1

Metal Ion Binding to RNA

*Pascal Auffinger,\textsuperscript{1,*} Neena Grover\textsuperscript{2} and Eric Westhof\textsuperscript{1}*

\textsuperscript{1}Architecture et Reactivité de l’ARN, Université de Strasbourg, IBMC, CNRS, 15 rue René Descartes, F-67084, Strasbourg, France
\textless p.auffinger@ibmc-cnrs.unistra.fr\textgreater
\textless e.westhof@ibmc-cnrs.unistra.fr\textgreater

\textsuperscript{2}Department of Chemistry and Biochemistry, The Colorado College, Colorado Springs, CO 80903, USA
\textless ngrover@coloradocollege.edu\textgreater

ABSTRACT

1. INTRODUCTION

2. DETAILS OF ION COORDINATION

3. PHYSIOLOGICAL RELEVANCE OF METAL IONS

4. MONOVALENT CATIONS

4.1. The Sodium and Potassium Cations

4.2. Other Monovalent Cations

4.2.1. Lithium Cations

4.2.2. Rubidium Cations

4.2.3. Cesium Cations

4.2.4. Thallium Cations

4.2.5. Ammonium Cations

5. DIVALENT CATIONS

5.1. The Magnesium Cation

5.1.1. General Properties

5.1.2. Outersphere Binding

5.1.3. Innersphere Binding

5.2. Manganese(II) as a Magnesium Substitute
ABSTRACT: RNA crystal structures have provided a wealth of information on localized metal ions that are bound to specific sites, such as the RNA deep groove, the Hoogsteen face of guanine nucleotides and anionic phosphate oxygens. With a number of crystal structures being solved with heavy metal derivatives and other “reporter” ions, sufficient information is available to estimate global similarities and differences in ion binding properties and to begin determining the influence of RNA and ions on each other. Here we will discuss the ions that are observed bound to RNA, their coordination properties, and the roles they play in RNA structural studies. Analysis of the crystallographic data reinforces the fact that ion interactions with nucleic acids are not easily interchanged between similarly charged ions. The physiological relevance of RNA-ion interactions, mainly involving K⁺ and Mg²⁺ cations, needs to be analyzed with care as different structures are solved under very diverse ionic conditions. The analysis is complicated by the fact that the assignment is not always accurate, often done under sub-optimal conditions, which further limits the generalization about the types of interactions these ions can establish.

KEYWORDS: anions · cations · crystallography · divalent metal ions · hexamines · ions · lanthanides · magnesium · manganese · metal binding · monovalent metal ions · potassium · RNA · sodium · solvation · transition metals · trivalent metal ions
1. INTRODUCTION

The roles of metal ions in the formation of RNA structures and catalysis are inherently intertwined (for reviews, see [1–5]) and have been extensively investigated since the sixties [6]. Only a fraction of the ions that are involved in forming functional RNA molecules show up in RNA crystal structures. In addition, the X-ray structures may not necessarily correspond to active forms [7–9], as crystallography often requires non-physiological ionic conditions. Therefore, a series of biochemical and biophysical studies are needed, in conjunction with structural studies, to fully understand the role of metal ions in RNA structure formation and to clarify their potential role in RNA function [10]. To achieve this goal, knowledge of ion binding motifs is required and can only be gained by a thorough examination of existing crystal structures.

In this chapter, we are examining, primarily from a structural point of view, interactions of various monovalent, divalent, trivalent, and tetravalent ions seen in RNA crystal structures.

2. DETAILS OF ION COORDINATION

For precisely characterizing ion binding sites, it is essential to first gather structural details on ion binding modes [11]. Coordination distances and geometries are two of the criteria that are used for assigning ionic species to electron density spots [12–17]. For example, Li⁺ and Mg²⁺, display average Li⁺/Mg²⁺...Ow coordination distances below 2.2 Å. Other ions, like Rb⁺ and Cs⁺ display Rb⁺/Cs⁺...Ow coordination distances around 3.0 Å (even 3.6 Å for iodide). Statistical analysis of the structures deposited in the Cambridge Structural Database (CSD), the world repository of small-molecule crystal structures [18,19], provides estimates of coordination distances with a better accuracy than the larger and generally less precise structures deposited in the Protein Data Bank (PDB).

Several web tools are now available for analyzing metal ion distributions in the PDB and other databases. Programs, such as MeRNA, MINAS, and SwS, were designed for mapping metal ion interactions with nucleic acids. These tools are based on an exploration of sub-ensembles of nucleic acid structures with specific goals in mind.

MeRNA (http://merna.lbl.gov) is a database for metal ion binding sites in RNA (see Chapter 2 of this volume) that currently scans 389 PDB files to characterize RNA interactions with 23 different metal compounds [20].
MINAS (http://www.minas.uzh.ch) contains exact geometric information on the first and second-shell ligands of metal ions present in nucleic acid structures (see Chapter 2 of this volume).

SwS (http://www-ibmc.u-strasbg.fr/arn/sws.html) is a Solvation web Service designed to provide a statistical overview of the structure of the first solvation shell of nucleic acid base pairs that is useful for gaining a better understanding of crystal structures and for validating the results of molecular dynamics simulations [21]. Among all the features of this web service, SwS characterizes cation and anion binding sites located in the first solvation shell of selected base pairs extracted from RNA and DNA structures (Figure 1).

![Figure 1](image_url)

**Figure 1.** Na\(^+\) (top) and Mg\(^{2+}\) (bottom) ion distributions around an rG=C base pair as calculated from PDB structures by the SwS web service [21]. These views depict the ion positions around an average rG=C base pair. For Na\(^+\) and Mg\(^{2+}\), 1227 and 1164 ions are shown, respectively.

*Met. Ions Life Sci. 2011, 9, 1–35*
3. PHYSIOLOGICAL RELEVANCE OF METAL IONS

A significant number of RNA structures show metal ions interacting with nucleic acids, either directly or through their first solvation shell. Almost all possible ion combinations have been tested in crystallization assays. Many of these metal ions are not biologically relevant and some are even toxic. Others are known cofactors for proteins with specific enzymatic functions, are found only in trace concentrations in cells and are unlikely to interact with nucleic acids \textit{in vivo}. Indeed, the two cations that are the most frequently found in the vicinity of nucleic acids are the divalent Mg$^{2+}$ and monovalent K$^+$ ions. Compared to all other cations, Mg$^{2+}$ ($\approx 1$ mM) and K$^+$ ($\approx 140$ mM) display the highest concentrations of divalent and monovalent ions in intracellular fluids (cytosol) of practically all living cells and are consequently the cations with the greatest likelihood to interact with nucleic acid components [22–25]. The most prevalent monovalent cation in extracellular fluids is Na$^+$ ($\approx 12$ and $150$ mM in cytosol and extracellular fluids, respectively). Therefore, it is assumed that Na$^+$ is rarely in contact with nucleic acid elements. Indeed, besides Mg$^{2+}$ cations, numerous experiments point to the necessity of having K$^+$ cations in buffer solutions for activating RNA systems. For instance, the importance of K$^+$ ions for ribosome structure and function became evident when significant unfolding of mammalian ribosomes was observed in their absence [26–28].

