Shape of the Ion Atmosphere?

Names like ion atmosphere, ion cloud, or ion sheath used here and in the literature denote a group of ions that is fluctuating but localized near the surface of the nucleic acid. Most theories predict that a region of high counterion density is localized to approximately ~10 Å around the nucleic acid at typical ion concentrations. This corresponds to 1–10 M cation concentrations to achieve charge neutrality (57). Although the BE-AES experiments above have confirmed overall charge neutrality and other basic expectations (51), those
Anomalous small-angle X-ray scattering (ASAXS) can probe the shape and spatial extent of the ion atmosphere. It uses the differential scattering properties of individual atomic species at different X-ray energies and thus can determine the ion–nucleic acid scattering for the probed (or isolated) ion. The anomalous scattering signal is related to the sum of the scattering from all pairs of atoms of the probed ion and the nucleic acid. ASAXS has revealed several basic properties of the ion atmosphere. There is a tighter distribution of divalent cations than monovalent cations around a nucleic acid (58, 59). This difference is predicted based on the stronger attractive forces for the divalent cations (see the next section). The overall scattering signals revealed about half as many ions for an atmosphere made up of divalent cations than for an atmosphere made up of monovalent cations, as expected based on the amount of charge needed to achieve charge neutrality (58, 60).

What Is the Relative Screening from Divalent and Monovalent Cations?

Most biochemical experiments in in vitro, cellular, or viral contexts have mixtures of divalent and monovalent cations. Biochemists are accustomed to maintaining a constant ionic strength, $I$, to maintain constant charge screening, using the simple relationship of the sum of each ion’s concentration weighted by the square of its valency [e.g., $I = [\text{Na}^+]·(+1)^2 + [\text{Mg}^{2+}]·(+2)^2 + [\text{Cl}^-]·(-2)^2$]. Ionic strength is related to how sharply the electric field is screened as a function of distance from a weakly charged point source. However, nucleic acids are strongly charged. For these systems the approximation of a point charge does not hold, and equal ionic strength does not correspond to equal screening by their atmospheres; there is no universal relationship for the degree of screening as a function of ion concentrations. Nevertheless, certain qualitative statements can be made, as elaborated in the following section.

What Is the Relative Ion Atmosphere Occupancy by Divalent and Monovalent Cations?

In terms of balance of monovalent and divalent ions present within the atmosphere, simple polyelectrolyte theories predict—and BE-AES experiments have confirmed—that divalent ions constitute a larger fraction of the ion atmosphere than would be predicted from the simple mole fraction present in bulk (Figure 5). For example, in a background of 20 mM Na$^+$, only $\sim0.5$ mM Mg$^{2+}$ is needed to replace half of the Na$^+$ ions in the ion atmosphere of a simple DNA duplex (51). As a consequence, nearly all of the accumulated counterions are divalent if mono- and divalent ions are present in approximately equal concentrations (Figure 5). A rough empirical rule of thumb is that $\sim10$ mM Mg$^{2+}$ gives similar behavior to $\sim1$ M Na$^+$, for example, in RNA folding in cases without specific ion binding (61, 62) or for DNA duplex stability (63–65), but this is far from exact and presumably depends on the properties and charge density of the particular nucleic acid or nucleic acid conformation.

The exact composition of a nucleic acid’s ion atmosphere and the relative energetics of different nucleic acid conformations depend on the bulk concentrations of all ions. These properties cannot yet be accurately predicted from theory.

ENERGETIC EFFECTS OF THE ION ATMOSPHERE

How Strong Is the Electrostatic Interaction Between Nucleic Acid Helices?

In the absence of counterions, the strength of electrostatic interactions is large, even between two phosphate groups. At the level of nucleic acid helices, the repulsive energy between two 10-bp double helices with their point of closest contact within 10 Å would be $\sim60$ kcal/mol ($\sim100$ times the thermal energy) in the absence of an ion atmosphere. This is substantially larger than the scale of energies from RNA or
The interplay of nucleic acid structural ensembles and ion interactions. The tethered duplex system has been used as a model system to study how ion interactions modulate the conformational ensemble of nucleic acids (17, 40, 41). (a,b) Sequence and schematic of the tethered duplex system. The system consists of two 12-bp DNA duplexes (red) joined by a flexible polyethylene glycol tether (green). (c) Visualization of the computationally derived ensemble of the tethered duplex system at various ionic conditions, from left to right and top to bottom: 0.02, 0.06, 0.17, 0.3, 2.0 M monovalent ions; the last image (bottom right) shows the ensemble in the absence of electrostatics (i.e., steric effects only). One duplex is rendered in gray, and the colored balls represent the distal end of the other duplex. Colors represent the energetic difference between the conformer and the minimum-energy conformer observed in the ensemble, from red (<1 kBT) to blue (>3 kBT). At a low salt concentration, electrostatic repulsion leads to repulsion between the duplexes; at a higher salt concentration, the repulsion is reduced, and a larger conformational ensemble is explored. Adapted with permission from Reference 40. Copyright 2008, American Chemical Society.