Other studies report that ribosomal activity is dependent on the presence of K$^+$ or NH$_4^+$ cations but is inhibited by Na$^+$ [29–31]. A recent study suggests that Na$^+$ and K$^+$ cations allosterically regulate cooperative binding of human progesterone receptor to DNA [32]. The cooperative binding of the receptor to DNA is activated or repressed by Na$^+$ and K$^+$, respectively. Interestingly, the apparent binding affinities of Na$^+$ and K$^+$ are comparable to their intracellular concentrations. Furthermore, progesterone receptors directly regulate the genes of a number of ion pumps and channels. This type of regulation seems analogous to that of Mg$^{2+}$ riboswitches which are Mg$^{2+}$-sensing RNAs regulating Mg$^{2+}$ concentration in cytosols [33–37]. A majority of Mg$^{2+}$ in the cell is bound to cellular components and a significant portion of Mg$^{2+}$ is bound to phosphometabolites, such as ATP [38], where its role is to neutralize the charge density of phosphates and to assist in the enzymatic hydrolysis reactions. Magnesium also stabilizes ribosome structures [28,39,40]. The growth of \textit{Escherichia coli} cells under conditions of Mg$^{2+}$ starvation results in ribosome depletion [41] and the \textit{in vitro} association of the small and large ribosomal subunits to form intact ribosomes depend strongly on Mg$^{2+}$ concentration [42–44].
4. **MONOVALENT CATIONS**

4.1. The Sodium and Potassium Cations

K\(^+\) cations prevail over Na\(^+\) cations in intracellular fluids and, consequently, are more likely to interact with nucleic acids. Despite this knowledge, Na\(^+\) cations were (and still are) used in NMR and crystallographic buffers [40]. An analysis of ion distributions, based on crystallographic structures, around an rG=C pair using SwS web service [21] reveals that 2218 Na\(^+\) and only 88 K\(^+\) cations are detected in the solvation shell of this pair.

The first monovalent cations to be tentatively detected in nucleic acid structures were Na\(^+\) cations in the vicinity of a DNA/drug complex [45]. Analysis of coordination distances indicates that assignment of K\(^+\) cations would also have been appropriate. When a hybrid solvent model was applied to a DNA dodecamer, the solvent sites that were previously associated with water, were defined as hybrid sites that were alternatively occupied by water molecules and cations [46–50]. These findings were rapidly followed by the detection of bound cations through the use of crystallographic difference maps [51] and confirmed by Tl\(^+\) substitution experiments [48].

Use of Na\(^+\) cations in crystallographic experiments might be associated with improved crystallization conditions although this has not been documented. It was noted that substituting K\(^+\) ions for Na\(^+\) slightly decreased RNA hairpin stability [52] as well as the stability of some ribozymes [53]. However, using crystallization mother liquors containing predominantly Na\(^+\) ions seems odd since K\(^+\) cations are much easier to distinguish from water molecules for several reasons: first, because a K\(^+\) cation displays an excess of 8 electrons over Na\(^+\) cations; second, Na\(^+\) ions are almost iso-electronic to water molecules; third, K\(^+\) cations provide anomalous signals whereas Na\(^+\) cations do not. K\(^+\) specific anomalous signals have been successfully used in several studies [48,54]. For structural assignments one can solely rely on the specific Na\(^+\)...Ow coordination distance of 2.4 Å to detect sodium cations [46]. For potassium cations, similarity between the average K\(^+\)...Ow and O/N...Ow coordination distances (2.8 Å) may, however, slightly complicate the detection of K\(^+\) cations.

The first specific monovalent binding site in RNA was identified crystallographically in the vicinity of an AA platform of a tetraloop receptor [55] (Figure 2). The presence of K\(^+\) cations was inferred from Tl\(^+\) and Cs\(^+\) soaking experiments. The specificity of the AA platform for K\(^+\) was confirmed by the fact that the related *Azoarcus* intron is six times more active in K\(^+\) instead of Na\(^+\) buffer. Another strong and specific K\(^+\) binding site is buried deep inside the structure of a 58-nucleotide ribosomal RNA
In this structure, six oxygen atoms from buried phosphate groups contact a K$^+$ cation [56] (Figure 2). A large number of monovalent ion binding sites (86 Na$^+$ and 2 K$^+$) contacting primarily the rRNA have been identified in the crystal structure of a large ribosomal subunit [40]. Although some interactions with anionic phosphate oxygens were identified, like the AA platform motif binding a Na$^+$ cation (Figure 2), the most common monovalent cation-binding sites appears to be located in the major groove of G·U wobble pairs. Binding of a Na$^+$ cation to an AA platform is similar to that observed for a tetraloop receptor [55].

Interestingly, a K$^+$ binding site has also been identified in close vicinity to the ribosomal peptidyl transferase active site [57] (Figure 2). In the ribosome,
it is very likely that under physiological conditions most, if not all, of the monovalent binding sites are occupied by K\(^+\) rather than Na\(^+\) [40]. Indeed, functional specificity for K\(^+\) has been observed in several natural RNAs [55]. For binding of K\(^+\) and Na\(^+\) to RNA see also [34,54,58–60].

Molar concentrations of monovalent cations often stabilize RNA structures [53,61,62]. Although they allow charge neutralization that collapses the polynucleotidic chains into a more compact structure, monovalent ions do not necessarily lead to the active fold of large RNAs in the absence of divalent cations [63]. For ribozymes Na\(^+\) and K\(^+\) cations do not seem to be directly involved in catalytic mechanisms but are necessary for the activity of all ribozymes [64].

4.2. Other Monovalent Cations

4.2.1. Lithium Cations

Lithium ions are often used in molar concentrations in crystallization solutions. Yet their low numbers of electrons (3 e\(^-\)) makes detection of Li\(^+\) ions very difficult. There are 35 structures in the PDB database in which Li\(^+\) was assigned to an electron density spot and only one of these is a nucleic acid structure, a DNA quadruplex that contains a Li\(^+\) cation coordinated to an O2\(^-\) atom and three water molecules. A tetrahedral arrangement with coordination distances under 2.0 Å is likely to be the most precise signature for these cations [15,17,65]. Regular tetrahedral arrangements should be systematically sought to identify Li\(^+\) when this ion is present in the crystallization solution.

Interestingly, Li\(^+\) cations were shown to promote catalytic activity of various ribozymes such as the hammerhead, the hairpin and the VS ribozymes at molar concentrations [62,66]. Li\(^+\) cations are also considered being good at RNA folding [67]. An in vitro vitamin B\(_{12}\) RNA aptamer selected in 1 M LiCl buffer needed Li\(^+\) cations; these ions could not be substituted by Na\(^+\) or K\(^+\) indicating that this aptamer is adapted to the conditions in which it was selected [53].

Li\(^+\) cations are used to control bipolar disorders. It has been suggested that competition between Li\(^+\) and Mg\(^{2+}\) ions for divalent ion binding sites in cellular components is the underlying theme in putative mechanisms of Li\(^+\) action [68].

4.2.2. Rubidium Cations

Rubidium cations could be identified in four different DNA structures. Rubidium is not normally found in living organisms. Due to its similarity
with potassium cations, Rb$^+$ concentrates into intracellular fluids and short-lived radioisotopes can be used as biomarkers. Crystallographers have used rubidium as a probe for detecting, otherwise difficult to characterize, Na$^+$ cation binding sites with the assumption that both cations would bind at the same locations [40,69,70]. Water coordination distances for Rb$^+$ are not well defined but are expected to be around 3.0 Å with fluctuating coordination numbers between 6 and 8 [17]. With the exception of Na$^+$ and Li$^+$ cations, K$^+$, Rb$^+$, Cs$^+$ and Tl$^+$ cations exhibit significant anomalous signals, especially the latter three cations, providing an opportunity to use single wavelength anomalous diffraction (SAD) technique to solve nucleic acid structures. The combination of high resolution and anomalous diffraction data can also pinpoint partially occupied binding sites [69]. Yet, Rb$^+$ and Cs$^+$ substitutions are, at present, seldomly used for detecting monovalent cation binding sites in biomolecular systems.

4.2.3. Cesium Cations

Cesium cations have a slightly larger ionic radius than rubidium cations ($\approx 3.1$ Å [17]). These have been characterized in one DNA and three RNA structures deposited in the PDB. Use of Cs$^+$ in crystallography is similar to that of Rb$^+$ [55]. Cs$^+$ displays larger anomalous signals but its coordination properties differ more from those of K$^+$ than Rb$^+$ thus limiting its use. Cs$^+$ cations, like [Co(NH$_3$)$_6$]$^{3+}$ and unlike Mg$^{2+}$ cations, bind to the major groove of single and tandem G·U. This specificity led to the development of a strategy to solve the phase problem using the insertion of appropriate G·U motifs [71].

4.2.4. Thallium Cations

Tl$^+$ cations represent a more appropriate K$^+$ mimic than Rb$^+$ and Cs$^+$ cations, given very similar coordination distances ($\approx 2.8$ Å) and enthalpies of hydration ($\approx -78$ kcal/mol) [48]. Tl$^+$ cations are nearly interchangeable with K$^+$ in Na$^+$/K$^+$ pumps [72] as well as in several other biochemical systems [73,74]. However, one must be careful in interpreting Tl$^+$ containing structures, as Tl$^+$ is not a biologically relevant cation. It has $d$ electrons that favor covalent bonding. Hence, soft atoms such as sulfur might interact differently with Tl$^+$ than with K$^+$ [48].

Despite their high toxicity, Tl$^+$ cations are used in crystallography since they display significant anomalous scattering of X-rays. Moreover, Tl$^+$ is a strong scatterer and can be identified simply by its peak height in standard electron density maps for fully occupied sites [59]. Hence, the inclusion of Tl$^+$ cations in crystallization liquors provides a sensitive monovalent cation
detection system. \( K^+/\text{Tl}^+ \) substitution was used in the interpretation of several nucleic acid structures such as RNA and DNA duplexes [48,69,70,75–77], the *Tetrahymena* ribozyme P4-P6 domain [55], a 58 nucleotide ribosomal fragment [56], the signal recognition particle (SRP) [78], a viral RNA pseudoknot [59] and the HDV ribozyme [64]. With strong scattering and strong anomalous signal, it is not surprising that more monovalent binding sites are observed with \( \text{Tl}^+ \) than with any other monovalent ion [48]. By using anomalous diffraction techniques, partial occupancies can be easily attributed to \( \text{Tl}^+ \) [48,55]. In a crystal structure of a DNA duplex, 13 partial \( \text{Tl}^+ \) binding sites could be identified. It was inferred that the number of monovalent cation binding sites substantially exceeds previous observations using less strongly scattering Rb\(^+\) (one 50\% occupied position) or Cs\(^+\) ions (four 20\% occupied positions). Indeed, the cumulative occupancy for the 13 sites is close to 2.26 [48].\(^{205}\)Tl can be detected by NMR techniques and is useful in characterizing bound monovalent cations that exchange slowly [79].

4.2.5. Ammonium Cations

Ammonium cations are non-metal ions that are very difficult to observe by crystallography and are certainly not physiologically relevant given their cellular toxicity. It has been reported that the hammerhead, hairpin and VS ribozymes are all catalytically active in solutions containing ammonium ions and no divalent cations [62]. The hairpin ribozyme, in particular, shows a fourfold increase in cleavage rate when magnesium ions are replaced by ammonium ions at a 1.0 M minimal concentration pointing to the fact that NH\(_4^+\) cations have a great structuring potential. NH\(_4^+\) cations mainly bind to phosphate groups and display coordination distances comparable to those of water molecules (2.8–3.0 Å).

5. DIVALENT CATIONS

5.1. The Magnesium Cation

5.1.1. General Properties

Mg\(^{2+}\) cations are physiologically the most relevant divalent ions for the formation of RNA structures and its functions. Mg\(^{2+}\) has a smaller ionic radius than Na\(^+\), K\(^+\), and Ca\(^{2+}\). It is a hard ion with only ten electrons. Mg\(^{2+}\) predominantly binds six ligands in an octahedral geometry and prefers hard ligands, such as oxygen, over nitrogen and sulfur atoms. The Mg\(^{2+}\)...Ow distance is 2.1 Å. Mg\(^{2+}\) binding to RNA is difficult to study as this cation is...
not amenable to direct spectroscopic studies, and is difficult to distinguish from Na$^+$ and H$_2$O in crystal structures (given that they have almost equal numbers of electrons). High resolution (1.5 Å or better) and high occupancy are necessary to accurately identify Mg$^{2+}$ ions in crystal structures. The metal-water distance (bond length) is what distinguishes Na$^+$ from Mg$^{2+}$ and is only interpretable at high resolution. Deprotonated water ligands further complicate unambiguous assignment of Mg$^{2+}$ ions in crystal structures.

5.1.2. Outersphere Binding

Mg$^{2+}$ interacts with RNA in both outersphere and innersphere modes. In the former Mg$^{2+}$ associates with the negatively charged phosphate backbone in a diffuse manner [80,81]. Mg$^{2+}$ is preferred for charge neutralization over the monovalent ions (K$^+$, physiological) because of its greater charge density and a lower entropic cost for localization of these ions [2,3,82]. A compactly folded RNA has a greater charge density and is likely to accumulate a greater number of Mg$^{2+}$ ions in its surroundings. The assumption is that Mg$^{2+}$ ions that are freely exchanged with bulk Mg$^{2+}$ (diffusely bound) are not likely to show up in the crystal structures due to their short occupancy times. A few Mg$^{2+}$ outersphere interactions that show up in RNA are presumed to be playing a role that may be different from the bulk ions.

Preferred ligands for outersphere interactions with Mg$^{2+}$ are anionic phosphate oxygens and electronegative atoms of bases. Outersphere interactions are primarily seen in the major groove of A-form helices, that display a pronounced negative electrostatic surface potential. Hydrated Mg$^{2+}$ ions bind to the floor of the deep groove of A-form helices by forming hydrogen bonds to acceptors atoms (mostly G(N7) and G(O6)) especially in GpG and GpU steps. The residence time for these ions is long enough to differentiate them from bulk ions [83,84]. The hexahydrated Mg$^{2+}$ prefers to bind to the Hoogsteen face of guanines, especially when the guanine is part of a non-canonical base pair. This is seen in the HIV-1 dimer initiation site [54], in the NMR structure of a small construct containing 5’ GU$^3$/3’ UG$^5$ base pairs [85], and in the P4-P6 domain of Tetrahymena (at a tandem 5’ GG$^3$/3’ UU$^5$) [86]. A water-mediated charge transfer from guanine to magnesium is also thought to play a role in stabilization of interactions with guanine and is thought to occur through a cooperative mechanism [87].

Although, magnesium outersphere interactions with guanine, especially at the non-canonical G · U base pairs are often cited, it should be noted that no magnesium binding is seen in 5’ GU$^3$/3’ UG$^5$ in the crystal structure of the recognition domain [78] or in the NMR structure of the tetraloop receptor complex [88]. Expectation of outerspace ion interactions with the G · U base
pairs is due to the presence of G(N7), G(O6), and U(O2) atoms in the major groove, creating an unusual hydrogen bond acceptor environment. A variety of ions interact with this site while magnesium hexahydrates are not systematically located there [58,89]. Several different explanations are possible for this observation, including low-resolution of the structures or use of magnesium analogs, making the assignment of magnesium binding to a particular site problematic. Hence, caution is advised before overanalyzing specific ion binding properties.

In a 2.4 Å resolution structure of the 50 S subunit of *Haloarcula marismortui*, nine out of 116 Mg$^{2+}$ ions are bound via outersphere interactions to the 3045 nucleotides of the RNA (Figure 3). These Mg$^{2+}$ ions were characterized by visually identifying the octahedral arrangement of water molecules located in the major groove of A-form helices [40]. As expected, hexahydrated Mg$^{2+}$ cations are much more mobile than partially dehydrated cations that form direct bonds with RNA atoms [90,91].

NMR and crystallographic data often yield slight differences in metal binding sites [58,85,88]. The data for magnesium binding are reliable when magnesium ions are being used, instead of substitutes, as demonstrated throughout this chapter.