DNA base pairing and stacking (1–3 kcal/mol). The ion atmosphere reduces this electrostatic repulsion. Theoretical calculations (see the section below titled Can We Compute Ion Atmosphere Properties and Energetics from Poisson–Boltzmann Theory?) indicate that this repulsion is decreased enormously (very roughly 10-fold to ~6 kcal/mol) in 150 mM monovalent salt, and even further in millimolar concentrations of divalent ions. The energetic estimates above come from theoretical calculations that have been difficult to test in the context of nucleic acid folding events. Important work in DNA liquid crystals has successfully measured the energies versus the distance between long helices but in the context of a complex, condensed phase (66–69). To isolate electrostatic energetics, we used a simple model system: two DNA duplexes tethered by a short polyethylene glycol (PEG) linker (Figure 6a,b) (17, 40). The preformed helices have known structures and greatly reduced the number of conformations present, and the use of a PEG instead of nucleic acid linker removed the potentially complicating features of linker electrostatics, stacking, and hydrogen bonding. We followed the conformational ensemble of states by small-angle X-ray scattering (SAXS); compared the SAXS results with predictions from PB electrostatic calculations for each ionic condition used; and coupled that with the conformational entropy from the PEG tether, determined by extensive molecular dynamics (MD) simulations. (As the PEG tether has only rotatable C–C and C–O bonds, it represents the simplest challenge for MD, much simpler than with proteins and nucleic acids.)

With an increasing concentration of salt, the tethered duplex relaxed from a highly repelled state (Figure 6c). Quantitative comparisons were made to SAXS profiles predicted from the PB-MD modeling. [A strength of SAXS is that experimental scattering profiles can be directly predicted from a given structure (70, 71).] These comparisons revealed similar behavior.
Can RNA Folding Energies Be Estimated from Mg$^{2+}$ Titrations?

A quantitative description of how an RNA folds requires estimating the transition’s folding free energy ($\Delta G$) and how it changes with mutations ($\Delta \Delta G$ values). $\Delta G$ is related to the ratio of folded and unfolded molecules. But in typical experiments that explore RNA catalysis or function, RNAs are well folded, and the unfolded population cannot be easily isolated or measured. Fortunately, RNAs often attain their functional tertiary structures upon addition of Mg$^{2+}$ (72–77). Analogous to estimating protein $\Delta G$ values through temperature or chemical denaturant titrations, RNA $\Delta G$ values can, in principle, be estimated from precisely characterizing folding fractions near Mg$^{2+}$-induced midpoints and extrapolating to the standard solution conditions used for functional experiments. Using Mg$^{2+}$ is a particularly convenient choice for RNA because, unlike temperature or denaturant, Mg$^{2+}$ titrations often isolate RNA tertiary structure formation events with limited or no change in RNA secondary structure. Unfortunately, there remains substantial confusion in interpreting and analyzing RNA Mg$^{2+}$ titrations, with many complexities arising from effects of the ion atmosphere, as elaborated below.$^1$

Experimental data monitoring the ion-dependent folding or function of RNA molecules [e.g., chemical footprinting data, the radius of gyration, or activity as a function of [Mg$^{2+}$] (72–77, 80–84)] are often well described by a Hill equation (85) in the form of Equation 2,

$$f_r = \frac{1}{1 + (K_{\text{mid}}/[\text{Mg}^{2+}])^n},$$

and Equation 3,

$$\text{RNA}_U + n\text{Mg}^{2+} = \text{RNAs} \bullet (\text{Mg}^{2+})_n,$$

where $f_r$ is the fraction folded at a given Mg$^{2+}$ concentration, $K_{\text{mid}}$ is the Mg$^{2+}$ midpoint, and $n$ is the Hill coefficient. With integer coefficients, the Hill model corresponds to a two-state equilibrium in which one state has $n$ more ligands bound than the other. Consequently, fits to Equation 2 have been described in terms of $n$ Mg$^{2+}$ ions binding to the RNA, the equilibrium represented by Equation 3. But this interpretation can be misleading. First, as mentioned above, ion–RNA interactions are governed by charge neutrality of the ion atmosphere. Thus, there are many more ions associated with the RNA than the $n$ ions implicated from Equation 2, although the Hill coefficient $n$ does read out the average difference in ions of the type being varied between the folded and unfolded state at the Mg$^{2+}$ atmosphere.

$^1$A major complication that occurs independently of the ion atmosphere (and, indeed, occurs widely for protein systems) is the possibility of multiple folding transitions that are mistakenly lumped together by a low-resolution experimental observable (non-two-state behavior). For example, native gel electrophoresis or X-ray scattering can detect the compaction of RNAs as a function of Mg$^{2+}$ but can miss folding of subdomains, leading to inconsistencies in thermodynamic models. Information-rich folding readouts, such as chemical footprinting, can resolve folding transitions and give self-consistent thermodynamic models; for an example, see Reference 73.
midpoint (8–10, 56, 86, 87). At an extreme, \( n \) can be a small fraction of unity, but the actual RNA molecule binds hundreds of Mg\(^{2+}\) ions in both the unfolded and folded states. In addition, and also because of charge neutrality, other ions are released for every Mg\(^{2+}\) ion that associates with RNA as it folds. Such an exchange process upon RNA folding is conceptually distinct from a typical ligand-binding event. Finally, observation of Mg\(^{2+}\) -induced RNA folding generally does not imply that the Mg\(^{2+}\) binds to a specific site on the folded RNA; the folding can be (and is often) due to the additional screening from atmospheric Mg\(^{2+}\). For example, the formation of a tetraloop–receptor tertiary contact in the *Tetrahymena* ribozyme and in isolated systems has been extensively studied via Mg\(^{2+}\) titration. However, there is no evidence for site binding to this RNA, and the same interaction forms stably in the presence of sufficiently high monovalent ion concentrations (61, 78, 88).