5.1.3. **Innersphere Binding**

For an RNA fragment to form direct interactions with metal ions, the local electrostatic environment and the energetics of metal dehydration are among the many factors that play a role. Partial dehydration of Mg$^{2+}$ is more likely in the non-helical region where neighboring phosphates are in close proximity [92]. As seen in the P4-P6 group I intron and 58-nucleotide rRNA crystal structure [86,92,93], chelated metal ions would be expected to be buried and relatively inaccessible to solvent due to phosphate charge neutralization. Even though dehydration of hexahydrated magnesium is energetically expensive, crystal structures of RNA show that specific environments allow for the formation of multiple direct bonds between magnesium and RNA. The properties of metal ions in the presence of RNA are likely to be different from the free metal ion; hence, estimates of the energetic requirements of forming innersphere coordination are rather imprecise. Studies on model systems such as nucleotides and small RNAs will lead to better understanding of direct interactions between metal ions and RNA [94,95]. Advances in theoretical and computational modeling, such as non-linear Poisson-Boltzmann calculations, microenvironment analysis, and molecular dynamic simulations (in concert with improved NMR approaches) will play a key role in understanding the dynamic behavior of magnesium-RNA interactions [96,90–92].
Figure 3. An example of each of the six binding modes of interaction between Mg$^{2+}$ and RNA derived from a crystal structure of the large ribosomal subunit of *Haloarcula marismortui* crystallized at 2.4 Å resolution involving six to three water molecules and zero to three direct contacts with RNA hydrophilic atoms. Reproduced by permission from [174]; copyright (2001).
Several high-resolution small RNA structures show potential magnesium binding sites [54,59,97]. A few comparative analysis of ion binding sites have also been performed [54,88]. Differences in crystallization conditions along with the use of magnesium analogs makes it challenging to generalize magnesium innersphere interactions. Some examples of magnesium binding sites and types of interactions are discussed here to provide insight into the role of innersphere interactions in RNA structures.

The loop E motif of the ribosomal RNA is a classical example of inner-sphere coordination of Mg\(^{2+}\) ions with RNA. This structure has been solved at 1.5 Å and shows Mg\(^{2+}\) bound in its octahedral and water in its tetrahedral geometry. Due to the high resolution, Mg\(^{2+}\) and water are distinguishable by their bond lengths of 2.1 Å versus 2.7 Å, respectively [97]. Two of the bound Mg\(^{2+}\) are hexahydrated and interact with the backbones of both strands. Three of the bound Mg\(^{2+}\) show one direct interaction each with RNA. One Mg\(^{2+}\) ion is primarily coordinated to G(N7), G(O6), and anionic phosphate oxygens. Three additional waters are proposed to bridge interactions between two Mg\(^{2+}\) ions that are seen bound to adjacent phosphate groups [97].

A classification of Mg\(^{2+}\) binding sites was derived from the observation of 116 Mg\(^{2+}\) binding events to the large ribosomal subunit of Haloarcula marismortui [40]. Ten different binding modes were observed: those with zero (type 0), one (type I), two (type II), three (type III), four (type IV), or five (type V), bonds between the Mg\(^{2+}\) ion and RNA. Each of these were found 9, 37, 45, 18, 1, and 1 times, respectively (Figure 3). The Type II and III interactions were further subdivided into two groups based on the three-dimensional arrangement of their RNA, protein, and/or water ligands. Complete dehydration of Mg\(^{2+}\) cations has not been reported yet. No octahedral coordination geometry was observed in five cases leading to description of a type X. Most innersphere interactions involve phosphate moieties that are in close proximity and require charge neutralization to form a particular structure. Anionic phosphate oxygens serve as ligand for Mg\(^{2+}\) ions in a majority of innersphere interactions. The N7 of purines or O6 of guanine are good outersphere ligands for Mg\(^{2+}\) and form innersphere interactions with Mg\(^{2+}\) in approximately 20% of the cases.

Recently, two different magnesium sensing riboswitches have been reported [33,34,37]. In the M-box RNA, six Mg\(^{2+}\) binding sites are seen, three of which are considered crucial for forming the compact structure [37]. One Mg\(^{2+}\) is bound via four innersphere interactions in the loop 5 region and is considered to be the most significant. It is likely that additional magnesium binding motifs exist in RNA that have not yet been identified.

Experimental approaches that allow studying RNA under physiologically relevant conditions, such as fluorescence resonance energy transfer (FRET),
transient electric birefringence (TEB) or gel shift assays provide crucial insights into changes in RNA structures under varying concentrations of magnesium ions and built a link between RNA structure and function [95,98]. Examining site-specific binding of magnesium is remains a challenging problem as Mg\(^{2+}\) cannot be perfectly substituted by any other ions.

5.2. Manganese(II) as a Magnesium Substitute

Divalent manganese (Mn\(^{2+}\)) is similar to Mg\(^{2+}\) cations in some of its chemical properties. Yet, these cations have different biological roles. An early survey of crystal structures from the CSD and PDB indicated that Mg\(^{2+}\) cations bind preferentially to oxygen atoms while Mn\(^{2+}\) prefers nitrogen [99]. It has also been reported that the affinity of Mn\(^{2+}\) and Mg\(^{2+}\) cations for A(N7) is similar, while that of Mn\(^{2+}\) for G(N7) is larger than that of Mg\(^{2+}\) [94,100]. The fact that Mn\(^{2+}\), as a softer metal, is better able to coordinate to the softer sulfur than Mg\(^{2+}\) has been quite successfully used for rescue experiments [10].

Manganese has two major functions in enzymes: (1) as a Lewis acid, for which its properties can be compared with those of magnesium, zinc, and calcium, and (2) as an oxidation catalyst, for which it can be compared with iron and copper. Divalent manganese has a radius of approximately 0.75 Å, somewhat larger than that of magnesium (0.65 Å) leading to Mn\(^{2+}\)...O coordination distances larger by 0.1 Å than the Mg\(^{2+}\)...Ow coordination distance (∼2.2 and 2.1 Å, respectively [14]). When divalent manganese replaces magnesium in the active site of a magnesium-utilizing enzyme, the catalytic activity of the enzyme is often maintained. This observation and the known paramagnetic properties of manganese have led to the widespread use of Mn\(^{2+}\) in NMR [88,101] and EPR studies [10,102–104]. Magnesium, however, is rarely a competent replacement for divalent manganese in enzymes. Given this higher affinity for nitrogens, Mn\(^{2+}\) prefers, as a transition metal with \(d\) electrons, to bind to lysine and histidines [99] and, therefore, probably discriminates between proteins and RNA on this basis.

The effects of the substitution of Mg\(^{2+}\) by Mn\(^{2+}\) cations or of Mn\(^{2+}\) soaking experiments on RNA structure and catalytic activity are diverse. Sometimes, Mn\(^{2+}\) ions appear as effective as Mg\(^{2+}\) ions for folding and catalytic activity. For the RNase P [105] and the HDV ribozyme [106], Mn\(^{2+}\) can substitute for Mg\(^{2+}\) with only a slight change in activity. In other instances, the reaction rate can be accelerated by such substitutions as seen for the minimal hammerhead ribozyme [107]. Mn\(^{2+}\) cations are also able to inhibit the catalytic reaction of ribozymes such as the hairpin ribozyme. In this case, Mn\(^{2+}\) cations are not able to correctly fold this RNA, whereas Sr\(^{2+}\) and Ca\(^{2+}\) can [108]. Additional non-specific cleavage with Mn\(^{2+}\) was also observed with group II introns [109].
Mn$^{2+}$ cations can induce localized conformational changes. For the SRP system [78] and a DIS sub-type [54], localized conformational changes were reported when Mg$^{2+}$ is replaced by Mn$^{2+}$. For the glycine riboswitch, Mg$^{2+}$, Mn$^{2+}$, and Ca$^{2+}$ facilitate glycine binding while other divalent cations did not [110]. For the SRP particle, Cs$^+$ and Mn$^{2+}$ significantly increase the protein-RNA binding affinity over that observed in the presence of K$^+$ and Mg$^{2+}$ [78] and biochemical studies assessed that Mn$^{2+}$ cations are more efficient than Mg$^{2+}$ for folding an mRNA pseudoknot [67]. Mn$^{2+}$ soak of the lysine riboswitch did not replace Mg$^{2+}$ [60]. Replacement of Mg$^{2+}$ by Mn$^{2+}$ cations cause ribosomal crystals to become twinned [40].

Overall, the data with these substituent metal ions may lead to incorrect assumptions about the role of magnesium ions in RNA structure and functions, since Mn$^{2+}$ and Mg$^{2+}$ do not always bind to the same sites as a result of different affinities for hard and soft atoms.

### 5.3. Other Alkaline Earth Metal Cations

Calcium is by far the most abundant divalent ion in the human body. Yet, its concentration in the cytosol is below the micromolar range while that of Mg$^{2+}$ is in the millimolar range. Hence, Ca$^{2+}$ cations are not likely to interact with nucleic acid components. The Ca$^{2+}$...Ow coordination distance (2.4 Å) is close to that of Na$^+$. The intermediate Sr$^{2+}$ cations have no particular biological functions. Non-radioactive strontium is not toxic in low doses. In the human body, Sr$^{2+}$ can substitute for Ca$^{2+}$ cations in bones. The Sr$^{2+}$...Ow coordination distance is $\approx$ 2.6 Å. Ba$^{2+}$ is the largest cation in the alkaline earth series with Ba$^{2+}$...Ow coordination distances around 2.8 Å (close to that of K$^+$) and is poisonous.

Ca$^{2+}$, Sr$^{2+}$, and Ba$^{2+}$ display anomalous signals [69] and were reported to behave similarly. DIS crystal structures indicated that Ca$^{2+}$, Sr$^{2+}$, and Ba$^{2+}$ cations display a clear preference for outersphere coordination to deep groove Hoogsteen sites of guanines leading to much less localized binding, although, Ba$^{2+}$ establishes in one occurrence innersphere coordination to a G(N7) atom [111]. These group II cations induce a distortion of the DIS helix leading to a lack of polymorphism with Mg$^{2+}$ structures [54]. On the other hand, in a high-resolution leadzyme structure crystallized in the presence of a mix of Mg$^{2+}$ and Sr$^{2+}$, strontium cations did not displace or perturb bound Mg$^{2+}$ cations seen in the absence of Sr$^{2+}$ [7]. A subdomain of the hepatitis C virus displays similar conformations when crystallized in the presence of Sr$^{2+}$ or Mg$^{2+}$ cations [112] (Figure 4).

Like other divalent cations, Ca$^{2+}$, Sr$^{2+}$, and Ba$^{2+}$ do not bind to G·U motifs [71]. Ca$^{2+}$ binds directly to phosphate groups [113,114] (Figure 4).
the TAR RNA structure (1.3 Å), Ca$^{2+}$ binding hinders the formation of an active structure [113]. The crystallization was performed in the presence of Mg$^{2+}$, indicating that the structure formed in the presence of Ca$^{2+}$ is preferred over the presumably active structure in the presence of Mg$^{2+}$. Thermodynamic analysis of TAR RNA constructs shows differences in stability in the two divalent ions [95]. In the crystal structure of the malachite green aptamer Sr$^{2+}$ is required for crystallization (Ca$^{2+}$ and Ba$^{2+}$ did not substitute for it). Indeed, a Sr$^{2+}$ cation bridges two symmetry-related molecules in the crystal lattice through inner sphere contacts with an U(O2) and a phosphate group oxygen atom [115]. In another structure, a Sr$^{2+}$ cation interacts with the O3' and O2' atoms of two symmetry-related terminal nucleotides and completes its coordination sphere with five additional water molecules [116] (Figure 4).

Sr$^{2+}$ and Ba$^{2+}$ maintain the catalytic activity of the HDV ribozyme [106]. Yet the affinity of Ba$^{2+}$ and Sr$^{2+}$ for the P4-P6 domain of the Tetrahymena group I intron ribozyme was two orders of magnitude smaller than that of other cations [117]. Similarly, these cations are less effective than Mg$^{2+}$ in folding a mRNA pseudoknot, with Ca$^{2+}$ displaying the weakest activity [67].

**Figure 4.** Examples of binding of Ca$^{2+}$ and Sr$^{2+}$ to RNA. (A) Superposition of two structures (one with Mg$^{2+}$, the other with Sr$^{2+}$) of the subdomain IIa of the hepatitis C virus. Reproduced by permission from [112]; copyright (2008). (B) A Ca$^{2+}$ cation stabilizing a sharp turn by interacting with three anionic oxygen atoms. Reproduced by permission from Ref. [113]; copyright (1998). (C) A Sr$^{2+}$ ion involved in crystal packing interactions. Reproduced by permission from Ref. [116]; copyright (1999).
5.4. Divalent Transition Metal Cations and Lead(II)

5.4.1. Cobalt(II) Cations

Co$^{2+}$ is a key component of cobalamin (vitamin B$_{12}$) and is only found in trace amounts in most organisms. Hence, it is not likely to interact with nucleic acid components. It is used in crystallography and biochemical assays. DIS crystal structures show Co$^{2+}$ binding to the major groove N7/O6 atoms of guanine Hoogsteen sites [54]. However, the site occupancies are weak (between 0.16 and 0.82). Crystal structures of the hammerhead ribozyme display Co$^{2+}$ in both inner- and outersphere coordination [62,118]. Most crystal structures of DNA molecules indicate a strong direct binding to G(N7) atoms (see also an early tRNA$^{Phe}$ structure [119]). Direct contacts with hydrophilic atoms imply average coordination distances around 2.1 Å. Note that a quantum mechanical study emphasizes the preference of transition metals (Co$^{2+}$ and Mn$^{2+}$) over Mg$^{2+}$ for binding to N7 atoms of guanines [100].

5.4.2. Nickel(II) Cations

Nickel plays important roles in the biology of microorganisms and plants. As found in six DNA structures, Ni$^{2+}$ coordinates exclusively to G(N7) atoms (see for example [120]) (Figure 5). In protein/nucleic acid structures, nickel atoms prefer to interact with amino acid residues [121]. Average Ni$^{2+}$...OW coordination distances are around 2.1 Å.

![Figure 5](image)

**Figure 5.** Binding of Cd$^{2+}$ and Ni$^{2+}$ ions to guanine residues. (A) Binding of Ni$^{2+}$ and associated waters to a guanine Hoogsteen site as observed in a netropsin bound DNA duplex at 1.58 Å resolution. Reproduced by permission from [120]; copyright (1999). (B) A Cd$^{2+}$ cation that stabilizes sharp turns in a four-way junction as observed in a ribosomal RNA/protein structure. Reproduced by permission from [130]; copyright (1999).
5.4.3. Copper(II) Cations

Copper is the third most abundant transition metal in the body and in the brain. The major oxidation states for copper ions in biological systems are cuprous Cu\(^{+}\) and cupric Cu\(^{2+}\); intriguingly, the former is more common in the reducing intracellular environment and the latter in the oxidizing extracellular environment [122].

No RNA structure associated with Cu\(^{2+}\) has been deposited into the PDB. It has been stated that Cu\(^{2+}\) displays a higher affinity for G(N7) than Zn\(^{2+}\), Cd\(^{2+}\), and Mg\(^{2+}\) [123]. In this respect, one Z-DNA structure shows clear interactions of Cu\(^{2+}\) with G(N7) atoms [124]. Average coordination distances are close to 2.0 Å and Cu\(^{2+}\) binds, like Hg\(^{2+}\) and Au\(^{3+}\), to the N3 atom of a modified purine [125].