Beyond conceptual confusion, serious problems arise when using Hill fits to estimate folding free energies, \( \Delta G \), and free-energy differences, \( \Delta G \), between mutants. The Hill coefficient is often assumed to be a constant independent of [Mg\(^{2+}\)]. This assumption a priori seems to be the simplest choice but, in fact, holds only in a few RNA systems. The number of Mg\(^{2+}\) ions, even around a single nucleic acid structure, changes with ion concentration (Figure 5), and because the Hill coefficient reads out the difference of such Mg\(^{2+}\) numbers between unfolded and folded states, this difference is likely to vary with the bulk ion concentrations. Changes of structure within each state (e.g., upon electrostatic relaxation of an unfolded state) (Figure 6) provide additional sources for Hill coefficient variation. Beyond these theoretical issues, there are the experimental observations: For example, RNA mutants that require higher Mg\(^{2+}\) concentrations to fold also often are fitted to apparent Hill coefficients that differ from the wild type (75, 80–82).

If the Hill coefficient is not constant, how then can one extrapolate \( \Delta G \) from the midpoint of a Mg\(^{2+}\) titration up to the Mg\(^{2+}\) concentration used for functional experiments? How can one compare \( \Delta G \) between mutants if they give different Hill coefficients? First of all, \( \Delta G \) should not be extrapolated outside the range of the actual measurements. For example, it is dangerous to use a standard state of 1 M Mg\(^{2+}\) when folding fractions can only be measured precisely at low millimolar Mg\(^{2+}\) concentrations. Such extrapolations amplify errors in Hill slopes to produce quantitatively or qualitatively wrong conclusions about energetic effects of mutations. Instead, an experimenter can use a thermodynamic model with a varying Hill coefficient that matches the observed Hill coefficient for the measured mutants and assumes a \( \Delta G \) that is independent of Mg\(^{2+}\) concentrations. For example, if the value of \( n \) obtained from each mutant varies linearly with each mutant’s midpoint, a simple expression can be applied as an approximation in Equation 4 (79, 87, 89):

\[
\Delta G = -n_{\text{WT}} + n_{\text{mutant}} \frac{1}{2} k_B T \ln \left( \frac{K_{\text{WT}}}{K_{\text{mutant}}} \right).
\]

where \( n_{\text{WT}} \) and \( n_{\text{mutant}} \) are the apparent Hill coefficients measured at the folding midpoints \( K_{\text{WT}} \) and \( K_{\text{mutant}} \) for the wild-type and mutant RNAs, respectively. The assumptions behind this and other models are given in Reference 79. If the apparent Hill coefficient \( n \) does not vary smoothly in the experiments, or if independent information indicates that \( \Delta G \) may vary strongly with [Mg\(^{2+}\)], then the experimenter must find simplifying solution conditions (see the below section titled Can We Get Rid of the Ion Atmosphere?) or use methods that permit measurement of \( \Delta G \) for all the mutants under identical experimental conditions. For the latter route, it is now feasible to use single-molecule approaches to measure unfolded fractions of \( \sim 10^{-2} \) and thus make
unambiguous comparisons for RNA variants with folding free-energy differences of up to $k_B T \cdot \ln(10^{-2}) \sim 3$ kcal/mol (90, 91).

**Can Direct Binding of Mg$^{2+}$ Be Separated from the Effects of Atmospheric Mg$^{2+}$?**

The direct or specific binding of RNA with divalent metal ions, as opposed to its association with a fluctuating ion atmosphere, has been difficult conceptually and problematic experimentally. On the one hand, the vast majority of Mg$^{2+}$ titrations following RNA folding or function have used simple site-binding models described by the Hill equation (see the previous section). On the other hand, it has been stated that specific metal ion binding to RNA in inner sphere complexes is extremely rare because the cost of ion desolvation is so high. Both misrepresent the physical and energetic nature of these binding and association events.

When metal ions site-specifically bind to small molecules, proteins, or nucleic acids, one or more water molecules that solvate the metal ion are lost but are replaced by ligands in the binding site. Although it is true that there is an enormous cost to stripping a solvating inner-sphere water molecule from a metal ion in solution, there is a similar enormous “payout” in replacing that water molecule with a functional group from the receptor molecule. Consider that Mg$^{2+}$ ions lose their solvating water molecules to bind ethylenediaminetetraacetic acid (EDTA), yet bind with nanomolar affinity. Many factors are involved in determining the position and affinity of a bound metal ion, including the ligand type and charge, as well as whether there are nearby ligands positioned or easily rearranged to provide a multidentate binding. There are well-documented cases of site-specific binding of metal ions to RNA (24, 92–98). Strong positive evidence for functionally important binding of Mg$^{2+}$ to a specific RNA site can be obtained from experiments involving thio or amine substitution for oxygen ligands to weaken the binding of Mg$^{2+}$ and allow rescue of function with “soft” metal ions like Mn$^{2+}$ (99–101). Inconsistent thermodynamic or kinetic frameworks have caused widespread confusion in the metal ion rescue literature, but quantitative modeling—with error analysis—can now be achieved to expose and eliminate such problems and limitations (79; also see the above section titled Can RNA Folding Energies Be Estimated from Mg$^{2+}$ Titrations?).

**Can We Get Rid of the Ion Atmosphere?**

Given the complexities introduced by the ion atmosphere of understanding nucleic acid energetics, it would seem desirable to find conditions in which the atmosphere is simply removed, so that a nucleic acid can be studied independent of its complicating effects. However, in accord with the principle of charge neutrality, “disappearing” the atmosphere is not possible. The net charge around a negatively charged nucleic acid must total zero, and only ions can accomplish this. In particular, dialysis into deionized water of nucleic acid samples, such as RNA purified from cells or viruses, does not strip away ions associated with the nucleic acids. Instead, such dialysis can result in the opposite of what is desired. The most strongly charged counterions, polyamines, basic peptides, or, worse, unknown species associate with the nucleic acid to keep it neutral. A better strategy to purify nucleic acid samples is to dialyze into buffers that contain known concentrations of known salts. Similarly, experimental conditions or discussions that refer only to having MgCl$_2$ present or titrating in Mg$^{2+}$ ignore the preexisting ion atmosphere and the inability of the classical site binding and Hill analysis to describe ion effects with nucleic acids (see the above section titled Can RNA Folding Energies Be Estimated from Mg$^{2+}$ Titrations?).

Although the ion atmosphere around a nucleic acid can never be removed, the atmosphere’s effects can be minimized through a counterintuitive strategy: using conditions with very high concentrations of monovalent salt, such as 2 M NaCl, to “swamp” the atmosphere. Under such conditions, the atmosphere composition is dominated by the monovalent ions. As
a further simplification, the atmosphere gives efficient screening between phosphate groups, so that electrostatic forces between nucleic acid segments can, to a first approximation, be neglected. Under these conditions, if titration of Mg\(^{2+}\) leads to a structural change, that transition would most likely correspond to specific binding of an Mg\(^{2+}\) ion; the alternative, better electrostatic screening of the folded state by atmospheric Mg\(^{2+}\), is unlikely due to the dominance of monovalent ions in the atmosphere. In high monovalent salt conditions, Mg\(^{2+}\) titrations become conceptually closer to standard ligand binding. Indeed, spectroscopic and biochemical identification of site-bound divalent ions, Hill fits that can be simply interpreted as the number of site-bound metal ions, and high-precision ion counting of those ions become feasible (52, 62, 97, 102, 103). A current, unmet challenge is to connect these unambiguous structural insights and energetic measurements under artificially high monovalent salt backgrounds to more standard conditions (79).

THE STATE OF THEORY

Because the fluctuating, volatile nature of the ion atmosphere makes visualization and experimental dissection difficult, its predominant description has come from theoretical models. Importantly, any assessment of the thermodynamic consequences of the ion atmosphere, such as for RNA folding or protein–nucleic acid binding, requires a quantitative theoretical description of nucleic acid–ion interactions. An adequate theoretical description should make quantitative and nontrivial predictions, and, in turn, there is a need for experimental benchmarks against which to test theoretical predictions. Here, we discuss commonly used models of the ion atmosphere and evaluate them against the available experimental benchmarks.

Does Manning Theory Account for Ion Atmosphere Effects?

A simple and yet influential model for DNA–ion interactions was proposed by Manning decades ago (104–106). Manning calculated that DNA and other polyelectrolytes with sufficiently high charge density induce a condensation of cations at the polyelectrolyte surface until its effective charge is reduced below a critical value. According to these simple electrostatic calculations, this condensed or bound layer is predicted to effectively neutralize \(\sim 60\% \)–70\% of the DNA backbone charge.

The strength (and, at the same time, the limitation) of Manning or so-called counterion condensation (CC) theory is its simplicity. Its consistency with a number of experiments that can be interpreted by assigning an effective charge suggests that it captures an important property of polyelectrolyte interactions (80, 107–114), but as with any simplified approach, there are also aspects of the actual system that are sacrificed. For example, atomic-level properties and interactions of the ions, solvent molecules, and hydration of the polyelectrolyte are not included. The theory also can imply a seemingly rigid distinction between a condensed layer and the rest of the ion atmosphere, whereas these together are responsible for full charge neutralization and form a continuum. On a practical level, the notion that full neutralization occurs in solution is often not realized by researchers who are aware of the predicted \(\sim 0.7\) neutralization by the condensed layer but lack a full understanding of this theory. Moving forward, concepts from Manning’s theory are being used in next-generation and more detailed theories (4, 11, 115, 116), and other approaches are also under development, as outlined in the following sections.

Can We Compute Ion Atmosphere Properties and Energetics from Poisson–Boltzmann Theory?

Poisson–Boltzmann (PB) theory is widely used to describe the ion atmosphere associated with nucleic acids and proteins (1, 12, 31, 33, 114). This theory has been used to quantitatively predict the composition and shape of the ion atmosphere around nucleic acids and has provided insight into the electrostatic component of the

---

**Poisson–Boltzmann (PB) theory:**
a differential equation describing the average (i.e., mean-field) electrostatic interactions of ions with a charged solute in solution

---
free-energy landscape governing RNA folding. Because PB theory has been the predominant ion atmosphere theory, we first introduce how PB theory models the ion atmosphere, then discuss its approximations and shortcomings, and finally evaluate its validity in light of recent rigorous experimental tests.

Poisson–Boltzmann theory background and equation. PB theory describes the interaction of mobile ions with fixed charges in solution (1, 33, 117–122). The fixed charge density, denoted $\rho_{\text{fixed}}(\vec{r})$, is typically defined by the structure of the protein or nucleic acid of interest. PB theory considers the situation where the charged molecule is surrounded by solvent, approximated as a continuum, with a dielectric constant $\epsilon$ and mobile ions of charge $z_i$ that interact with fixed charge density and with the average potential of all ions. This average potential of all ions is determined by their interactions of ions beyond that expected from the average ion concentration.

\[
\Phi(\vec{r}) = -\frac{\rho_{\text{fixed}}(\vec{r})}{\epsilon} - \sum_i c_i z_i \epsilon \exp\left(\frac{-z_i e \Phi(\vec{r})}{k_B T}\right)
\]

where $e$ is the elementary charge, $k_B$ is the Boltzmann constant, and $T$ is the absolute temperature. The term $c_i$ is the bulk concentration of ion species, $i$, namely the concentration far from the charged macromolecule. The term $\lambda(\vec{r})$ is an accessibility factor that defines the regions in space that are accessible to ions (where $\lambda = 1$) or inaccessible to ions (e.g., inside the macromolecule, where $\lambda = 0$). The exponential function from the Boltzmann factor (the last term in Equation 5) makes the equation nonlinear; consequently, Equation 5 is referred to as the nonlinear PB equation. For computational simplicity, a linearized version of the equation, referred to as the Debye–Hückel theory (125), is sometimes used but is not adequate for polyelectrolytes. The general approach is to compute $\Phi(\vec{r})$ by numerically solving Equation 5.