5.4.4. Zinc(II) Cations

Next to iron, zinc is the second most abundant transition metal in the human body. Its concentration is the highest in the brain where it matches that of magnesium [122]. Zinc cations have a different binding profile than Mg\(^{2+}\) or Mn\(^{2+}\) cations. The coordination number of Zn\(^{2+}\) varies between four and six with Zn\(^{2+}\)...Ow distances around 2.1 Å. These cations display a strong preference for interacting with nitrogen and sulfur atoms [14] and, in nucleic acids, with G(N7) atoms [123].

Zn\(^{2+}\), like all non-Mg\(^{2+}\) divalent ions, is not likely to interact with RNA molecules in vivo. Their preference for binding to proteins arises from their increased affinity for nitrogen and sulfur atoms. In DIS structures, Zn\(^{2+}\), like most transition metals, interacts with nucleobases, especially with G(N7) atoms [54]. It has been reported that the minimal hammerhead is catalytically active with Zn\(^{2+}\) cations in the presence of spermine [126]. Zn\(^{2+}\) binding is similar to Mg\(^{2+}\) at tandem G*A steps [127].

5.4.5. Cadmium(II) Cations

Cadmium has only been found in marine diatoms living in zinc-depleted environments. In all other organisms cadmium is considered to be toxic. The Cd\(^{2+}\) cation shows a clear preference for binding to G(N7) atoms when compared to Mg\(^{2+}\) cations [123]. Cadmium (\(^{113}\)Cd) is used in NMR spectroscopy [79] and in EPR silencing experiments [128]. Given their thiophilic character, Cd\(^{2+}\) cations are used, similarly to Mn\(^{2+}\), in phosphorothioate rescuing experiments [129]. However, it is important to note that addition of a thiophilic metal ion rescues some reaction steps but has deleterious effects on others.

Three Cd\(^{2+}\) cations are found in a ribosomal fragment bound to protein L11. One ion is thought to stabilize the association of the two complexes in

*Met. Ions Life Sci. 2011, 9, 1–35*
the asymmetric unit, but the other two are integral to the RNA structure [130] (Figure 5). Cd\(^{2+}\) was shown to bind like Mg\(^{2+}\) to tandem G \cdot A steps [127]. Its average coordination distance is close to 2.1 Å.

5.4.6. Mercury(II) Cations

Although considered to be toxic, mercury, is an important component of Chinese, Tibetan, and Ayurvedic medicine. Mercury has no known physiological role as a trace element. One weak Hg\(^{2+}\) ion binding site is observed in a crystal structure of a 5S rRNA domain [97]. Hg\(^{2+}\) intercalates also between a T-T noncanonical base pair [131] that resembles the binding of an Au\(^{3+}\) between a guanine and a cytosine of a GxC base pair (see Figure 8 in Section 7.1). In such complexes, mercury induces deprotonation of the N3 atom of each thymine, which yields a highly stable complex. In a ribosomal fragment bound to protein L11, a mercury cation binds to the N3 atom of a U with a coordination distance of 2.4 Å [130]. Interestingly, Hg\(^{2+}\) is bound to nucleobases and not to amino acids in this structure. Hg\(^{2+}\) was also observed in an early structure of tRNA\(^{Phe}\) [132].

5.4.7. Lead(II) Cations

Lead is a poisonous metal with no known medicinal applications [133]. It was recognized very early that Pb\(^{2+}\) is involved in RNA-specific position cleavage [134,135]. Hence, it is commonly used as a chemical probe for determining RNA three-dimensional structures.

A low-resolution crystallographic structure reveals three Pb\(^{2+}\) tRNA binding sites [136]. In another low-resolution structure of the specificity domain of the RNase P, 23 Pb\(^{2+}\) atoms were identified [137]. Finally, in a DIS structure, three partially occupied binding sites were described [54]. Unfortunately, given the cleaving activity of Pb\(^{2+}\) on the \textit{in vitro} selected leadzymes, no crystal structures with lead were obtained [7]. Hence, no clear consensus on Pb\(^{2+}\) binding sites can be proposed. Pb\(^{2+}\) has been used to obtain anomalous diffracting derivatives of an RNase P fragment [138].

6. TRIVALENT CATIONS

6.1. Hexammine Cations ([Co/Ru/Rh/Os/Ir(NH\(_3\))\(_6\)]\(^{3+}\))

Hexammine cations have no biological functions but are sometimes used to aid crystallization [139] and as heavy atom derivatives that can provide phase information through multiple or single wavelength anomalous
diffraction (MAD or SAD) methods [140,141]. They are often observed to
bind to nucleic acids through direct contacts involving the ammine groups of
the cations. Hexammine cations bind essentially to major groove atoms, to
oxygen atoms of phosphate groups and less frequently at other sites.
Interestingly, they are primarily used for solving the structure of nucleic acid
and nucleic acid analogs (129 structures) and are much less useful in protein
crystallography (12 structures). The use of \([\text{Co(NH}_3\text{)}_6\text{]}^{3+}\) cations prevails
(109 structures) over the use of the heavier ruthenium, rhodium, osmium,
and iridium hexammines (20 structures in total).

For DNA systems, hexammine cations are known inducers of the left-
handed Z-form and are claimed to stabilize this form much more effectively
than \(\text{Mg}^{2+}\) cations [142,143]. This justifies their use in the crystallization of a
large number of Z-DNA double helices. Yet, for RNA systems, there is no
real evidence that these cations drastically alter conformational equilibria.
Examination of the SRP in the presence of various cations showed that this
RNA structure accommodates a variety of metal ions without structural
changes [78].

In *Neurospora* VS ribozyme, \([\text{Co(NH}_3\text{)}_6\text{]}^{3+}\) ions induce a “similar” fold to
\(\text{Mg}^{2+}\) cations at a 33-fold lower concentration but without leading to an
active structure [144]. In mixed-metal kinetic experiments of this ribozyme,
\([\text{Co(NH}_3\text{)}_6\text{]}^{3+}\) does not inhibit \(\text{Mg}^{2+}\)-induced self-cleavage. More generally
it was reported that hexammine cations are efficient at folding RNA [63].
When included in crystallization solutions, low concentrations of cobalt
hexammine chloride (0.1–1.0 mM) dramatically increased the number, size,
and growth rate of P4-P6 crystals [139]. In the hairpin ribozyme, cobalt(III)
hexammine can replace \(\text{Mg}^{2+}\) cations for all folding and catalytic functions
implying that \(\text{Mg}^{2+}\) cations are not directly involved in the catalytic reaction
[108,145,146]. On the other hand, cobalt hexammine is a potent inhibitor of
the HDV ribozyme [64].

The question of whether hexammine cations are appropriate mimics of
\([\text{Mg(H}_2\text{O})_6\text{]}^{2+}\) cations is often debated. It is generally understood that these
cations are not perfect analogs of magnesium hexahydrate. Although their
ionic radii are close (4.1 Å for \(\text{Mg}^{2+}\) hexahydrates and 3.9 to 4.5 Å for \(\text{Co}^{3+}\)
and \(\text{Ir}^{3+}\) hexammines, respectively), their charge (+3/+2) and the number
of associated hydrogen atoms (18/12) differ. In several instances,
\([\text{Co(NH}_3\text{)}_6\text{]}^{3+}\) was observed to bind at locations also occupied by
\([\text{Mg(H}_2\text{O})_6\text{]}^{2+}\) and \([\text{Mn(H}_2\text{O})_6\text{]}^{2+}\) ions [78]. In many structures, cobalt
hexammine coordinates to GpG steps in the major groove [139]. However,
the details of coordination of hexammine and hexahydrate cations are dif-
ferent [78]. Clearly, some sites can discriminate between these ions, as
observed in the P4-P6 domain, the SRP [78], and the smaller HIV-1
dimerization initiation site [54]. Tandem G·U pairs were reported to be
good hexammine cation binders [71,78] but displayed no specificity for
divalent cations [89,147]. However, an NMR investigation found that tandem G·A pairs displayed a good affinity for divalent cations and no affinity for hexammine cations [148]. In at least one instance, a cobalt hexammine cation had all six of its ammine groups engaged in hydrogen bond contacts with RNA atoms [149]. A rare innersphere coordination to the RNA by a [Co(NH₃)₄]³⁺ cation has also been reported [54]. Besides recent Raman microscopic analysis of RNA crystals supported the fact that [Co(NH₃)₄]³⁺ cations can displace innersphere bound magnesium cations [150], indicating that the properties of these coordination complexes is affected by the RNA environment.