From the solution of the PB equation and $\Phi(\vec{r})$, all other quantities of interest can be determined. Most importantly, knowing the electrostatic potential at all points in space [i.e., $\Phi(\vec{r})$] allows the electrostatic contributions to free energies to be calculated for any process of interest, including association events, folding, and conformational changes. PB also allows calculation of the number of associated ions, and, as described below, this provides a powerful link to experiment.

Software packages for Poisson–Boltzmann calculations. A number of freely available software packages have been developed to solve the PB equation (Equation 5) numerically and have enabled widespread application of PB theory to macromolecular systems of interest. Examples of PB software packages include DelPhi (118, 126, 127), APBS (122, 128, 129) [with the useful PDB2PQR (130) interface], MEAD (131), and AQUASOL [employing the Langevin dipole PB approach, a scheme extending the PB framework by representing the solvent as self-orienting dipoles of variable density (44, 132–135)].

Limitations of the Poisson–Boltzmann theory. PB theory makes several simplifying approximations in its treatment of the ion atmosphere. First, it does not consider the ions as discrete objects but treats their interactions in an averaged or mean-field theory fashion. Thus, ion–ion correlations are neglected. Second, in the PB framework, ions do not have a discrete size and are only characterized by their valency or charge. The mean-field nature of PB theory also prevents other atomic-level properties, including ion polarizability and desolvation, from being considered. Similarly, it treats the aqueous environment and the interior of the nucleic acid as media with
uniform dielectric constants, $\varepsilon(\vec{r})$, neglecting their discrete nature and polarizability. Despite these limitations, PB theory has been widely applied to nucleic acid systems, in part because of its relative ease of implementation.

**Experimental tests of Poisson–Boltzmann theory.** Although PB theory introduced the concept of the ion atmosphere to many and provides a means to calculate electrostatic interaction energies, its use as a quantitative, predictive tool required testing. Incisive tests of any theory require experimental systems that are simple enough to avoid extraneous features, as these can complicate relating experiment to theory and introduce factors that can inadvertently serve as free fitting parameters. For example, early tests of tRNA folding demonstrated that PB models could account for $\text{Mg}^{2+}$-dependent folding but also incorporated assumptions about the nature of the unfolded state and included an adjustable parameter for the folding contribution from the formation of tertiary interactions (39, 55, 56, 62, 136–138).

Similarly, folding assays for group I ribozymes revealed that folding is much more efficiently induced by divalent than monovalent cations, as predicted by PB theory and other polyelectrolyte theories (32, 80, 139, 140), but the complexity of large RNAs precludes rigorous evaluation of the PB model. We therefore turned to simplified experimental systems. Much of this work has been reviewed, so we briefly describe it and refer readers to the prior review and the original papers (17, 33, 40–42, 51, 53, 141).

As noted above, BE-AES, or ion counting, provides a powerful window into the ion atmosphere. For PB theory, ion counting could test the effect of ion identity (e.g., size), as well as the competition of the monovalent and divalent cations for occupancy of the ion atmosphere. Simple DNA duplexes could be used to minimize accompanying folding events or conformational rearrangements. As the valency of ions is the only property taken into account in PB, all ions of the same valency, such as Li$^+$, Na$^+$, K$^+$, and Rb$^+$, are predicted to behave identically. Experimentally, the smaller ions modestly out-compete, by two- to threefold, the larger ions for ion atmosphere occupancy (33, 51). Some differences can be accounted for by incorporating additional terms into the PB free-energy functional, but a fully successful treatment is not yet available and may not be possible (141, 142). In addition, careful studies with more complex RNA systems reveal different effects from different monovalent cations, with trends that do not match the preferential ion atmosphere occupancy order for the simple helix, and these ion-specific effects differ from system to system (e.g., References 24, 143, and 144). These observations suggest that more specific interactions, even with monovalent cations, need to be taken into account to fully understand the ion atmosphere and RNA energetics.

For the physiologically important divalent ions, ion counting experiments with simple DNA helices demonstrate that the preferential divalent ion association occurs at two- to threefold smaller divalent ion concentrations than predicted from PB theory (51). In addition, SAXS-based experiments with the simple tethered two-helix model system, referred to above (Figure 6), revealed discrepancies in the divalent ion concentration dependences of conformational relaxation from charge screening of more than an order of magnitude, relative to PB predictions (33, 40). In contrast, the deviations from PB theory for relaxation induced by monovalent cations were modest. The stronger deviations for divalent cations are consistent with ion–ion correlation effects, as higher valency (and higher charge density) is expected to increase correlations in the positions of the ions with respect to one another (16, 57, 145, 146).

Overall, these and other results (58, 59, 114, 147, 148) suggest that PB theory provides a good guide for the general expected behavior of the ion atmosphere. However, it does not provide a quantitative description of the contents or energetics of that atmosphere. Deviations appear to arise from the fundamental limitations of the mean-field approximation, specifically that ions in PB theory lack size, relevant atomic or solvation properties, and spatial correlations with one another.
Overcharging: a phenomenon whereby a highly charged molecule (e.g., negatively charged DNA) attracts so many counterions that its overall charge changes sign

Can Expansions of Poisson–Boltzmann Fix Its Problems with Divalent Ions?

There are a number of approaches to modeling ion–nucleic acid interactions that go beyond PB theory and try to overcome its limitations. Some theoretical treatments stay close to the PB framework and enhance it by adding size-dependent terms (141, 142, 149–152) or an improved treatment of the solvent (44, 133–135, 153). Other approaches employ Monte Carlo schemes with an explicit treatment of (part of) the ions (13, 154–157). Such approaches are clearly necessary given the known failures for PB theory, in particular the lack of ion–ion correlations, for divalent and higher-valency ions. Currently, none of these models can be readily applied to a range of systems, and none have been subjected to stringent experimental tests. The development of a tractable theory that can accurately predict the properties of the divalent and higher-valency ions in the atmosphere remains an important outstanding challenge.

Can We Predict Specific Metal Ion Binding Sites?

Understanding the energetics of metal ion binding requires modeling not only the electrostatic interaction between the metal ion and the RNA moieties but also water desolvation, RNA–metal interactions mediated by discrete water molecules, electronic polarization of these moieties, and the effects of the other ions in the atmospheric background. The ability to predict these interactions remains limited. Currently, the simplest and best-defined metal ion prediction puzzle involves taking the exact atomic coordinates of an RNA from its crystal structure and predicting the locations of crystallographically observed metal ion sites. Even here, only ~50% of the crystallographic sites can be recovered by current knowledge-based algorithms (27, 170, 171). Predicting the number and location of site-bound metal ions in nucleic acids of unknown structure or in solution versus crystals is still more difficult and remains an unmet challenge. Experimental methods can, in favorable cases, identify bound metal ions and their binding sites, but care and caution are needed in carrying out and interpreting these experiments (79, 99, 101).

Can Molecular Dynamics Just Simulate the Ion Atmosphere?

In contrast to PB theory, MD simulations can treat the solvent explicitly by representing both water molecules and ions as discrete particles (172, 173). In addition, MD simulations naturally treat the macromolecule as flexible; in
other words, the positions of the atoms that make up the protein or nucleic acid are updated in the simulation, whereas PB-type approaches consider mostly a static macromolecular structure and require calculations for a series of discrete structures. As a consequence, MD simulations can potentially be more accurate than PB theory and related approaches.

However, several issues require careful consideration. First, the added detail (explicit solvent, flexible macromolecule) makes MD much more computationally costly than PB methods. Second, for realistic simulations, it is important to have a sufficiently large simulation volume and number of explicit ions such that the ions far from the nucleic acid in the simulation have bulk-like properties regarding concentrations, fluctuations, and ion relaxation. In particular, it is insufficient to only include just enough counterions to make the overall simulated system charge neutral. For example, Pappu and coworkers found that they had to include 800 ion pairs and 55,000 water molecules in addition to 30 neutralizing counterions in their simulation of the 32-residue Tar-Tar* RNA kissing loop motif.

MD simulations also, by necessity, involve a large number of parameters that describe the interactions (van der Waals, electrostatics, etc.) between different kinds of atoms. Several authors have recently pointed out that some of the popular choices of parameter sets used for ion simulations may be flawed and produce unphysical artifacts, such as the formation of salt clusters at concentrations well below the concentration limit or unphysical attraction and clustering of DNA. But changing parameters to prevent such artificial behavior does not ensure an ability to predict other behaviors. Physically realistic ion parameters are a prerequisite to correctly addressing more challenging problems, such as ion interactions with nucleic acids and nucleic acid–protein complexes. Recent efforts have been aimed at improving parameters for MD simulations by comparison with experimental results, such as osmotic pressures, crystal phase parameters, Raman spectroscopy, and BE-AES data.

Most basically, any computational approach requires rigorous experimental tests and predictions of nontrivial properties to evaluate the theory, and the need for computation in a truly unbiased fashion is becoming more and more apparent. The Critical Assessment of Protein Structure Prediction consortium revealed, through blind predictions of previously unknown protein structures, much greater limits in our ability to predict protein structures than previously recognized; multiple structure prediction algorithms, developed using previously determined structures with portions of these data set aside for testing and validation, failed in the face of truly blind predictions. Other communities, including the RNA structure community, are following the example set by CASP and are developing their own blind prediction tests.

For nucleic acid electrostatics, we have an emerging suite of experimental tools that can provide benchmarks for further developing computational approaches and, ultimately, rigorous tests of new predictions from those approaches. We anticipate a continued and highly integrated feedback loop of computation and experiment using ion counting, ASAXS, and energetic tests with simple nucleic acid systems that will greatly advance our understanding of and ability to predict the properties, behavior, and structural and energetic consequences of ions and the ion atmosphere.

**SUMMARY POINTS**

1. Nucleic acid folding, structure, and interactions are central to biology and cannot be understood without a deep appreciation of electrostatics.
2. Nucleic acids and other polyelectrolytes are surrounded by an ion atmosphere. Its properties are distinct from standard site binding, and it is extremely challenging to study and understand.

3. As many biochemists and molecular biologists are unfamiliar with the underlying physics and manifestations of the ion atmosphere, we have attempted to provide a basic conceptual and practical guide, discussing both current experimental and theoretical approaches used to probe and describe the ion atmosphere.

4. Theoretical descriptions of the ion atmosphere, including the Hill model, Manning condensation, PB theory, and Monte Carlo and MD approaches, have been used to explain the ion atmosphere and ion–nucleic acid interactions. Of these, the most common, the Hill model, is the most limited, and there is currently no theory that quantitatively accounts for all aspects of nucleic acid–ion interactions.

5. The principle of charge neutrality is a powerful concept. It serves as a guide for the interpretation of experiments and development of intuition about ion atmosphere behavior.