The observation that hexammine and monovalent cations bind more efficiently to the minor groove of G·U pair motifs, such as tandem G·U pairs [78,89] led to an efficient general strategy for solving the phase problem in RNA crystallography [71]. In this method, G·U tandems are grafted into the target sequences with the expectation to create a highly specific hexammine binding site. This “rational” soaking strategy is an improvement over earlier “soak and pray” strategies and has been successfully used in the resolution of large RNA structures [151].

Along with cobalt hexammine, four other hexammine compounds, based on ruthenium, rhodium, osmium, and iridium have been used as crystallization agents. In the resolution of an hybrid RNA duplex [152], [Co(NH₃)₆]³⁺, [Rh(NH₃)₆]³⁺ and [Ir(NH₃)₆]³⁺ complexes were isomorphous, although the resolution increases from the lighter to the heavier cation (2.2 Å for cobalt hexammine; 1.8–1.6 Å for rhodium hexammine; 1.5 Å for iridium hexammine). A Z-DNA structure crystallized in the presence of ruthenium hexammine, suggested that the interaction of these cations with adenines induced a tautomeric shift from the amino to the imino form, disrupting the canonical A-T base-pairing scheme [153]. Indeed, metal ion binding to nucleobases can alter their pKₐ values [154]. Iridium hexammine is observed in six RNA and one DNA structure (see for example [155]). These heavy and bulky cations were used to obtain the heavy atom derivatives needed for solving the RNA nucleic acid structures (Figure 6). Osmium hexammine was used in several studies [56,139]. Even though osmium hexammine is almost isoelectronic to [Ir(NH₃)₆]³⁺, the observed electron density does not allow detecting with precision the ammine groups (see for example [138]). The only exception are the unpublished structures of the P4-P6 introns in which a specific site could bind a putative osmium pentammine cation [139].

6.2. Lanthanide Cations (La³⁺ to Lu³⁺)

The lanthanide series (from Ce³⁺ to Lu³⁺) was systematically tested to create heavy atom derivatives in the P4-P6 structure [139], as lanthanides
proved useful in the determination of the first tRNA [119,132] and hammerhead ribozyme crystal structures [156]. Diffraction experiments on lanthanide-soaked crystals focused on Sm\(^{3+}\) as a potential heavy-atom derivative for two reasons. First, the unit-cell dimensions changed as a function of decreasing ionic radius for lanthanides in the series from Lu\(^{3+}\) to Sm\(^{3+}\) after which they remained constant (Sm\(^{3+}\) to Ce\(^{3+}\)). Second, the mosaic spread of the diffraction pattern increased as a function of increasing ionic radius for all lanthanides except for Sm\(^{3+}\), for which the mosaicity was only slightly worse than that of native crystals. Yet, due to poor isomorphism and weak diffraction, the samarium derivative was abandoned in favor of the osmium hexammine derivative [139].

Lanthanide ions such as samarium bind specifically between phosphate oxygens in an adenosine-rich corkscrew structure at the junction of three helices by displacing a magnesium ion in this region; this ion is critical to the folding of the entire P4-P6 domain [139]. In order to find crystallization conditions that included metal cations that were able to generate anomalous scattering derivatives for a large RNase P fragment, the La\(^{3+}\), Pr\(^{3+}\), Sm\(^{3+}\),

---

Figure 6. Binding of twelve [Ir(NH\(_3\))\(_6\)]\(^{3+}\) cations to a flavin mononucleotide riboswitch inferred from a crystal structure at \(\approx 3.2\)\(\) Å resolution. Reproduced by permission from [155]; copyright (2009).
Gd$^{3+}$, Dy$^{3+}$, Yb$^{3+}$, and Lu$^{3+}$ lanthanides were tested. From this series, only Sm$^{3+}$, Gd$^{3+}$, and Yb$^{3+}$ produced anomalously diffracting derivatives where lanthanide cations bound solely to phosphate oxygen atoms [138]. In a SRP fragment, a Lu$^{3+}$ cation was found to bridge two phosphate groups [157]. It seems that lanthanide cations do not form direct coordination complexes to bases but bind exclusively to backbone phosphate oxygens, in agreement with earlier observations dealing with Pr$^{3+}$, Eu$^{3+}$, Gd$^{3+}$, Tb$^{3+}$, and Lu$^{3+}$ cations. These cations were found to bind to the same sites as Sm$^{3+}$ [119]. Regarding catalytic reaction rates, the addition of small quantities of Nd$^{3+}$ in the presence of Pb$^{2+}$ increased significantly the yield of the RNA cleavage reaction in a leadzyme, although other rare earth ions or divalent ions did not promote the reaction suggesting that subtle recognition phenomena are at play [158].

It was proposed that terbium inhibits the hammerhead ribozyme’s catalytic activity by competing with Mg$^{2+}$ cations leading to the characterization of three Tb$^{3+}$ binding sites that displayed, like almost all other metal ions bound minimal hammerhead ribozyme structures, occupancies below 0.5 [159–161]. These Tb$^{3+}$ ions are located in close proximity to electropositive and are reminiscent of SO$_4^{2-}$ binding sites. They display coordination distances of 3.8–4.0 Å compatible with sulfate binding and suggesting a possible error in the assignment of these electronic densities [162]. Moreover, lanthanides are expected to coordinate to phosphate oxygens than to base atoms. Note that the average “lanthanide . . . Ow” coordination distance is around 2.4 Å.

Unfortunately, most of these lanthanide atom derivatives have not been made public (only three nucleic acid structures containing terbium, ytterbium or lutecium cations are archived in the PDB) probably because of the modest resolution of these derivatives [58] (Figure 7).

7. OTHER TRIVALENT AND TETRAVALENT CATIONS

7.1. Gold Cations

Au$^{3+}$ is sometimes included in soaking experiments in order to produce anomalous diffracting derivatives [138]. Soaking experiments of a DIS subtype revealed the formation of an intriguing base pair: The Au$^{3+}$ cation induced a deprotonation of a G(N1) atom leading to a gold mediated G=C pair [54] (see also Sections 5.4.3 and 5.4.6 devoted to the Cu$^{2+}$ and the Hg$^{2+}$ cations). This very unique base pair is emblematic of the difficulty of predicting cation binding modes, as some can be induced by soaking experiments themselves (Figure 8). Note that in tRNA$^{\text{Asp}}$ Au$^{3+}$ was found to interact with more common binding sites [6].

Met. Ions Life Sci. 2011, 9, 1–35
7.2. Platinum Cations

Pt$^{4+}$ cations provide anomalous signals. They seem to have a good affinity for 5’ApC3’ sites as revealed by a DIS crystal structure [54] and early structures of tRNA$^{\text{Phe}}$ [119,163]. They were also tentatively located close to a G(N7) atom [119]. Hence, they seem to bind exclusively to nucleobases with an octahedral coordination and associated binding distances exceeding 2.5 Å.