6. Ion counting and other experimental tools, using simple nucleic acid systems, have been critical in testing electrostatic theories and provide benchmarks for the future development of electrostatic theories and computational approaches for nucleic acid electrostatics.

7. Development of quantitative descriptions of nucleic electrostatics and energetics will require a tight experimental–computational feedback loop with rigorous evaluation of blind and nontrivial predictions.

8. Understanding of the ion atmosphere and its energetic manifestations is fundamental to understanding the functions and interactions of nucleic acids in biology and to developing an ability to engineer nucleic acids for use in cellular regulation and other processes.

FUTURE ISSUES

1. A strength of the BE-AES ion counting method is that all ions in essentially arbitrary samples can be quantified. Extensions of BE-AES to complex folded RNAs and to nucleic acid–protein complexes have the potential to dissect the ion atmosphere and help understand its energetic consequences in these more complex settings.

2. The adequate treatment of ions in MD simulations remains a challenge, and further tests are needed. Unanswered questions are how well classical simulation protocols based on fixed charges can correctly capture ion behavior in solution and what properties will ultimately require polarizable force fields or quantum mechanical treatments.

3. Development and rigorous testing of MD force fields for nucleic acid–ion interactions are needed to ultimately provide a predictive and complete understanding of RNA and DNA folding and interactions.
4. Investigators need to develop quantitative, first-principles models for specific ion binding to nucleic acids. To date, only simplified systems using very high monovalent cation concentrations to effectively saturate the ion atmosphere have yielded direct, quantitative information about site-binding stoichiometry. Future challenges are to carry out such analyses under physiological salt conditions and to account for the site interaction energies.

5. Understanding the effect of the ion atmosphere and ion binding on nucleic acid dynamics is lacking. In addition, we do not know the response time of the ion atmosphere to the dynamics of the nucleic acid and its binding partners, or to what extent that response time influences nucleic acid kinetic, dynamic, and binding properties. Also needed is an understanding of whether specific metal ion interactions help nucleate RNA structures, and whether establishment of these interactions can be rate limiting.

6. The overall goal is to develop a quantitative and predictive understanding of nucleic acid folding and structure, alone or with other interaction partners.

7. A critical aspect of this grand challenge is determining whether ion atmosphere effects can be treated as independent energetic contributors that can then be combined in an additive fashion with contributions from, for example, tertiary motif interactions, conformational preferences of helix junctions, and protein binding.

DISCLOSURE STATEMENT
The authors are not aware of any affiliations, memberships, funding, or financial holdings that might be perceived as affecting the objectivity of this review.

ACKNOWLEDGMENTS
We thank Aleksei Aksimentiev, David Garcia, Serge Lemay, and the members of the Doniach and Herschlag labs for useful discussions. We acknowledge the enormous amount of research in this area by many groups and regret, owing to space and scope limitations, that it was not possible to cite a large amount of fascinating and high-quality work. We received funding from the Netherlands Organisation for Scientific Research (NWO), the Burroughs Wellcome Fund (Career Award at the Scientific Interface), and National Institutes of Health grant PO1 GM066275.

LITERATURE CITED
34. Lipfert et al.
49. Record MT Jr, Anderson CF, Lohman TM. 1978. Thermodynamic analysis of ion effects on the binding and conformational equilibria of proteins and nucleic acids: the roles of ion association or release, screening, and ion effects on water activity. Q. Rev. Biophys. 11:103–78


85. Hill AV. 1910. The possible effects of the aggregation of the molecules of haemoglobin on its oxygen dissociation curve. J. Physiol. 40:4–7


## Contents

Journeys in Science: Glycobiology and Other Paths
*Raymond A. Dwek* .......................... 1

Lipids and Extracellular Materials
*William Dowhan* .................................. 45

Topological Regulation of Lipid Balance in Cells
*Guillaume Drin* ..................................... 51

Lipidomics: Analysis of the Lipid Composition of Cells and Subcellular Organelles by Electrospray Ionization Mass Spectrometry
*Britta Brügger* ....................................... 79

Biosynthesis and Export of Bacterial Lipopolysaccharides
*Chris Whitfield and M. Stephen Trent* .................. 99

Demystifying Heparan Sulfate–Protein Interactions
*Ding Xu and Jeffrey D. Esko* .......................... 129

Dynamics and Timekeeping in Biological Systems
*Christopher M. Dobson* .................................. 159

Metabolic and Nontranscriptional Circadian Clocks: Eukaryotes
*Akhilesh B. Reddy and Guillaume Rey* .................. 165

Interactive Features of Proteins Composing Eukaryotic Circadian Clocks
*Brian R. Crane and Michael W. Young* .................. 191

Metabolic Compensation and Circadian Resilience in Prokaryotic Cyanobacteria
*Carl Hirschie Johnson and Martin Egli* .................. 221

Activity-Based Profiling of Proteases
*Laura E. Sanman and Matthew Bogyo* .................. 249

Asymmetry of Single Cells and Where That Leads
*Mark S. Bretscher* ................................... 275

Bringing Dynamic Molecular Machines into Focus by Methyl-TROSY NMR
*Rina Rosenzweig and Lewis E. Kay* .................. 291
Chlorophyll Modifications and Their Spectral Extension in Oxygenic Photosynthesis
Min Chen ................................................................. 317

Enzyme Inhibitor Discovery by Activity-Based Protein Profiling
Micah J. Niphakis and Benjamin F. Cravatt ........................................ 341

Expanding and Reprogramming the Genetic Code of Cells and Animals
Jason W. Chin ............................................................... 379

Genome Engineering with Targetable Nucleases
Dana Carroll .................................................................. 409

Hierarchy of RNA Functional Dynamics
Anthony M. Mustoe, Charles L. Brooks, and Hashim M. Al-Hashimi ........................................ 441

High-Resolution Structure of the Eukaryotic 80S Ribosome
Gulnara Yusupova and Marat Yusupov .............................................. 467

Histone Chaperones: Assisting Histone Traffic and Nucleosome Dynamics
Zachary A. Gurard-Levin, Jean-Pierre Quivy, and Geneviève Ahmouzni ........................................ 487

Human RecQ Helicases in DNA Repair, Recombination, and Replication
Deborah L. Croteau, Venkateswarlu Popuri, Patricia L. Opresko, and Vilhelm A. Bohr ........................................ 519

Intrinsically Disordered Proteins and Intrinsically Disordered Protein Regions
Christopher J. Oldfield and A. Keith Dunker ........................................... 553

Mechanism and Function of Oxidative Reversal of DNA and RNA Methylation
Li Shen, Chun-Xiao Song, Chuan He, and Yi Zhang .............................................. 585

Progress Toward Synthetic Cells
J. Craig Blain and Jack W. Szostak .......................................................... 615

PTEN
Carolyn A. Worby and Jack E. Dixon ......................................................... 641

Regulating the Chromatin Landscape: Structural and Mechanistic Perspectives
Blaine Bartbolomew ........................................................................... 671

RNA Helicase Proteins as Chaperones and Remodelers
Inga Jarmoskaite and Rick Russell ........................................................... 697

vi Contents
Selection-Based Discovery of Druglike Macro cyclic Peptides
   Toby Pasioura, Takayuki Katoh, Yuki Goto, and Hiroaki Suga .................. 727

Small Proteins Can No Longer Be Ignored
   Gisela Storz, Yuri I. Wolf, and Kumaran S. Ramamurthi .......................... 753

The Scanning Mechanism of Eukaryotic Translation Initiation
   Alan G. Hinnebusch .................................................................................. 779

Understanding Nucleic Acid–Ion Interactions
   Jan Lipfert, Sebastian Doniach, Rbiju Das, and Daniel Herschlag .............. 813

Indexes

Cumulative Index of Contributing Authors, Volumes 79–83 ....................... 843

Cumulative Index of Article Titles, Volumes 79–83 ..................................... 847

Errata

An online log of corrections to Annual Review of Biochemistry articles may be found at
http://www.annualreviews.org/errata/biochem
New From Annual Reviews:

**Annual Review of Statistics and Its Application**

Volume 1 • Online January 2014 • http://statistics.annualreviews.org

Editor: **Stephen E. Fienberg**, Carnegie Mellon University
Associate Editors: **Nancy Reid**, University of Toronto  
**Stephen M. Stigler**, University of Chicago

The Annual Review of Statistics and Its Application aims to inform statisticians and quantitative methodologists, as well as all scientists and users of statistics about major methodological advances and the computational tools that allow for their implementation. It will include developments in the field of statistics, including theoretical statistical underpinnings of new methodology, as well as developments in specific application domains such as biostatistics and bioinformatics, economics, machine learning, psychology, sociology, and aspects of the physical sciences.

Complimentary online access to the first volume will be available until January 2015.

**TABLE OF CONTENTS:**

- A Systematic Statistical Approach to Evaluating Evidence from Observational Studies, David Madigan, Paul E. Stang, Jesse A. Berlin, Martijn Schuemie, J. Marc Overhage, Marc A. Suchard, Bill Dumouchel, Abraham G. Hartzema, Patrick B. Ryan
- The Role of Statistics in the Discovery of a Higgs Boson, David A. van Dyk
- Brain Imaging Analysis, F. DuBois Bowman
- Statistics and Climate, Peter Guttorp
- Climate Simulators and Climate Projections, Jonathan Rougier, Michael Goldstein
- Probabilistic Forecasting, Tilmann Gneiting, Matthias Katzfuss
- Bayesian Computational Tools, Christian P. Robert
- Bayesian Computation Via Markov Chain Monte Carlo, Radu V. Craiu, Jeffrey S. Rosenthal
- Build, Compute, Critique, Repeat: Data Analysis with Latent Variable Models, David M. Blei
- Structured Regularizers for High-Dimensional Problems: Statistical and Computational Issues, Martin J. Wainwright
- High-Dimensional Statistics with a View Toward Applications in Biology, Peter Bühlmann, Markus Kalisch, Lukas Meier
- Next-Generation Statistical Genetics: Modeling, Penalization, and Optimization in High-Dimensional Data, Kenneth Lange, Jeanette C. Papp, Janet S. Sinsheimer, Eric M. Sobel
- Breaking Bad: Two Decades of Life-Course Data Analysis in Criminology, Developmental Psychology, and Beyond, Elena A. Erosheva, Ross L. Matsueda, Donatello Telesca
- Event History Analysis, Niels Keiding
- Statistical Evaluation of Forensic DNA Profile Evidence, Christopher D. Steele, David J. Balding
- Using League Table Rankings in Public Policy Formation: Statistical Issues, Harvey Goldstein
- Statistical Ecology, Ruth King
- Estimating the Number of Species in Microbial Diversity Studies, John Bunge, Amy Willis, Fiona Walsh
- Dynamic Treatment Regimes, Bibhas Chakraborty, Susan A. Murphy
- Statistics and Related Topics in Single-Molecule Biophysics, Hong Qian, S.C. Kou
- Statistics and Quantitative Risk Management for Banking and Insurance, Paul Embrechts, Marius Hofert

Access this and all other Annual Reviews journals via your institution at [www.annualreviews.org](http://www.annualreviews.org).