8. ANIONS

A statistical survey of nucleic acid structures [162] established that anions can bind directly to nucleotides despite the negative charges carried by the...
polyanionic RNA and DNA molecules. This survey led to a map of anion binding sites to nucleotides. Among the possible anions, $\text{SO}_4^{2-}$, $\text{SeO}_4^{2-}$ [164], and acetate anions ($\text{CH}_3\text{CO}_2$) [165] were found to bind to the same guanine Watson-Crick and Hoogsteen sites as $\text{Cl}^-$ anions. It is worth noting that some metals are found in their anionic rather than cationic form. For instance, vanadium and tungsten atoms form vanadate ($\text{VO}_2^-$) and tungstate ($\text{WO}_4^{2-}$) anions that are iso-structural with $\text{SO}_4^{2-}$, $\text{SeO}_4^{2-}$, and $\text{PO}_4^{2-}$ anions.

The presence of well localized electron densities around nucleic acid structures has generated some significant assignment errors where metal cations such as $\text{Mg}^{2+}$ instead of anions such as $\text{Cl}^-$ or $\text{SO}_4^{2-}$ were placed incorrectly in the vicinity of nucleic acid electropositive atoms [162].

9. SUBJECTIVITY IN THE STRUCTURE DETERMINATION PROCESS

At the end of this analysis, it is important to recall that a certain level of subjectivity is inherent in the structure determination process and thus it is likely that some errors are present in macromolecular crystal structures. If uncorrected, these errors become a part of databases and statistical surveys. Consequently, reports based on locally deficient crystal structures, integrate some of these errors, like errors in the assignment of electron densities to water, cations or anions. Many of these potential sources of errors have been reviewed in the following articles [162,166–170]. Thus, there is still a need for revising existing structural databases, and for the development of efficient validation techniques. The SwS web service is one such validation tool [21].

Paraphrasing Georg Cristoph Lichtenberg (1742-1799) [171], it is worth remembering that, “crystallographers (scientists in general) are ‘creative’ and manage to produce new types of errors with regularity”.

10. SUMMARY

A significant number of crystal structures that include metal cations have been solved, and deposited into the PDB. These structures provide important data on water and ion binding sites [172]. Many of these metal-RNA interactions have been further investigated through biochemical and biophysical studies that help us to better understand the role played by mono- and multivalent cations in maintaining the structure and the activity of RNA systems.
The specific roles played by the diffuse ionic atmosphere surrounding nucleic acids are still difficult to address despite numerous studies devoted to this topic [92,173]. With the advent of more sophisticated detection techniques, it appears that a significant part of the “diffuse ionic cloud” is rather localized, and that very specific binding sites are associated with RNA motifs such as, for example, the AA platform for K\(^{+}\) cations. It is quite clear that much of the progress in understanding specific ion binding to RNA lies ahead of us. In the meantime, experiments on model systems such as nucleosides, nucleotides [6,94], and small RNAs [95] are still necessary to understand the effect of metal ions on RNA and vice versa.

While a large variety of ions are reported to associate with nucleic acids (close to 30 ions interacting with RNA and DNA are discussed in this review), only two of them, K\(^{+}\) and Mg\(^{2+}\), are physiologically relevant and interact directly with RNA in vivo (although rare associations with other ions cannot be fully excluded). All the other ions are used as crystallographic probes in order to provide heavy atom derivatives for solving crystal structures or for specific spectroscopic purposes, with the most common use being associated with specific paramagnetic properties.

Monovalent cations are easier to dehydrate and can therefore intrude into the first hydration shell of RNA nucleotides. They interact essentially with the deep groove G(N7/O6) atoms and less with other electronegative atoms. They are also seen close to the anionic oxygen atoms of phosphate groups in turn regions or participating in crystal packing interactions. Monovalent cations that do not directly bind to nucleic acids are difficult to detect. Even though various monovalent ions can be used, they are not interchangeable – for example, Tl\(^{+}\) is considered to be the best mimic for K\(^{+}\) (beside Na\(^{+}\)) but it does not always occupy the same binding sites or, if it is the case, coordination details may change. One interesting use of non-physiological monovalent cations is related to the high affinity of Cs\(^{+}\) ions for G·U containing motifs leading to targeted insertion of these base pairs into large RNAs for solving phase issues.

Probably the most important and best characterized metal-RNA interactions are those involving Mg\(^{2+}\) cations. Again, the G(N7/O6) binding sites seem to be preferred along with anionic oxygen atoms. Single dehydration of the hexahydrated Mg\(^{2+}\) cations is observed more often than multiple dehydration. Mg\(^{2+}\) seems to be a part (along with K\(^{+}\) cations) of several specific metal binding motifs that have to be more thoroughly characterized. Fully hydrated Mg\(^{2+}\) cations are often observed in the electronegative major groove of GpG steps.

Since Mg\(^{2+}\) cations are difficult to distinguish from water molecules in crystal structures, soaking experiments with heavier cations are often utilized.
The best mimic for Mg\(^{2+}\) is Mn\(^{2+}\) that displays anomalous properties. Yet, sufficient evidence exists that magnesium cannot be perfectly substituted by other ions. When Mg\(^{2+}\) is substituted by other cations, such as for example Mn\(^{2+}\), Ca\(^{2+}\), Sr\(^{2+}\), Co\(^{2+}\), Ni\(^{2+}\), Zn\(^{2+}\), Cd\(^{2+}\), or [Co(NH\(_3\))\(_6\)]\(^{3+}\) to study general ion binding properties or to locate a particular ion binding site, the results need to be interpreted carefully due to differences in coordination chemistry and binding properties of these cations. Many of these transition metals bind exclusively to G(N7) atoms. Such substitutions can lead to incorrect conclusions regarding the RNA structure or the role of the associated metal ions [94].

Lanthanides generally bind to phosphate oxygen atoms and not to electronegative nucleobase atoms. Their use is generally limited to obtaining heavy atom derivatives since they quite often generate crystal structures of low resolution and/or deform the crystal cell in soaking experiments.

In conclusion, it is highly desirable that crystallization and other biophysical experiments focus more on the use of the two biologically relevant K\(^+\) and Mg\(^{2+}\) ions and limit the use of Na\(^+\) cations in crystallization buffers. The detection of cations in the vicinity of nucleic acids remains a challenge that is complicated by partial occupancy issues. Indeed, anomalous diffraction techniques often provide occupancies below one for the larger monovalent cations (K\(^+\), Rb\(^+\), Cs\(^+\), Tl\(^+\)) suggesting that many ion-binding sites are alternatively occupied by monovalent cations, divalent cations and water molecules.

**ACKNOWLEDGMENTS**

Neena Grover is funded by a grant from the National Science Foundation MCB-0950582.

**ABBREVIATIONS AND DEFINITIONS**

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>ATP</td>
<td>adenosine 5'-triphosphate</td>
</tr>
<tr>
<td>CSD</td>
<td>Cambridge Structural Database</td>
</tr>
<tr>
<td>DIS</td>
<td>HIV dimerization initiation site</td>
</tr>
<tr>
<td>EPR</td>
<td>electron paramagnetic resonance</td>
</tr>
<tr>
<td>FRET</td>
<td>fluorescence resonance energy transfer</td>
</tr>
<tr>
<td>HDV</td>
<td>hepatitis delta virus ribozyme</td>
</tr>
<tr>
<td>HIV</td>
<td>human immunodeficiency virus</td>
</tr>
<tr>
<td>MAD</td>
<td>multiple anomalous diffraction</td>
</tr>
</tbody>
</table>

*Met. Ions Life Sci.*, 2011, 9, 1–35
MeRNA metal binding to RNA database
mRNA messenger RNA
NMR nuclear magnetic resonance
Ow oxygen of water
PDB Protein Data Bank
SAD single anomalous diffraction
SRP signal recognition particle
SwS solvation web service for nucleic acids
TAR trans-activation region
TEB transient electric birefringence
tRNA transfer RNA
VS virus satellite ribozyme

REFERENCES


*Met. Ions Life Sci.* 2011, 9, 1–35

Met. Ions Life Sci. 2011, 9, 1–35